19e Colloque International de l'ADMEE-Europe

L'évaluation au 21e siècle : Vers de nouvelles formes, modélisations et pratiques de l'évaluation ?

11-13 septembre 2006, Luxembourg

Résumés des ateliers

	Titre de la communication	Auteur(s)
Atelier 1 : Animateur : Dieudonné Leclercq	COMMENT EVALUER L'ENSEIGNEMENT DES CONCEPTS (AVEC LES TAXONOMIES DE BLOOM ET DE GAGNE ?	Sylviane Bachy
	L'AUTOÉVALUATION DE LA VOLITION ET DE LA PERCEPTION DU SENS	A. La Marca, G. Zanniello & G. Cappuccio
	L'EVOLUTION DES QCM ET CRITERES DE QUALITE ETIC –PRAD	Dieudonné Leclercq
Atelier 2 : Animateur : Romain Martin	DE LA NECESSITE D'UN DOUBLE REGARD. LE CAS DE L'EVALUATION D'UN MOYEN D'ENSEIGNEMENT DES MATHEMATIQUES EN SUISSE ROMANDE.	Chantal Tièche Christinat
	DEVELOPPEMENT D'UN OUTIL POUR LA CONSTRUCTION ET L'EVALUATION DE STRATEGIES DE RESOLUTION DE PROBLEMES EN MATHEMATIQUES A L'AIDE DU TESTING ASSISTE PARORDINATEUR.	Gilbert Busana & Romain Martin
	L'ÉVALUATION EN MATHÉMATIQUE : À LA RECHERCHE D'UNE CONCEPTUALISATION DE L'ACTIVITÉ ÉVALUATIVE DES ENSEIGNANTS	Maria Palmira Alves, José António Fernandes & Eusébio André Machado
Atelier 3 : Animateur : Jean-Guy Blais	PRECODEM, UN EXEMPLE DE TAO MIS AU SERVICE DE L'EMPLOI	Jadoul Raynald & Mizohata Sachié
	EFFETS DE L'INTRODUCTION D'UN DISPOSITIF INFORMATISE D'EEE DANS UNE UNIVERSITE FRANCAISE.	Nathalie Younes
	PRINCIPES ET OUTILS DE L'EVALUATION ACADEMIQUE DES COMPETENCES EN FORMATION INITIALE D'INFIRMIER-IERES	Brigitte Kampel & Salomé Anderfuhren Fiaux
Atelier 4 : Animateur : Gérard Figari	VAE EN LIGNE ET PRATIQUES REFLEXIVES : LES NOUVELLES PRATIQUES DE L'EVALUATION.	Padma Emilie Gouiller & Jacques Serizel
	REPRESENTATIONS ET TRADUCTIONS DANS LA VALIDATION DES ACQUIS DE L'EXPERIENCE EN FRANCE	Mireille Pouget
	LA PLACE DES VARIABLES D'EVALUATION DANS LA PLANIFICATION ET LA GESTION DE PROCESSUS D'ENSEIGNEMENTAPPRENTISSAGE A L'ECOLE PRIMAIRE	Teresa Grange Sergi
Atelier 5 : Animateur : Léopold Paquay	EVALUER LES COMPETENCES ? LES RESSOURCES ET LES CONTRAINTES DU PORTFOLIO NUMERIQUE.	Baillat Gilles & Connan Pierre-Yves
	LE PORTFOLIO AU SERVICE DES COMPÉTENCES TRANSVERSALES	Hervé Bernard & Joëlle Vlassis
	UN PORTFOLIO CERTIFICATIF, CLE DE VOÛTE D'UN PROGRAMME DE FORMATION PEDAGOGIQUE OBLIGATOIRE DES ENSEIGNANTS DE L'ENSEIGNEMENT SUPÉRIEUR. VERS UNE ALLIANCE ENTRE UNE FONCTION FORMATRICE INTÉGRATIVE ET UNE FONCTION CERTIFICATIVE ?	Léopold Paquay & Catherine Van Nieuwenhoven

Atelier 6 : Animateur : Marc Demeuse	EQUITE ET EFFICACITE : QUELQUES PROBLEMES DE MESURES DANS UN CONTEXTE D'ANALYSE LOCALISE DE PISA	Jean Moreau & Christian Nidegger
	BASES DE DONNÉES INTERNATIONALES SUR LA QUALITÉ DE L'ÉDUCATION (1964/2005)	Nadir Altinok
	ROLE DU FACTEUR LINGUISTIQUE DANS LES PRATIQUES PEDAGOGIQUES RAPPORTÉES AU PROGRAMME INTERNATIONNAL DE RECHERCHE EN LECTURE SCOLAIRE (PIRLS)	Marielle Simon & Catherine Turcotte
Atelier 7 : Animateur : Christian Monseur	DATA ENVELOPMENT ANALYSIS (DEA) : UNE NOUVELLE MÉTHODE D'ÉVALUATION DE LA PERFORMANCE SCOLAIRE.	Djily Diagne
	UTILISATION D'UN MODELE DE REPONSE A L'ITEM POUR LA VALIDATION D'UNE ECHELLE DES ATTITUDES PROFESSIONNELLES DES MEDECINS GENERALISTES	Etienne Le Bihan, Cédric Baumann, François Alla & Michèle Baumann
	UNE ALTERNATIVE EXPLORATOIRE A L'AJUSTEMENT PONDÉRAL POUR COMPENSER LA NON-RÉPONSE DES ÉLÈVES.	Christian Monseur
Atelier 8 : Animateur : Jean-Marc Braibant	AVIS PEDAGOGIQUE, EVALUATION DES ENSEIGNEMENTS : QUELLES PRATIQUES, QUELS ENJEUX ?	Nathalie Delausnay & Michel Sylin
	MAIS POURQUOI LES ÉVALUATEURS N'UTILISENT-ILS PAS LES GRILLES D'ÉVALUATION ?	Marie Noëlle De Theux, Jean Marc Braibant, Benoit Raucent, Pascale Wouters & Thérèse Bouvy
	COMMENT EVALUER EQUITABLEMENT LES PROJETS ? L'EVOLUTION DES CONCEPTIONS DE L'EQUITE AU SEIN D'UNE GRANDE ECOLE D'INGENIEURS.	Caroline Verzat
Atelier 9 : Animateur : Palmira Alves	BESOINS ET QUALIFICATIONS DANS LE CONTEXTE DE LA FORMATION UNIVERSITAIRE DE L'ÉDUCATEUR/ ÉDUCATRICE GRADUÉ(E) ET DANS LE DOMAINE SOCIO-ÉDUCATIF AU LUXEMBOURG	Arthur Limbach-Reich, Ulla Peters & Regina Gaitsch
	L'ÉVALUATION DE LA PORTÉE DE DISPOSITIFS INNOVANTS DE FORMATION CONTINUE D'ENSEIGNANTS DU PRIMAIRE ET DU SECONDAIRE SUR LE DÉVELOPPEMENT, LA MOBILISATION ET LE TRANSFERTS DE LEURS SAVOIRS ET COMPÉTENCES PROFESSIONNELLES	Annie Presseau & Stéphane Martineau
	ETAT DES LIEUX DES PRATIQUES FORMELLES ET INFORMELLES DES ACTEURS DE LA FORMATION.	Leclerc Claude
Atelier 10 : Animateur : Nathalie Droyer	REDEFINIR LES COMPETENCES ACQUISES PAR LES ELEVES A L'ECOLE PRIMAIRE POUR MIEUX COMPRENDRE LA STRUCTURE DES APPRENTISSAGES	Sophie Morlaix
	ÉVALUATION ET AUTOÉVALUATION DES COMPÉTENCES COMME MOYEN DE QUALIFICATION DE L'INSTRUCTION A DISTANCE.	Gaetano Domenici

	L'ÉVALUATION DES APPRENTISSAGES DANS UN ENVIRONNEMENT SCOLAIRE MARQUÉ PAR LA DIFFÉRENCIATION PÉDAGOGIQUE ET L'ENSEIGNEMENT EN FAVEUR DU TRANSFERT : QUELQUES RÉSULTATS D'UNE RECHERCHE-ACTION MENÉE AVEC DES ENSEIGNANTS DU PRIMAIRE.	Annie Presseau & Stéphane Martineau
Atelier 11 : Animateur : Claude Houssemand	PARASCHOOL, EVALUATION D'UNE EXPERIENCE D'ENSEIGNEMENT ASSISTE PAR ORDINATEUR AU SECONDAIRE 1 : QUELS ENJEUX ? QUELLES IMPLICATIONS POUR L'ENSEIGNEMENT ?	Anne-Marie Broi
	UN OUTIL D'EVALUATION DU PROFIL ET DES COMPETENCES DES DEMANDEURS D'EMPLOI	Claude Houssemand, Romain Martin & Raymond Meyers
	EVALUATION D'UN LOGICIEL D'(AUTO)APPRENTISSAGE DES LANGUES	Audrey Thonard
Atelier 12 : Animateur : Lucie Mottier Lopez	EVALUER L'EFFICACITE PEDAGOGIQUE D'UNE FORMATION OU D'UN COURS A L'AIDE D'UN OUTIL D'AUTOEVALUATION	François-Marie Gerard, Jean-Marc Braibant & Thérèse Bouvy
	EVALUER LA QUALITE DE L'ENSEIGNEMENT SUPERIEUR : UNE CULTURE PROFESSIONNELLE EMERGENTE	Alain Hertay & Pierre Stegen
	EVALUER AUTREMENT LA CAPACITE D'ANALYSER SA PRATIQUE D'ENSEIGNANT AU SECONDAIRE SUPERIEUR SOUS L'ANGLE DE LA PSYCHOLOGIE EDUCATIONNELLE	Delcomminette Séverine, Leclercq Dieudonné & Sente Isabelle
Atelier 13 : Animateur : Marlyse Pillonel	CONSTRUCTION ET UTILISATION D'UN OUTIL D'ANALYSE DES RESULTATS A MI-PARCOURS POUR PROMOUVOIR LA REFLEXION METACOGNITIVE DE L'ETUDIANT DIAGNOSTIQUE EN DIFFICULTE EN PREMIERE ANNEE DE L'ENSEIGNEMENT SUPERIEUR	Didier Salmon
	PROFILS DE PERFORMANCES ATTEINTS EN ORTHOGRAPHE ALLEMANDE ET FRANÇAISE : LES ENFANTS A DIFFICULTES	Sylvie Bodé & Joyce Serres
	LA VALIDATION DES ACQUIS DE L'EXPÉRIENCE : VERS UN RENOUVELLEMENT DES REGARDS SUR L'AUTO-ÉVALUATION	Borhène Chakroun
Atelier 14 : Animateur : Marthe Hurteau	L'AUTORITÉ DU "BON PROFESSEUR": EVALUATION DES ATTITUDES ET COMPORTEMENTS FAVORABLES AU RESPECT DES REGLES	Françoise Welbes & Robert A.P. Reuter
	LES PROCESSUS DE PENSEE DES ENSEIGNANTS DURANT LA PLANIFICATION ET L'IMPLEMENTATION DE LEUR ENSEIGNEMENT	Philippe Wanlin & Cindy Bodeux
	UNE CLASSE COOPERATIVE VERTICALE DANS L'ENSEIGNEMENT SUPERIEUR : EVALUATION DES IMPACTS DU DISPOSITIF AUPRES DE FUTURS ENSEIGNANTS.	Voz Grégory
	ET SI ON PARLAIT DE LA PERTINENCE DE L'EVALUATION ?	Carette Vincent

Atelier 15 : Animateur : André Machado	DEUX DECADES POUR BROUILLER LES FRONTIERES ENTRE L'EVALUATION FORMATIVE ET L'EVALUATION SOMMATIVE	Aldina Silveira Lobo
	L'ÉVALUATION FORMATIVE : VERS UNE CLARIFICATION CONCEPTUELLE	Julie Desjardins, Colette Deaudelin & Catherine Allen
Atelier 16 : Animateur : Pascal Bressoux	RÉDUCTION DE LA TAILLE DES CLASSES ET IMPLICATION DES ÉLÈVES DANS LES TÂCHES SCOLAIRES	Pascal Bressoux & Lima Laurent
	L'EVALUATION D'UN DISPOSITIF DE PROMOTION DE LA REUSSITE EN PREMIERE ANNEE D'ETUDES UNIVERSITAIRES EN Belgique FRANCOPHONE : DE NOUVELLES PISTES ENVISAGEES.	Alice Salcin, Dorothée Baillet & José-Luis Wolfs
	DISPOSITIF DE FORMATION-EVALUATION SOCIOCONSTRUCTIVISTE CENTRE SUR LES COMPETENCES, EN FORMATION INITIALE D'INFIRMIER-IERES	Salomé Anderfuhren Fiaux & Brigitte Kampel
Atelier 17 : Animateur : Matthis Behrens	LES REPRÉSENTATIONS SOCIALES DES 36 INTERVENANTS DIRECTS ET INDIRECTS DU SYSTÈME SCOLAIRE SUR LA GESTION DE L'HÉTÉROGÉNÉITÉ DANS LE SYSTÈME SCOLAIRE LUXEMBOURGEOIS: ANALYSE DU CONTENU DE DONNÉES ISSUES D'UN DISPOSITIF DE FOCUS GROUP.	Christian Meyers, Georges Barthel, Danielle Zerbato & Romain Martin
	GENERALISABILITE ET COMPETENCE SOCIALE : UNE ETUDE CHEZ DES ELEVES QUEBECOIS	Marie-Hélène Hébert, Martin Gendron, Éric Frenette, Égide Royer & Céline Morand
	COMPRÉHENSION DU BULLETIN SCOLAIRE DANS UN CONTEXTE D'IMPLANTATION D'UN NOUVEAU PROGRAMME DE FORMATION AU QUÉBEC : ANALYSE DES REPRÉSENTATIONS DES PARENTS DU 1ER DU 3E CYCLE DU PRIMAIRE D'UNE COMMISSION SCOLAIRE	Micheline-Joanne Durand & Nathalie Loye
Atelier 18 : Animateur : Walter Tessaro	PRATIQUES D'ÉVALUATION DANS LE CADRE D'UN FORUM : LE RETOUR DE L'ENSEIGNEMENT MUTUEL ?	Loïc Clavier
	EVALUATION DE L'INTÉRÊT POUR LES BRANCHES SCIENTIFIQUES PRÉSENTÉES DANS UN CONTEXTE MASCULIN ET DANS UN CONTEXTE FÉMININ : CONTENUS NON-PARALLÉLISÉES ET CONTENUS PARALLÉLISÉES	Kerger Sylvie
	POSITIONNEMENT DES PARENTS LUXEMBOURGEOIS FACE A PLUSIEURS SCENARIOS D'EVOLUTION POSSIBLE DU SYSTEME EDUCATIF	C. Dierendonck, D. Poncelet & R. Martin
	METHODOLOGIE D'EVALUATION DES COMPETENCES DE CONDUITE D'UN GROUPE D'ANALYSE DE PRATIQUES	Richard Étienne

Atelier pratique 1 : Dresser le portrait des apprenants pour améliorer leurs apprentissages	DRESSER LE PORTRAIT DES APPRENANTS POUR AMELIORER LEURS APPRENTISSAGES : UNE INTRODUCTION AUX MODELES PSYCHOMETRIQUES VISANT LE DIAGNOSTIC COGNITIF EN EDUCATION	André Rupp & Nathalie Loye
Atelier pratique 2 : Le Testing Adaptatif par Ordinateur : outil d'évaluation formative	LE TESTING ADAPTATIF PAR ORDINATEUR : OUTIL D'EVALUATION FORMATIVE	Patrick Plichart & Gilbert Busana
Atelier pratique 4 : Application Web pour la création et l'administration en ligne de tests adaptatifs	APPLICATION WEB POUR LA CREATION ET L'ADMINISTRATION EN LIGNE DE TESTS ADAPTATIFS	Nancy Boiteau

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

COMMENT EVALUER L'ENSEIGNEMENT DES CONCEPTS (AVEC LES TAXONOMIES DE BLOOM ET DE GAGNE ?

Sylviane Bachy, Université Catholique de Louvain, Institut Libre Marie Haps*

Mots clés : concept, dispositif, taxonomie de Gagné et de Bloom.

Résumé :
L'apprentissage de la phonétique par les étudiants de BAC 1 en Haute Ecole est souvent rendu austère par le nombre de nouveaux concepts qui demandent une haute précision notamment pour le diagnostic logopédique. La taille des groupes (250 étudiants) complique souvent l'enseignement des concepts. Aussi en analysant nos examens de 2004, nous n'avons pas été étonnées de constater que les questions de définition et de classification qui évaluent les concepts étaient moins bien réussies par les étudiants que les autres types d'évaluation (choix multiples, schémas à compléter, etc). Une analyse de la qualité avec le modèle intégré de Lebrun (2005) : « un modèle pour réfléchir à la qualité d'un dispositif pédagogique », a ciblé les contenus à adapter. Sur base des taxonomies de Bloom, revisitée par Anderson et Krathwohl (2001) et de Gagné (dans Rieunier 2000) des pauses méthodologiques ont été introduites. Pour évaluer l'impact du nouveau dispositif nous proposons une évaluation quantitative (résultats aux examens) et qualitative (avec un questionnaire d'évaluation sur l'activité d'enseignement). L'ensemble des résultats nous permet de croire que le nouveau dispositif favorise l'apprentissage des concepts

Introduction

Enseigner un concept représente 50% de l'activité d'enseignement. Reconnaître les concepts et connaître les stratégies pour les enseigner permet à l'enseignant de **concevoir un dispositif** d'enseignement adapté et ainsi de préparer ses leçons plus rapidement (Rieunier, 2001).

Trois façons d'évaluer le dispositif basé sur l'apprentissage des concepts sont présentées. L'évaluation porte sur une analyse de la qualité du dispositif, une analyse quantitative et une analyse qualitative.

L'évaluation de la qualité de l'ancien dispositif avec le modèle intégré de Lebrun (2005) a ciblé de nouveaux objectifs à atteindre en modifiant le contenu des

* bachy@rom.ucl.ac.be Université Catholique de Louvain, place Blaise Pascal 1, 1348 Louvain-La-Neuve (Belgique), Institut Libre Marie Haps, 11 rue d'Arlon 1050 Bruxelles

informations présentées aux cours. Les taxonomies de Gagné (Rieunier, 2000) et de Bloom (Anderson et Krathwohl, 2001) ont été utilisées comme grille de lecture pour définir les types de connaissances et les types de processus cognitifs mis en jeu. Les étudiants ont appris a reconnaître les concepts et à développer des stratégies pour les apprendre. Pour mesurer l'impact de ces pauses méthodologiques nous proposons une évaluation quantitative et qualitative.

Ce document est divisé en quatre parties. La section 1 expose notre contexte d'enseignement et définit le concept de concept. La section 2 analyse l'ancien dispositif avec le modèle intégré de Lebrun et décrit les adaptations réalisées sur base du tableau des variables de Raynal et Rieunier (Rieunier 2000). Les sections 3 et 4 évaluent le nouveau dispositif.

1. Contexte et concept

1.1 Le contexte

Nous enseignons le cours de phonétique dans la Haute Ecole Léonard de Vinci, à l'Institut Libre Marie Haps dans le département de logopédie et d'audiologie. Le cours de phonétique est un cours de base dispensé le premier semestre de la première année (BAC1).

Parmi les types de connaissances évalués lors de l'examen en juin 2004 les deux questions concernant les concepts (questions de classification et de définition) ont été moins bien réussies par les étudiants.

1.2 Question de concept

Un « concept » c'est l'association d'une étiquette avec un ensemble de propriétés (Barth, 1987) définissant une idée (Rieunier, 2000) qui reprend les informations les plus importantes (Vial, 2001). S'inscrivant sous forme de nœuds dans des réseaux sémantiques (de Vecchi et Carmona-Magnaldi, 1996), les concepts sont liés entre eux. Ils *servent* à catégoriser/classifier/définir les connaissances (Bloom dans Anderson et Krathwohl 2001) et donc à interpréter notre environnement comme un ensemble d'objets reconnaissables. Résultant d'une construction des connaissances de l'individu au gré des interactions avec son environnement, les concepts peuvent différer d'un individu à l'autre par des combinaisons différentes d'attributs (attributs

internes ou externes de Robert 2006, attributs essentiels et non essentiels de Barth, 1987).

Par exemple « la résonance » peut s'expliquer par l'observation du verre qui se brise lorsqu'il entre en résonance avec une fréquence très aigue, par la bouteille remplie d'eau dans laquelle on souffle et qui ne produit pas le même son si la quantité d'eau est modifiée, etc. Ces exemples observables permettent de tirer une série de propriétés : il faut une cavité (résonateur) qui a des caractéristiques physiques (volume, matière, etc), il y a une fréquence de résonance propre à l'objet qui entre en vibration lorsqu'elle rencontre une fréquence proche. La modification du résonateur entraîne une modification de la fréquence de résonance.

La résonance est le phénomène se produisant lorsqu'un résonateur est excité par une vibration à une fréquence donnée, appelée fréquence de résonance. Elle est liée aux caractéristiques physiques du résonateur (volume, forme, matière).

Grâce à d'autres exemples comme le pont de Takoma (pont qui se brise par effet de résonance avec le vent), la balançoire, etc l'observateur peut ajouter aux propriétés déjà définies que le phénomène de résonance, amplifie un mouvement (la vibration) lorsqu'il rencontre une fréquence proche de sa fréquence de résonance.

Ces exemples indiquent que *la fréquence de résonance d'un système est une fréquence naturelle de vibration à laquelle l'énergie donnée au système (« poussée » dans le cas de la balançoire, « onde sonore » pour le verre, « onde du vent » pour le pont) est amplifiée.*

Quand nous parlons, notre conduit vocal[1] peut être schématisé par une suite de tubes dont les fréquences de résonance nous en apprennent beaucoup sur les sons produits…

[1] Le conduit vocal est le terme qui désigne la cavité remplie d'air qui s'étant de la glotte (larynx) aux lèvres. Cette cavité n'est pas rigide, et sa forme peut-être modifiée par le mouvement d'articulateurs (la langue par exemple).

Le conduit vocal, du larynx aux lèvres (et aux narines) est une structure tridimensionnelle complexe (forme courbe). C'est la position des articulateurs qui va déterminer le mode de propagation de l'air et donc les sons produits. En parlant, l'individu transmet une série de fréquences à nos quatre résonateurs : le pharynx, la bouche, le nez et les lèvres. Ils ont des formes différentes que nous pouvons modifier avec la langue, le voile du palais, les lèvres et le maxillaire inférieur. La modification des cavités de résonance détermine les fréquences qui sont amplifiées. Celles-ci s'appèlent les formants. La différence entre un « i », un « a », un « u » s'explique par ce phénomène d'amplification de certaines fréquences par le passage du son dans les résonateurs.

2. Analyse de l'ancien dispositif et adaptations

Différents outils d'évaluation sont présentés dans cette partie. Dans le point 2.1 nous présentons le modèle intégré de Lebrun qui a servi de base à notre évaluation. Le point 2.2 reprend le sous-modèle de Biggs. Ce point introduit également les taxonomies de Bloom et de Gagné utilisées pour le nouveau dispositif. Le point 2.3 présente le sous-modèle de Barbier et Lesne. Le point 2.4 traite du pentagone d'apprentissage de Lebrun lui-même. Enfin le point 2.5 propose le nouveau dispositif avec le tableau des variables de Raynal et Rieunier (2000).

2.1 Présentation du modèle

Lebrun croise trois modèles (2005, p. 191) pour déterminer un dispositif de qualité:

- L'alignement constructiviste de Biggs (objectifs-méthodes-outils) (Lebrun 2005) que nous pouvons évaluer en analysant la cohérence des interactions entre ces trois objets,
- Les trois champs de Barbier et Lesne (Laloux, 2002) :le champ de la vie quotidienne, professionnelle et sociale (objectifs), le champ des phénomènes relatifs au cadre institutionnel des programmes de formation (méthodes) , le champ des activités pédagogiques (outils) qui situent la démarche,
- Et son propre modèle « le pentagone d'apprentissage » (IMAIP) Informations, Motivation, Activités, Interactions et Productions que nous évaluerons par la présence des cinq ingrédients pour un dispositif de qualité (Lebrun 2005).

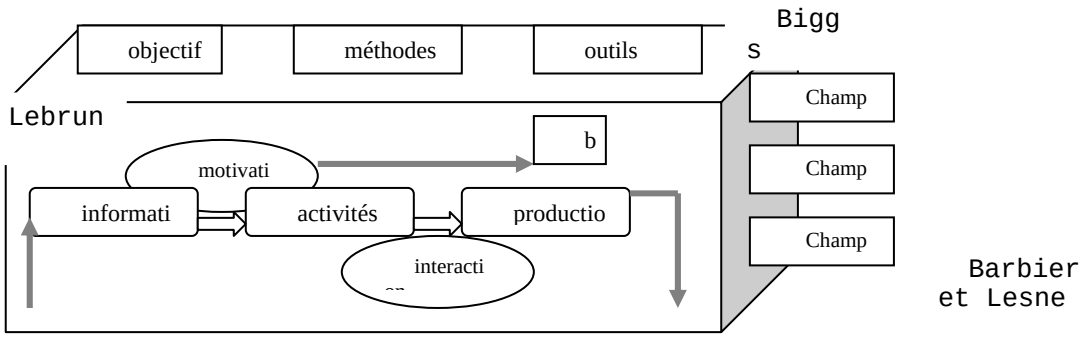


Figure 1 – Un modèle pour réfléchir à la qualité d'un dispositif pédagogique

2.2 L'alignement méthodes-objectifs-outils de Biggs

Les objectifs de cours en 2004 étaient essentiellement répartis en objectifs généraux et spécifiques

Un exemple **d'objectif** spécifique était formulé de la manière suivante :
« Après cette leçon de 2 heures, les étudiants de BAC1 en logopédie seront capables de définir la résonance ».

Les **méthodes** utilisées visaient à exposer, expérimenter et interroger les étudiants et les **outils** favorisaient les questions (questions-guides de TABA dans Rieunier 2000) ou les comparaisons (métaphore de de Vecchi et al, 1996).

Selon Biggs (Lebrun 2005), développer des compétences demande une cohérence entre les trois objets : objectifs, méthodes et outils.

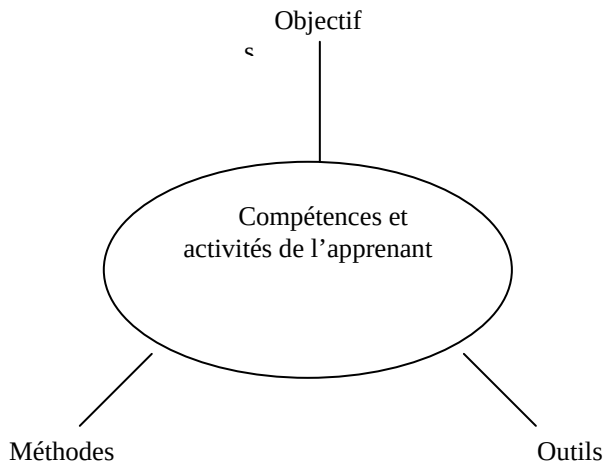


Figure 2 – Alignement constructiviste de Biggs par Lebrun (2005)

Si le dispositif de 2004 semble cohérent il ne met pas suffisamment l'accent sur ce qui est spécifique dans l'objectif de pouvoir « définir ». Développer des compétences spécifiquement pour l'apprentissage des concepts demande d'autres choix qui ne peuvent se faire que si on a précisé au préalable **les types de connaissances visées** (la taxonomie de Gagné) et **les types de processus cognitifs mis en jeu** (taxonomie de Bloom).

Le tableau ci-dessous reprend l'ensemble des deux taxonomies. Anderson et Krathwohl (2001) définissent une dimension qui concerne le processus cognitif (the cognitive process dimension), elle est issue des travaux de Bloom et une dimension qui concerne le type de connaissances (the knowledge dimension) proche des travaux de Gagné parce qu'on y reconnaît la distinction entre les connaissances des faits, des concepts (englobant les principes) et des procédures.

Tableau 1 - Taxonomie de Bloom revisitée simplifiée d'après le site http://coe.sdu.edu/eet/Articles/bloomrev/index.htm

The knowledge dimension	**The cognitive process dimension**					
	remember	understand	apply	analyse	evaluate	create
Factual knowledge						
Conceptual Knowledge						
Procedural Knowledge						
Meta-cognitive Knowledge						

Connaître les types de connaissances (faits, concepts, principes, procédures, stratégies) – dans ce cas-ci les concepts et les types de processus cognitifs (mémoriser, comprendre, appliquer, analyser, évaluer et créer) permet d'élaborer une stratégie d'apprentissage. Or, tout porte à croire que les étudiants deviennent plus compétents lorsqu'ils deviennent conscients des stratégies qu'ils utilisent (Rieunier, 2000 ; de Vecchi & Carmona-Magnaldi, 1996 ; Barth, 1987 ; Taba cité par Rieunier,

2001). Plus les étudiants sont conscients de leur apprentissage, plus facilement ils peuvent créer des stratégies et des modèles (Bachy 2006) pour mobiliser leurs compétences. C'est la raison pour laquelle nous avons décidé de faire des pauses méthodologiques.

Les pauses méthodologiques concernent essentiellement la prise de conscience de son propre fonctionnement intellectuel (métacognition) par l'étudiant. Les connaissances métacognitives sont les stratégies d'apprentissages et l'analyse des systèmes de traitement de l'information que tout individu met en œuvre pour apprendre. Pour les concepts, il s'agit de préciser aux étudiants les types de processus cognitifs mis en jeu, les méthodes (interrogatives, expérimentales, expositives), les démarches (inductives et déductives), les outils (métaphores, etc).

L'introduction des **pauses méthodologiques** répond à ce type de connaissances que nous n'exploitions pas auparavant. D'après Raynal et Rieunier, (Rieunier 2000) elles devraient améliorer de façon significative l'aptitude à traiter l'information. C'est ce que nous verrons grâce à l'analyse quantitative des résultats.

2.3 Les trois champs de Barbier et Lesne (1986 dans Lebrun 2005)

Schématiquement nous avons représenté les trois champs ci-dessous :

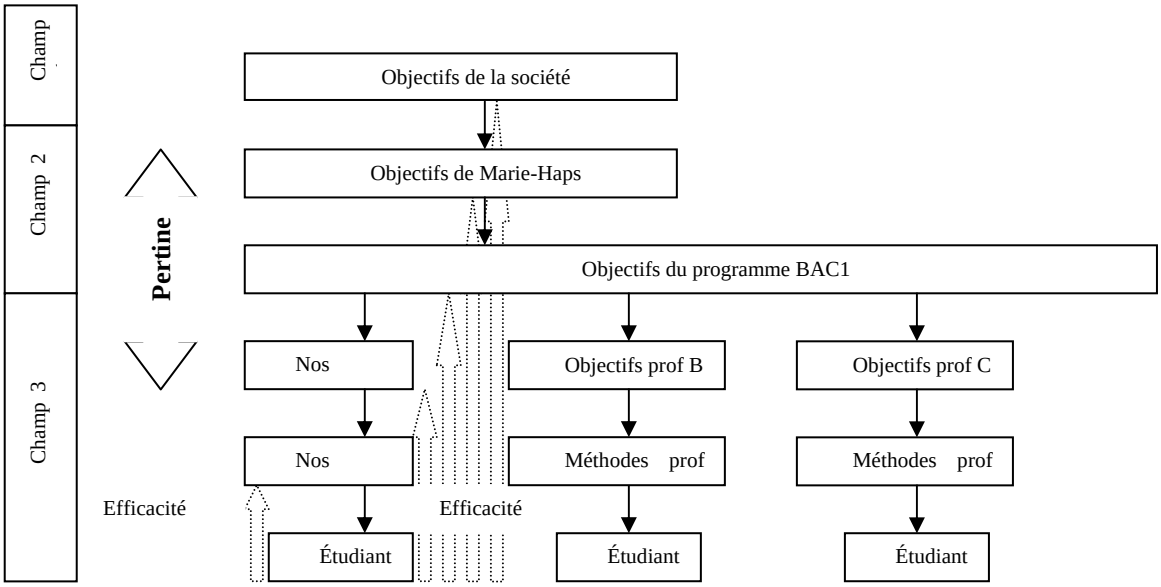


Figure 3 – Schéma des trois champs adapté de Barbier et Lesne par Laloux (2002)

Ce schéma représente les trois champs : champ 1 : les objectifs de la société, champ 2 les objectifs de l'Institut et du programme d'études et le champ 3 nos activités pédagogiques. Il faut un lien *pertinent* entre les champs. L'efficacité interne se situe dans le champ 3 : niveau du cours. L'efficacité externe représente le développement des compétences qui répondent à nos objectifs, aux objectifs du programme, aux objectifs de l'Institut et aux besoins de la société.

Le fait que les étudiants réussissent moins bien les questions concernant les concepts demande de reconsidérer l'efficacité interne de notre dispositif car sans doute que le transfert des connaissances lié à l'ancien dispositif n'a pas permis aux étudiants de répondre aux objectifs des champs 2 et 1. Les professeurs de pratique professionnelle (stage) ont notamment relevé certaines confusions entre des concepts ce qui confirmerait nos résultats et les liens entre ce qui se passe au cours et ce qui est transféré

.

L'introduction des pauses méthodologiques par rapport aux conceptions de Barbier et Lesne permettrait un meilleur transfert entre le cours et les activités sur le terrain. En effet, l'enseignant qui explique les stratégies, explique en fait les liens entre ses méthodes, les outils et les objectifs (champ 3) qu'il souhaite atteindre par rapport aux objectifs du programme (champ 2) et des besoins du terrain (champ 3).

2.4 L'IMAIP de Lebrun,

Le modèle IMAIP (informations, motivation, activités, interaction et production) permet d'évaluer le champ 3 de Barbier et Lesne : nos objectifs sont-ils atteints pour nous et pour les étudiants (efficacité interne) ?

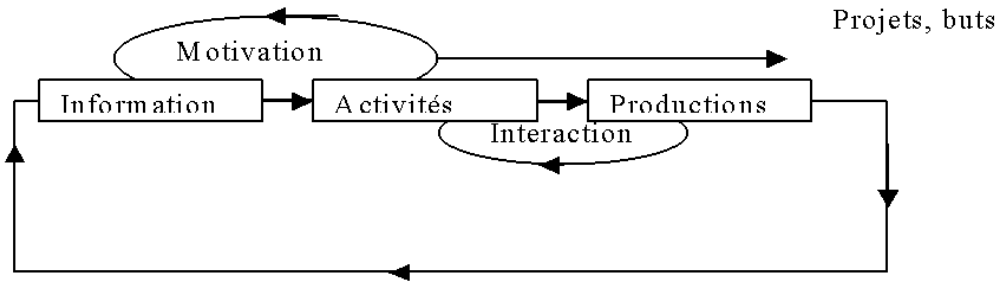


Figure 4 – Le pentagone d'apprentissage IMAIP (Lebrun 2005)

Il permet de réguler les ingrédients sur base des buts poursuivis : sont-ils atteints, sont-ils satisfaisants? Notre évaluation formative en fin de chaque cours dans une perspective d'une **régulation intentionnelle** (Perrenoud 1998) nous laissaient croire que notre dispositif était de qualité parce que nous obtenions un pourcentage élevé de bonne réponses (80%). La discussion porte dès lors sur ce qu'on entend par « qualité ».

L'analyse de l'évaluation certificative (présentée plus loin) nous a montré que notre ancien dispositif devait proposer d'autres types de contenu pour atteindre un nouvel objectif : apprendre à apprendre les concepts puisque les étudiants étaient moins performants pour ce type de connaissance et que nous y étions sensibles.

Le contenu a été analysé grâce aux taxonomies de Gagné et de Bloom. Après avoir identifié les concepts et les types de processus cognitifs, nous avons modifié l'*information* en insistant sur l'apprentissage de stratégies. Les autres « ingrédients » déjà présents en 2004 ont été maintenus.

2.5 Réflexions et adaptation

2.5.1 Réflexions

L'analyse de l'ancien dispositif par le modèle intégré de Lebrun permet les réflexions suivantes :

-Cohérence (Biggs) ne veut pas dire qualité : l'ancien dispositif était cohérent mais il ne répondait pas tout à fait à nos attentes qui n'étaient pas clairement formulées. La précision des objectifs grâce aux taxonomies de Bloom et de Gagné a permis l'élaboration de stratégies d'apprentissage spécifiques pour les concepts.

-Pertinence (Barbier et Lesne) ne veut pas dire qualité : l'ancien dispositif proposait des objectifs pertinents par rapport aux objectifs institutionnels et aux objectifs de la société. Toutefois, les compétences développées étaient insuffisantes pour l'apprentissage des concepts.

- IMAIP (Lebrun) ne veut pas dire qualité : l'ancien dispositif avait les cinq ingrédients : information, motivation, activité, interaction et production sans en avoir nécessairement le bon contenu. Préciser le type de connaissance et éplucher notre cours avec la taxonomie de Gagné en tête nous a permis de réajuster l'*information* donnée.

2.5.2 Adaptations

L'analyse du dispositif met en évidence les adaptations suivantes à réaliser pour préciser nos objectifs spécifiques :
- Analyser les types de connaissances (Gagné),
- Analyser les types de processus cognitifs (Bloom),
- Rendre conscient des stratégies (Rieunier, 2000; de Vecchi & Carmona-Magnaldi, 1996; Barth, 1987 ; Taba cité par Rieunier, 2001),
- Apprendre à enseigner les stratégies.

La grille de lecture proposée par Anderson et Krathwohl (2001) pour les taxonomies (présentée au tableau 1) se retrouve dans le tableau des variables de Raynal et Rieunier (1980 dans Rieunier, 2000, p. 156).

Tableau 2 – Tableau des variables de Raynal et Rieunier (1980 dans Rieunier, 2000)

Types d’apprentissage / Éléments de stratégies	faits	explications	concepts	principes	méthodes	stratégies	atti
Démarches	inductive	déductive	démonstrative	analogique	dialectique	proacti	
Méthodes	expositive		interrogative		active		
Techniques	Techniques d’animation : Questionnement Exemples Encouragement Consignes Synthèse, etc	exposé	démonstration	expérience par professeur	simulation	jeu de rôl	
		Expérience d’application	exercice	compagnonnage	enseignement programmé	expérienc élève	
Représentation du réel	Objet réel	photo	vidéo	maquette	dessin	schéma	symb
Support	papier	tableau à craie	tableau papier	transparents	diapositives	film	
Taille du groupe	travail individuel		travail à deux		travail en petit groupe	group	
Niveau de performance	citer	expliquer	classer	appliquer	inventer		

Cet outil permet d’analyser les stratégies d’enseignement et de concevoir des leçons mesurées aux travers des deux taxonomies : celle de Gagné (ligne du dessus) et celle de Bloom (ligne du dessous). Il répond donc aux questions suivantes : Comment enseigner les concepts ? Et comment concevoir une stratégie d’enseignement pour l’apprentissage des concepts ? En effet il permet de détailler la stratégie comme l’ensemble : démarche, méthode, technique, support et outils en fonction du type d’apprentissage et du type de processus cognitif visé.

Un exemple de leçon pour l’enseignement du concept « résonance » est repris ci-dessous dans le tableau des variables pour les **concepts**.

Tableau 3 – Tableau du dispositif pour le **concept** « résonance », groupe classe

Objectifs :	Après cette leçon de 2h, les étudiants de BAC1 seront capables de comprendre, analyser et évaluer le phénomène de résonance.
Motivation	Ce phénomène permet d'expliquer et d'évaluer des troubles phonétiques comme les accents, les imprécisions articulatoires…
Démarche : Arche c'est-à-dire déductive et inductive (Rieunier)	-Représentation : pour vous le phénomène de résonance représente quoi ? -Exemples de départ (flûte et bouteille) et contre exemples (pont) -définition -compréhension (corps humain)
Méthodes : Méthode expositive et interrogative recherche d'attributs (Barth)	Vidéo du verre qui se brise, pont de Takoma, sonagramme (Formants) Qu'est-ce que ces exemples ont en commun, ont de différent?
Technique : support	exposé, démonstration, expérience (Taba), avec vidéo, ordinateur, flûte, bouteille…
0Évaluation	Evaluation formative en fin de leçon : choix multiple à main levée Evaluation certificative en juin : par écrit : « expliquez le phénomène de résonance »

Les pauses méthodologiques visent l'explication de ce tableau auprès des étudiants.

3. Evaluation du dispositif sur base des résultats aux examens. Quel est l'impact de l'introduction des pauses méthodologiques ?

3.1 Précisions méthodologiques

Nous comparons deux tableaux de résultats issus de 2004 et de 2005 réalisés à partir des examens de deux groupes d'étudiants. Les deux groupes ont suivi le cours lors du premier semestre de leur première année. L'examen s'est déroulé en juin et comportait le même type de questions avec la même répartition des points. Les critères d'évaluation des questions touchant les concepts sont les mêmes. Nous

attendons que les étudiants expliquent certains concepts par des exemples, le contexte, et les attributs essentiels.

Les résultats sont à prendre avec une certaine précaution. En effet, les étudiants ne sont pas les mêmes et les questions, bien que calibrées de la même manière, sont également différentes.

3.2 Critères d'évaluation

Une amélioration aux questions 1 et 7 de l'examen de 2005 pourrait montrer un impact positif des pauses méthodologiques sur l'apprentissage des concepts.

3.3 Résultats

En 2004, le dispositif n'est pas axé spécifiquement sur les concepts. La moyenne des étudiants est très bonne mais on peut voir que les questions 1 et 7 qui évaluent les concepts sont largement moins bien réussies que les autres questions du style (schémas à compléter, choix multiples, synthèse, etc.)

Tableau 4– Résultats à l'examen de 2004

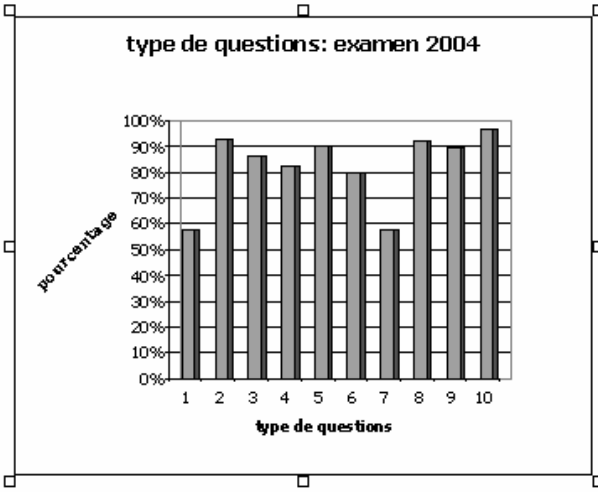
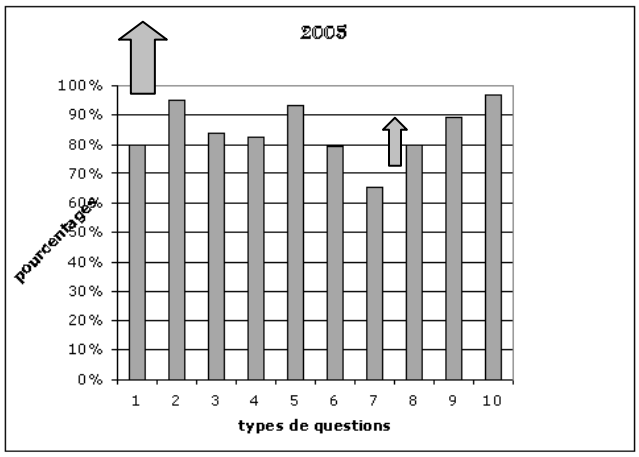


Tableau 5 – Résultats à l'examen de 2005



En 2005, le dispositif comprend des pauses méthodologiques pendant lesquelles les étudiants apprennent les stratégies pour apprendre les concepts. L'examen comportait

exactement la même répartition des questions qu'en 2004. Les questions 1 et 7 portaient sur les concepts. Le résultat est à prendre avec précaution et nuance puisque la cohorte des étudiants n'est pas la même qu'en 2004. Néanmoins, les résultats indiquent une amélioration de 59% à 80% pour les questions de définition et de 59% à 63% pour les questions de classification. Notons une perte de 10% pour les choix multiples (question 8).

4. Evaluation du dispositif sur base de l'évaluation par les étudiants. Les étudiants se trouvent-ils compétents et apprécient-ils le dispositif ?

Nous avons demandé aux étudiants d'évaluer l'activité d'enseignement en trois points : les objectifs, les méthodes et les outils et d'évaluer si ils se trouvaient compétents. Les énoncés sont issus du questionnaire de la Haute Ecole Léonard de Vinci (Belgique).

Les étudiants peuvent répondre par ++, +, -, -- aux affirmations telles que :
-les objectifs du cours sont présentés clairement
-les contenus abordés correspondent aux objectifs
-les exemples proposés illustrent bien la théorie
-la méthode à suivre est clairement proposée
-la méthode facilite le travail demandé
-les outils sont clairement définis
-etc

Tous les questionnaires reçus en 2005 (une vingtaine), c'est-à-dire très peu par rapport au grand groupe initial (250 étudiants) indiquent qu'ils sont satisfaits (++, +) de l'activité d'enseignement. Par contre le sentiment de compétence est mitigé. Les remarques justifient, à juste titre, qu'ils attendent d'être confronté à la pratique pour savoir s'ils sont compétents et s'ils pourront utiliser à bon escient les concepts appris. Ce type d'évaluation quoique toujours très intéressant n'est malheureusement pas suffisant pour avancer des conclusions étant donné la faible proportion des réponses obtenues.

Conclusion

Suite au constat que les étudiants réussissaient moins bien les questions portant sur les concepts nous avons analysé et adapté l'ancien dispositif. Puis nous avons évalué si l'introduction des pauses méthodologiques favorisait un bon apprentissage des concepts.

Nous avons d'une part utilisé le **modèle intégré de Lebrun** (2005) qui intègre l'alignement constructiviste de Biggs, les trois champs de Barbier et Lesne et son propre modèle d'apprentissage IMAIP. Ce modèle permet à l'enseignant de structurer

son dispositif pour qu'il soit cohérent, pertinent et de qualité. Toutefois, même si le changement dans notre dispositif peut venir renforcer les idées maîtresses des sous-modèles, il est difficile de mesurer l'impact de la métacognition par rapport à la notion de « qualité ».

Nous avons d'autre part **comparé les résultats quantitatifs pour mesurer l'impact des pauses méthodologiques** sur l'apprentissage des concepts. Les résultats sont à prendre avec précaution étant donné que méthodologiquement tous les paramètres ne sont pas semblables. Néanmoins, nous observons une amélioration de 5 à 20% en 2005 avec le nouveau dispositif pour les questions sur les concepts.

Enfin nous avons **présenté l'évaluation des étudiants par rapport à notre activité**. Nous avons sélectionné des phrases qui permettent de donner une appréciation sur les objectifs, les méthodes et les outils. Tous les étudiants qui ont participé (très peu) à cette évaluation non obligatoire sont satisfaits (+ et ++) de la façon dont se déroule le cours mais ils ont du mal à se considérer comme compétents par rapport à l'apprentissage des concepts.

L'association des trois types d'évaluation montre une nouvelle fois qu'évaluer est une tâche complexe et que même s'il nous est permis de penser que l'apprentissage conscient de stratégie pourrait faciliter l'apprentissage des concepts nous n'avons pas trouvé jusqu'à présent un moyen de l'évaluer de manière significative.

Listes des références bibliographiques

Anderson, L. & Krathwohl, D. (2001). *A taxonomy for learning, teaching and assessing : a revision of Bloom's taxonomy of educational objectives*. Addison Wesley Longman : Abridged edition.

Bachy, S. (2006). Conséquences épistémologiques de la dimension interactionnelle dans la linguistique : Présentation d'un modèle. *Actes du colloque de l'AIPU*. Monastir.

Bachy, S. & Simon, AC. (2006). Un DVD comme support d'apprentissage dans des cours de linguistique française. *Actes du colloque de l'AIPU*. Monastir.

Bandura, A. (2004). *De l'apprentissage social au sentiment d'efficacité personnelle*. Paris : l'Harmattan.

Barth, B.M. (1987). *L'apprentissage de l'abstraction*. Paris : Retz.

De Ketele, JM. & Chastrette, M. & Cie. (1989). *Guide du Formateur*. Bruxelles : De Boeck.

Frenay, M., Noël, B., Parmentier, P. & Romainville, M. (1998). *L'étudiant-apprenant*. Bruxelles : De Boeck Université.

De Vecchi, G. (1996). Faire construire des savoirs. Paris : Hachette livre.

Gagné, R. (1985). *Conditions of learning*. Ed: Holt. New York.

Laloux, A. (2002). Innovations pédagogiques. Evaluation. IPM Louvain-La-Neuve.

Lebrun, M. (2005). ELearning, pour enseigner et apprendre. Louvain-La-Neuve : Bruylant-Academia.

Paquay, L., Altet, M. & Charlier, E.(1996). *Former des enseignants professionnels*. Bruxelles : De Boeck Université.

Perrenoud, P. (1998). L'évaluation des élèves. Bruxelles : DeBoeck Université.

Prégent, R . (1990). *La préparation d'un cours*. Montréal : Editions de l'Ecole Polytechnique de Montréal.

Raynal, F. & Rieunier, A. (1997). *Pédagogie : dictionnaire des concepts clés*. Issy-les-Moulineaux : ESF.

Rieunier, A. (2000). *Préparer un cours 1. Applications pratiques*. Issy-les-Moulineaux : ESF.

Rieunier, A. (2001). *Préparer un cours 2. Les stratégies pédagogiques efficaces*. Issy-les-Moulineaux : ESF.

Robert, F. (2006). Comment se forgent et se dépassent les préconceptions: l'hypothèse de la généralisation abusive implicite appliquée aux concepts. *Actes du Colloque de l'AIPU* : Monastir 2006.

Verhaeghe, JC. & Wolfs, JL. (2004). *Pratiquer l'épistémologie : un manuel d'initiation pour les maîtres et formateurs*. Bruxelles : De Boeck Université.

Vial, M. (2001). Se former pour évaluer. Bruxelles : De Boeck Université.

Liste des figures

Figure 1 – Un modèle pour réfléchir à la qualité d'un dispositif pédagogique

Figure 2 – Alignement constructiviste de Biggs par Lebrun (2005)

Figure 3 – Schéma des trois champs adapté de Barbier et Lesne par Laloux (2002)

Figure 4 – Le pentagone d'apprentissage IMAIP (Lebrun 2005)

Liste des tableaux

Tableau 1 - Taxonomie de Bloom revisitée simplifiée

Tableau 2 – Tableau des variables de Raynal et Rieunier (1980 dans Rieunier, 2000)

Tableau 3 – Tableau du dispositif pour le concept « résonance », groupe classe

Tableau 4– Résultats à l’examen de 2004

Tableau 5 – Résultats à l’examen de 2005

L' AUTOÉVALUATION DE LA VOLITION ET DE LA PERCEPTION DU SENS

A. LA MARCA — G. ZANNIELLO — G. CAPPUCCIO

Résumé:

Pour cette recherche nous avons utilisé un outil qui permet aux étudiants universitaires d'évaluer leur volition et leur perception du sens, pour mieux comprendre ce qui se passe réellement et les motiver à s'engager dans les études.

Les finalités et les principes méthodologiques qui ont inspiré cette intervention peuvent être rattachées au fait qu'une profonde formation intellectuelle, reçue à travers une intervention éducative qui valorise les capacités de chaque élève, peut créer les conditions favorables qui facilitent l'apprentissage responsable.

La partie théorique de notre comunication est consacrée à l'apprentissage autonome, aux stratégies d'autorégulation et au processus d'autorégulation *de la motivation à s'engager ou à persévérer dans une tâche: rôle actif de la volonté et perception du sens de la vie.*

La partie expérimentale concerne l'inventaire sur la volition (J. Kuhl , A. Furmann) et le questionnaire sur la perception du sens (J. Battista, R. Almond), deux *instruments qui permettent de* stimuler la réflexion de l'étudiant sur sa croissance pour construire des occasions d'auto évaluation et de rétroaction réfléchie sur les progrès de son apprentissage. Dans notre recherche les deux questionnaires ont été considérés comme deux parties d'un unique instrument d'évaluation. On décrit ici la structure des deux questionnaires utilisés et les caractéristiques de l'échantillon des étudiants examinés.

On présente enfin les résultats obtenus parmi les deux instruments d'autoévaluation qui permettent de remarquer seulement ce que l'étudiant dit de lui-même. Pour obtenir une plus grande crédibilité des réponses nous avons comparé les résultats obtenus avec ceux des questionnaires ALM di A. La Marca et QPA di K. Polacek , déjà validés, pour vérifier aussi la validité externe de la version italienne des deux nouveax questionnaires.

L'apprentissage autonome

A travers un apprentissage autonome on permet aux étudiants universitaires de devenir des apprenants capables, autonomes et motivés, et cela tout au long de leur vie. Ce qu'on désire former sont des étudiants universitaires qui apprécient l'étude pour ce qu'il leur apporte sur le plan personnel et social.

L'apprentissage autonome est l'apprentissage par lequel l'étudiant universitaire est capable de prendre, avec les personnes appropriées, les décisions nécessaires pour répondre à ses propres nécessités.

L'apprenant est autonome lorsqu'il devient capable de prendre en charge son propre fonctionnement intellectuel.

Selon cette pensée, les étudiants universitaires développent un système de valeurs, d'attitudes de connaissances et d'habilités leur permettant de prendre des décisions responsables et d'agir selon ces décisions. L'apprentissage autonome est encouragé lorsque les enseignants permettent aux étudiants universitaires de faire des expériences qui stimulent leur motivation, leur curiosité, leur confiance en eux-même, leur indépendance et l'image positive qu'ils ont d'eux-mêmes; l'apprentissage autonome est basé sur l'idée que l'étudiant universitaire doit se rendre autonome *apprenant à apprendre*.

L'apprentissage autonome est un apprentissage continu qui se prolonge tout au long de la vie et porte la personne à réfléchir avant de prendre ses décisions, tout en favorisant l'amélioration de ses capacités.

Il est important de créer un environnement qui contribue à la démarche d'autonomie, permet aux étudiants d'accepter consciemment la responsabilité de leur propre apprentissage et de prendre des décisions dans ce domaine. L'enseignant joue un rôle important à ce niveau, en offrant aux étudiants universitaires un environnement propice qui encourage la motivation, la confiance en soi-même, la curiosité et le désir d'apprendre.

La dimension considérée essentielle pour l'autonomie repose sur le développement d'un système de soi positif en tant qu'apprenant. Selon certains chercheurs, le sentiment d'efficacité personnelle est un déterminant cognitif qui régule les relations fonctionnelles entre les facteurs personnels et la qualité de

l'adaptation et du changement. Il régule la motivation et l'action. Il est donc le mécanisme central de l'action humaine intentionnelle. L'apprenant croit pouvoir réaliser ce qu'il entreprend avec ce qu'il a comme capacités (Galand, 2006, 247-260).

L'autonomie des étudiants universitaires passe par deux points principaux: la possibilité qu'ils réalisent eux-mêmes une partie des tâches de l'enseignant; la possibilité qu'ils améliorent eux-mêmes leurs stratégies d'apprentissage. Parmi ces tâches figurent la programmation de l'enseignement, l'analyse et l'utilisation des stratégies d'apprentissage, l'évaluation des performances. Ainsi, aider les étudiants universitaires à réaliser ces tâches concourt à leur faire atteindre un bon degrés d'autonomie.

L'apprentissage est en effet le processus par lequel des informations, plus ou moins structurées dans les savoirs organisés, deviennent des connaissances et des habiletés intégrées aux dimensions de l'intelligence et de la volonté d'une personne, permettant à celle-ci d'exercer des compétences nouvelles.

L'apprentissage autorégulé

Le fait d'apprendre est un processus naturel qui consiste à poursuivre des objectifs ayant une signification pour l'apprenant. Ce processus est intérieur, volitif et actif; il se définit par une découverte et une construction du sens à partir d'une information et d'une expérience, l'une et l'autre filtrées par les perceptions, les pensées et les émotions propres de l'apprenant.

L'apprentissage autorégulé se caractérise par une interaction complexe entre de nombreux facteurs, d'ordre cognitif, affectif, motivationnel et volitif, et liés aux ressources internes et externes de la personne. Bien plus, la personne ne se contente pas d'être influencée par ces facteurs, mais opère une action de gestion sur ceux-ci, en vue de construire un nouveau apprentissage. Pintrich et Schunk (1996, 2002), expliquent que les processus autorégulateurs ont de grandes propriétés motivationnelles, car ces processus fournissent une direction et un état final aux comportements. Selon Zimmerman, Bonner et Kovach (2000) l'apprentissage autorégulé se déroule en quatre phases:

- *autoévaluation et autocontrôle, détermination du niveau*: autoévaluation du niveau des étudiants universitaires concernant le contenu à apprendre. Cette

évaluation utilise les traces écrites des étudiants universitaires (que l'enseignant doit favoriser). Ces traces écrites peuvent être un "journal de bord", des brouillons. Cette évaluation peut être reprise par l'enseignant.

- *analyse de la tâche d'apprentissage*: les étudiants universitaires fixent des objectifs concernant le contenu à apprendre et préparent et utilisent des stratégies métacognitives.
- *Application de la stratégie et contrôle*: les étudiants universitaires réalisent le plan prévu, en prenant des notes concernant leur démarche.
- *Contrôle des résultats*: Analyse des résultats en rapport avec la stratégie utilisée. L'élève pourra déterminer si la stratégie est adéquate en fonction des objectifs poursuivis.

Suffit-il d'encourager des étudiants à réfléchir sur leurs stratégies d'autorégulation pour que leurs perfomances s'améliorent ? Selon nous, non. L'étudiant universitaire devient autonome dans son apprentissage quand il décide avant tout d'être autonome; il s'agit d'une décision qui implique le sens même de sa vie.

Rôle actif de la volonté de l'élève

L'étude de la capacité d'autorégulation a donné lieu à un nombre impressionnant de travaux de recherche en psychologie de l'apprentissage et de l'éducation, et de nombreuses théories malgré certaines divergences.

Kuhl (1984) et Corno (1993) proposent la distinction entre la motivation et la volonté. La motivation concerne le moment de l'analyse pré-décisionnelle (prise de décision et détermination des buts), tandis que la volonté fait référence au temps du traitement post-décisionnel (activités engagées et régulées dans la poursuite des buts fixés). La personne doit d'une part générer des représentations motivationnelles positives au début de la tâche, afin de s'engager réellement dans son traitement (contrôle motivationnel pour *initier* l'effort); et d'autre part protéger par la suite ces buts poursuivis de buts alternatifs et compétitifs (contrôle volitif pour *maintenir* l'effort). La volonté est alors la capacité de s'atteler à la tâche dans la poursuite du but, et la persévérance face aux difficultés (Corno, 1986, 1993). Pendant la poursuite de son but, la personne doit en effet protéger son intention d'apprendre des

distracteurs et des intentions compétitives. Ces mécanismes ont été étudiés dans le cadre du contrôle volitif, notamment au sein de la théorie du contrôle de l'action (Corno, 2001, Kuhl, 2000).

La perception du sens de la vie

Dans la mesure où l'élève apprendra à raisonner, à chercher la vérité, à travailler avec responsabilité en acceptant les conditions qui dérivent de ses choix et à rectifier chaque fois ce que sera nécessaire, il aura mûri comme personne, en obtenant son harmonie intérieure et il se trouvera dans les meilleures conditions pour pouvoir sortir de lui-même et s'ouvrir aux autres.

Tout ceci manifeste un besoin fondamental de la personne: retrouver le sens de ce que l'on fait, se réapproprier le contrôle de la vie, s'ouvrir à de nouveaux champs de conscience pour sortir de la crise existentielle dans laquelle on se trouve. Dans la dimension sociale la question se pose de plus en plus intensément.

En tant qu'être la personne peut choisir librement sa perspective, sa position et son attitude face aux conditions intérieures et extérieures de son existence. La recherche de sens questionne l'homme dans chacune de ses actions, dès qu'il accomplit quelque chose, il s'interroge pour savoir si sa démarche est dotée de sens.

Dans la littérature relative aux recherches empiriques sur l'attribution de sens on fait souvent référence à la position philosophique et psychologique de Victor Frankl(1965, 1972); Battista et Almond (1973) en analysant critiquement les diverses positions théoriques, ont construit une définition de la nature de l'expérience d'une vie pleine de sens, qui se base sur quatre thèses:

1. Lorsque la personne affirme que sa vie est pleine de sens, cela implique qu'elle possède et adhère à quelque concept ou à un cadre conceptuel concernant le sens de la vie.

2. La conception du *sens de la vie* fournit un cadre de référence dont peut dériver un ensemble de buts spécifiques ou de buts existentiels, que la personne s'engage à réaliser et qui constituent la perspective à partir de laquelle les évènements de la vie peuvent être interprétés et évalués dans leur cohérence.

3. Lorsque la personne affirme que sa vie est riche ou pleine de sens, cela signifie qu'elle perçoit avoir réalisé ou cherché à poursuivre des éléments qui caractérisent son projet de vie personnel.

4. Ce processus d'autoperception et d'autoévaluation est à la base de la possibilité de la personne d'expérimenter effectivement que sa vie, dans ses diverses dimensions existentielles, a un sens.

L'autoévaluation de la volition et de la perception du sens

Les deux instruments utilisés évaluent la capacité de contrôler efficacement la volonté de la personne de mener à terme les engagements avec persévérance, en surmontant la fatigue et la frustration devant les difficultés rencontrées, et la capacité de s'engager dans le travail universitaire pour des raisons intrinsèques en contrôlant adéquatement les sollicitations externes et les intérêts alternatifs. Nous pensons que les deux instruments doivent être employés ensemble, comme un questionnaire unique; c'est que nous avons fait dans notre enquête.

Le premier instrument se réfère à la volition (items 1-68); le deuxième concerne la perception du sens (items 69-96). Le questionnaire est rédigé sous la forme classique d'autoperception à savoir une description du comportement suivie d'une échelle graduée sur quatre niveaux: pas du tout vrai, seulement en partie vrai, assez vrai, presque toujours ou toujours vrai).

1. Le premier instrument *l'Inventaire des Composantes Volitives* (ICV) (Selbststeuerungsinventar: SSI ; Volitional Components Inventory: VCI) a été élaboré par J. Kuhl et Fuhrmann (1998) pour mesurer la compétence volitive c'est-à-dire les composantes volitives qui sont à la base de la capacité d'autodirection. On a vu que la compétence volitive est très importante dans plusieurs domaines comme la psychologie clinique, la psychologie de l'école, la psychologie du sport ou la psychologie du travail ou des organisations. Dans tous ces domaines *l'Inventaire des*

Composantes Volitives peut être utilisé pour mesurer la compétence volitive qui est considérée une qualité- fondamentale[1].

2. Le deuxième questionnaire utilisé, le *Life Regard Index*, élaboré par J. Battista et R. Almond en 1973, mesure le niveau de satisfaction du besoin de sens de la personne à travers deux points de vue: *l'interprétation* de son existence et de ses perspectives de vie et la réalisation de ses choix et de leur valeur existentielle[2].

---

1 Pour vérifier la validité et la fidélité du questionnaire et pour décider si on pouvait l'utiliser de façon stable dans la version actuelle, on a effectué une recherche sur ses qualités psychométriques. L'élaboration statique des donnés recueillies et les différentes enquêtes de nature psychométrique ont été développées à plusieurs niveaux. La validité de construction du questionnaire a été vérifiée au moyen de l'analyse factorielle avec rotation orthogonale des facteurs extraits, en utilisant la technique Varimax . Une première application de la version italienne avec les relatives analyses statistiques a été effectuée par H. Müller sous la direction de M. Pellerey en 2004 avec un échantillon de 253 sujets. Les résultats sont en cours de publication.

2 Le questionnaire a été validé dans le contexte des Etats Unis (Nunnally, 1978; De Vogler et Ebersole, 1980; Chamberlain et Zika, 1988; Zika et Chamberalin, 1992) et Hollandais (Debats, 1990;1996;1998; Debats et al., 1993). Dans le domaine

Les résultats de l'enquête

Les interviewés sont des étudiants de la Faculté de Science de la Formation de Palermo, Enna et Agrigento[3]. L'échantillon comprend 638 sujets, de 19 à 25 ans, dont 83,4 % ont un âge compris entre 19 et 21 ans et 16,6 % entre 22 et 25 ans, repartis selon le tableau suivant:

Corso di Laurea	**Anné**	**ville**	**Numéro questionnaires**
Esperto dei processi formativi	1º	Palermo	102
Educatore Prima Infanzia	1º	Palermo	104
Educatore Prima Infanzia	2º	Palermo	120
Scienze della Formazione Primaria	3º	Palermo	71
Scienze della Formazione Primaria	2°	Palermo	81
Scienze umane e pedagogiche	4º	Palermo	33
Scienze della Formazione Primaria	2º	Enna	52
Tecnologie educative	1º	Agrigento	75
		Tôt	638

de cette recherche, on en a vérifiées positivement les qualités psicometriques même dans le contexte italien.

[3] Les entretiens ont été effectués par des pédagogues qui possédaient une expérience de recherche dans ce domaine. Après avoir été pleinement informés sur les finalités de la recherche et avoir examiné quelques recherches sur l'argument, ils ont personnellement testé le questionnaire, afin de pouvoir relever les difficultés que les étudiants auraient pu éventuellement rencontrer.

Les interviewers ont fourni les même explications initiales à tous les étudiants universitaires contactés et ils ont leur expliqué le sens et le but de la recherche.

D'abord nous présenterons un cadre global des résultats moyens pour chaque item. Ensuite nous donnerons une brève explication aux réponses positives et négatives plus fréquentes. Enfin nous exposerons les résultats de notre enquête sur la validité externe des deux parties du questionnaire d'autoévaluation proposé aux étudiants des cours de la Faculté de Sciences de la Formation.

Les moyennes des items

La lecture des valeurs moyennes des scores des différents items consent de diagnostiquer les défauts les plus fréquents dans les stratégies d'apprentissage des étudiants universitaires. Leur compréhension permet de projeter des interventions formatives selon leurs nécessités réelles. C'est cette compréhension une des raisons de notre enquête.

La moyenne des items indique les lignes de tendance à l'intérieur de l'échantillon. En considérant que pour chaque item, les scores peuvent varier d'un minimum de 1 à un maximum de 4, il faut tenir compte les moyennes supérieures à 2,5 et les moyennes des items inférieure à 2,5. La transformation en chiffres des réponses aux items (pas du tout vrai = 1; seulement en partie vraie = 2; assez vrai = 3; presque toujours ou toujours vrai = 4), met en évidence un problème en ce qui concerne le processus volitif et la perception du sens quand la moyenne des scores est significativement inférieure à 2.5; en revanche on peut affirmer qu'il n'existe pas de problèmes si la moyenne des réponses est supérieure à 2.5.

Si, pour des raisons de validité statistique, on considère seulement les scores moyennes supérieurs à 3.30 et ceux inférieurs à 1.80 on remarque que, en ce qui concerne la *dimension de la volonté*, est très fréquent que les étudiants programment les temps de leur travail; par contre c'est peu fréquent qu'ils ne portent pas à terme le travail, ils se plaignent pour l'excès de travail, ils travaillent pour être estimé des autres. En ce qui concerne la dimension du *sens de vie*, c'est très fréquent que les étudiants ont des buts claires qui leur rendraient heureux, ils savent ce qu'ils veulent faire vraiment dans la vie ; peu d'étudiants attribuent peu de valeur à la plupart des choses qui font, ils n'ont pas un projet de vie et ils attendent les évènements.

Inventaire des Composantes Volitifs														
Item	*m (1-4)*	*dev*		*item*	*m (1-4)*	*dev*		*item*	*m (1-4)*	*dev*		*Item*	*m (1-4)*	*Dev*
1.	3,03	0,66		***18.***	2,90	0,79		***35.***	2,69	0,85		***52.***	3,01	0,75
2.	2,59	0,68		***19.***	3,02	0,73		***36.***	2,50	0,71		***53.***	2,50	0,74
3.	2,42	0,79		***20.***	1,86	0,73		***37.***	2,20	0,73		***54.***	2,31	0,69
4.	**3,37**	0,72		***21.***	2,87	0,78		***38.***	3,18	0,70		***55.***	2,83	0,83
5.	2,91	0,85		***22.***	2,78	0,82		***39.***	3,03	0,83		***56.***	2,54	0,87
6.	2,85	0,84		***23.***	3,06	0,78		***40.***	2,98	0,78		***57.***	2,95	0,80
7.	2,44	0,80		***24.***	2,97	0,83		***41.***	2,25	0,89		***58.***	2,08	0,81
8.	2,48	0,86		***25.***	2,03	0,83		***42.***	1,81	0,81		***59.***	**1,67**	0,79
9.	2,59	0,86		***26.***	2,57	0,92		***43.***	2,52	0,90		***60.***	2,35	0,86
10.	2,65	0,86		***27.***	2,56	0,81		***44.***	2,10	0,78		***61.***	2,68	0,81
11.	2,39	0,92		***28.***	2,25	0,84		***45.***	2,09	0,79		***62.***	2,14	0,81
12.	1,97	0,86		***29.***	**1,73**	0,85		***46.***	2,06	0,87		***63.***	2,07	0,97
13.	2,49	1,01		***30.***	2,18	0,95		***47.***	2,15	1,01		***64.***	2,18	1,00
14.	2,81	0,77		***31.***	2,42	0,73		***48.***	2,82	0,77		***65.***	2,62	0,74
15.	2,02	0,87		***32.***	2,49	0,87		***49.***	2,47	0,79		***66.***	2,53	0,76
16.	2,41	0,98		***33.***	**1,69**	0,76		***50.***	1,92	0,78		***67.***	**1,79**	0,86
17.	2,10	0,95		***34.***	**1,79**	0,83		***51.***	1,94	0,88		***68.***	1,91	0,91

Life Regard Index (1-4)														
item	*m (1-4)*	*dev*		*item*	*m (1-4)*	*dev*		*item*	*m (1-4)*	*dev*		*Item*	*M (1-4)*	*Dev*
69.	2,89	0,86		***76.***	1,93	0,96		***83.***	2,84	0,90		***90.***	2,87	2,23
70.	2,47	0,82		***77.***	**1,42**	0,72		***84.***	3,13	0,84		***91.***	2,86	0,82
71.	3,22	0,84		***78.***	2,64	0,81		***85.***	**1,34**	0,60		***92.***	2,05	0,91
72.	2,09	0,90		***79.***	2,73	0,94		***86.***	2,99	0,83		***93.***	**1,15**	0,46
73.	2,75	1,51		***80.***	**1,58**	0,78		***87.***	1,96	0,86		***94.***	**1,38**	0,66
74.	3,02	0,85		***81.***	**1,23**	0,54		***88.***	2,55	0,80		***95.***	**1,48**	0,78
75.	1,95	0,89		***82.***	3,00	0,88		***89.***	1,92	0,88		***96.***	3,32	0,82

Les réponses les plus fréquentes

L'analyse des réponses à chaque item permet de recueillir des informations plus spécifiques sur les points forts et les points faibles que les étudiants universitaires reconnaissent dans leur façon d'étudier. Nous discutons seulement les valeurs des fréquences égales ou supérieures à 39%, c'est-à-dire au moins 14 points supérieurs à la distribution entièrement fortuite des réponses, parce que, comme nous avons déjà dit, chaque question prévoit quatre alternatives possibles de réponse.

Nous pouvons affirmer que la plus partie des d'étudiants examinés n'ont pas problèmes sérieux d'insuffisante volonté et de manque de sens pour ce qu'ils font.

Items Inventaire des Composantes Volitifs	*pas du tout vrai* ***(1)***	
	f	***%***
59. Souvent je commence à faire une chose sans la mener à terme	*319*	*50*
29. Mes actuelles conditions de vie sont assez dures	*308*	*48,3*
33. Souvent je me comporte sur la base de ce que les autres s'attendent de moi	*296*	*46,4*
67. Si je ne fais pas ce que les autres s'attendent de moi, souvent j'ai peur de perdre leur sympathie	*287*	*45,0*
34. Quelque fois les autres considèrent mes manières de faire contradictoires	*270*	*42,4*
68. J'ai beaucoup d'aspects contradictoires	*248*	*38,9*
42. je renvoie beaucoup de choses que j'ai à faire	*248*	*38.9*
Items Inventaire des Composantes Volitifs	***presque toujours ou toujours vrai (4)***	
	f	***%***
4. Si je suis plein de choses à faire, j'établis le temps le travail (c'est-à-dire quand je dois faire les différentes choses)	*319*	*50,0*

Items Life Regard Index	***pas du tout vrai (1)***	
	f	***%***
93. En réalité il n'y a rien que je veux faire vraiment	*567*	*88,9*
81. En réalité je ne crois pas profondément à rien qui concerne ma vie	*522*	*81,8*
85. Je n'attribue pas beaucoup de valeur à ce que je fais	*458*	*71,8*
94. Je passe la plupart de mon temps à faire des choses pour moi réellement peu importantes	*450*	*70,5*
77. Je ne sais pas vraiment ce que je veux faire dans ma vie	*443*	*69,4*
95. Je n'ai pas réellement un plan de vie	*424*	*66,5*
80. Je ne suis pas capable de porter à terme les choses qui sont vraiment importantes	*365*	*57,2*
76.Il me semble qu'il ne m'arrive jamais rien de vraiment extraordinaire	*262*	*41,1*
Item Life Regard Index	***presque toujours ou toujours vrai (4)***	
	f	***%***
96, J'ai quelques buts qui me donneraient beaucoup de satisfaction personnelle si je pouvais les mener à terme	*325*	*50,9*
84. Il y a des choses auxquelles je dédie toutes mes énergies vitales	*251*	*39,3*

Validité externe de l'instrument

Pour supporter les résultats obtenus par le questionnaire sur la volition et la perception du sens on a fait une corrélation avec d'autres outils d'évaluation déjà validés et que nous considérons liés au processus volitifs et au sens de la vie. On a utilisé les questionnaires *QPA* de Klement Polacek (2005) et *ALM* de Alessandra La Marca (2002).

Le questionnaire *QPA* comprend 90 items rédigés sous la forme classique d'autoperception à savoir une description du comportement suivie d'une échelle graduée sur cinq niveaux. Le but de cet instrument c'est celui de relever l'acquisition de stratégies métacognitives, de la motivation et de l'autorégulation de l'apprentissage. En effet pour encourager un apprentissage autodirigé, il est indispensable de focaliser l'attention sur le processus d'autorégulation et sur l'utilisation de stratégies d'apprentissage et en même temps d'exclure les modèles motivationnels inadéquats. Le questionnaire traite les cinq facteurs présentés suivant: Motivation intrinsèque de l'apprentissage, Métacognition et autorégulation de l'apprentissage, Stratégies de l'apprentissage, Consolidation de l'apprentissage, Apprentissage superficiel.

Le questionnaire *ALM* comprend 122 items rédigés sous la forme classique d'autoperception à savoir une description du comportement suivie d'une échelle graduée sur quatre niveaux (jamais ou presque jamais, quelquefois, souvent, presque toujours ou toujours). Le questionnaire est composé de deux parties. La première se réfère à la motivation, à la volonté, aux réactions et aux interférences émotives avec les échelles: Engagement motivé, Convictions d'efficacité Réaction émotives, Motivations extrinsèques, Valeur du travail scolaire, Manifestations d'intérêt et satisfaction pour le travail scolaire. La deuxième concerne les processus et les stratégies d'approfondissement avec autres échelles: Apprentissage actif, Organisation des connaissances, Autocontrôle et Organisation du mécanisme d'apprentissage, Métacomprension, Autorégulation pour la tache, Recherche d'aide. Tous les facteurs mettent en évidence des connaissances et des compétences de nature métacognitive puisqu'ils signalent d'un coté la sensibilité pour certains processus cognitifs, volitifs, motivationnels et affectifs et, de l'autre, la capacité de les régler de manière valable et productive pendant les apprentissages scolaires.

Corrélations des items du ICV et du LRI avec le total des échelles des questionnaires QPA et ALM.

Parmi les mesures effectuées on a remarqué des corrélations positives ($p < .01$) entre certaines échelles du QPA e de l'ALM et les items de l' *l'Inventaire des Composantes Volitives (ICV)* et du *Life Regard Index (LRI).*

A). A travers les corrélations réalisées entre les items de l'*Inventaire des Composantes Volitives* et certaines échelles du questionnaire QPA c'est très intéressant de remarquer que l'item 39 (souvent la peur d'un échec me stimule à m'engager particulièrement) se corrèle positivement avec les échelles : *Stratégies de l'apprentissage* ($r = .48$) et *Métacognition et autorégulation de l'apprentissage* ($r = .40$). L'analyse des items de ce deux échelles du questionnaire QPA nous permet d'affirmer que la peur de l'échec peut être, au-delà du stimulus à s'engager, une occasion pour développer des stratégies d'apprentissage et d'autorégulation.

En particulier, pour les étudiants interviewés, l'autorégulation aide à adapter la méthode d'étude à la matière à apprendre; à se rendre compte de la méthode d'étude employée; à chercher à comprendre si on assimile la leçon; à exposer des

différentes et réponses à une question importante; à se demander que déjà l'on sait sur le sujet qu'on va lire; à faire des schémas ou des resumés pour mieux apprendre; à accepter les conseils des professeurs pour mieux apprendre leurs disciplines.

La corrélation positive entre l'item 16 (je m'imagine souvent ce que pensent les autres de moi, si je fais ce qu'ils s'attendent de moi) et l'échelle *Apprentissage superficiel* (r = .40) permet d'affirmer qu'un étudiant qui se préoccupe excessivement du jugement des autres a une tendance à étudier de façon superficielle parce qu'il emploie une méthode d'étude sans en connaître la raison; il se limite à étudier seulement ce que suggère le professeur pendant sa leçon; il évite les tâches dans lesquelles il prévoit ne pas pouvoir obtenir un bon résultat; lorsqu'il rencontre une tâche très difficile l'évite; il ne distingue parfois pas quels sont les sujets les plus importants à étudier.

La corrélation entre l'échelle *Stratégies de l'apprentissage* et les item 43 (souvent je pense soudainement à des choses qui ne concernent pas ce que je fais) (r = -. 38) et 12 (dans cet instant le travail/l' étude me coûte beaucoup) (r = - .36) résulte négative. Les deux corrélations négatives confirment l'hypothèse qu'en général un étudiant qui utilise des stratégies adéquates d'apprentissage focalise l'attention sur ce qu'il fait et il ne perçoit pas excessivement le poids ou la fatigue du travail intellectuel.

B). Nous avons trouvé des corrélations significatives entre chaque item de *l'Inventaire des Composantes Volitives* (ICV) et certaines échelles du questionnaire *ALM*. L'item 3 (Lorsque je veux j'arrive à réduire mon anxiété) se corrèle positivement avec l'échelle *Engagement motivé* (r = .38). L'item 19 (Lorsque je déroule une activité difficile si je veux je réussis à me concentrer sur ses côtés positifs) se corrèle avec les échelles *Engagement motivé* (r = .35) et *Convictions d'efficacité* (r = .45). Un élève qui étudie avec un engagement motivé est conscient de l'importance de l'engagement personnel; il continue même si la tâche demande beaucoup de temps et de fatigue, ou s'il ne lui plaît pas, et il porte à terme un travail commencé; il a désir et curiosité de connaître; il sait donner une signification à ce qu'il étudie; il étudie pour s'améliorer comme personne; il est autonome et éprouve

satisfaction dans ce qu'il étudie en déterminant des motivations valables pour étudier. Il formule habituellement des jugements positifs sur sa capacité d'exécuter des tâches déterminées et sur ses connaissances, stratégies et capacité de gérer les situations d'apprentissage.

Une confirmation ultérieure provient de la corrélation positive existante entre l'item 36 (d'habitude je réussis à me motiver assez bien lorsqu'on affaiblit ma volonté à continuer) et les échelles *Convictions d'efficacité* (r = .37), *Manifestations d'intérêt et satisfaction pour le travail scolaire* (r = .36) et *Métacomprehénsion* (r = .35). C'est intéressant de remarquer en outre que l'échelle *Métacomprehénsion*[4] se corrèle positivement même avec les items suivants: 7 (Je suis capable de changer ma mauvaise humeur s'elle empêche mon action) (r = .36); 14 (je suis capable de trouver facilement la sérénité si je sais qu'ensuite je réussis à faire mieux) (r = .37); 40 (Beaucoup de choses vont bien parce que je les affronte énergiquement) (r = .34); 48 (Pour aller en avant tranquillement je suis capable de focaliser mon attention sur des choses positives) (r = .36); 49 (Quand je réfléchis sur la possibilité de faire ou ne pas faire quelque chose d'habitude je fais rapidement mon choix (r = .40). Cela signifie que d'habitude un étudiant qui sait autoréguler ses émotions affronte les situations avec sérénité et énergie, en focalisant l'attention sur des aspects positifs, il est en mesure d'accomplir ses choix avec rapidité, il a acquis une adéquate compétence métacognitive qu'on relève de ses habilités de métacomprehénsion.

L'item 12 (Dans cet moment le travail/l'étude me coûte beaucoup) et l'item 46 (Je dois affronter beaucoup de difficultés) se corrèlent (respectivement, r = .48 et r = .35) avec l'échelle *Motivations extrinsèques* probablement parce que l'élève qui étudie seulement pour le résultat ressenti beaucoup plus la fatigue de l'engagement.

[4] Nous rappelons qu'elle vérifie le degré de connaissance des aspects métacognitifs spécifiques qui influencent l'habilité de compréhension et qu'elle fait référence aux habilités que l'étudiant doit conscientement appliquer pour arriver à comprendre la valeur de ce qu'il lit ou il écoute -

C). Le score global de l'échelle *Apprentissage superficiel* du questionnaire QPA se corrèle avec les item suivants du questionnaire *Life Regard Index* (LRI): 85 (Je n'attribue pas beaucoup de valeur à ce que je fais) (r = .51), 87 (Les autres personnes semblent se sentir mieux que moi en ce qui concerne leur vie) (r = .41), et 94 (je passe la plupart de mon temps à faire des choses pour moi réellement peu importantes) (r = .42). Ceci confirme que l'élève qui étudie de manière superficielle, en plus de passer la majorité de son temps en activités qu'il considère peu importantes, n'attribue pas de valeur à ce qu'il fait et il retient que les autres soient plus contents que lui du genre de vie qu'ils conduisent.

D). Sept items du questionnaire LRI se corrèlent avec six échelles du questionnaire ALM.

Les corrélations positives entre l'échelle *Métacomprension* du questionnaire ALM et l'item 71 du LRI (Je suis profondément satisfait de vivre) (r =. 36); entre l'échelle *Organisation des connaissances* du ALM et les items: 72 (Les autres semblent avoir une idée meilleure sur ce qu'ils veulent faire dans leur vie) (r = .37) et 79 (J'ai un système ou un cadre de référence qui me fait comprendre vraiment pourquoi je vis) (r = .36) du questionnaire LRI confirment l'hypothèse qu'être satisfaits de sa propre vie, même si l'on pense que les autres ont des idées plus précises en ce qui concerne le futur, et posséder un cadre de référence qui donne du sens à l'existence, sont indubitablement des facteurs qui soulignent particulièrement la capacité d'organiser ses connaissances et l'acquisition de stratégies et des habiletés qui permettent une efficace métacomprehénsion. Cela est confirmé même par les corrélations positives qui émergent entre l'item 78 (Je suis aussi impliqué dans ce que je fais que je trouve des nouvelles sources d'énergie dont je n'étais pas d'abord conscient) et les échelles *Manifestations d'intérêt et satisfaction* (r = .37) et *Métacomprension* (r = .35). L'élève qui s'efface positivement à ce qu'il fait, manifeste intérêt et satisfaction et acquiert, avec une plus grande facilité, des habiletés qui facilitent l'apprentissage et qui font référence à la métacomprehénsion.

On observe, en outre, que même l'item 83 (J'ai une philosophie de vie qui fournit de la signification à ma vie) se corrèle avec les échelles *Manifestations d'intérêt et satisfaction* (r = .38) et *Apprentissage responsable* (r = .36). Les étudiants qui donnent valeur à leur vie manifestent intérêt et satisfaction pour l'étude, préfèrent

apprendre ce qui permet de connaître des choses nouvelles; ils choisissent des tâches défiantes et apprennent de manière responsable parce qu'ils discriminent les tâches qui demandent plus d'attente et de programmation consciente; ils savent comprendre et organiser des plans d'action ou des procès d'apprentissage; ils savent prévoir les conséquences de leurs actions; ils sont capables de se interroger sur les raisons d'un comportement déterminé ou d'une manière de procéder, en se posant des questions sur le travail qu'ils font.

Comme on pouvait facilement supposer, l'item 90 (j'ai vraiment trouvé ce qui est important dans ma vie, se corrèle avec l'échelle *Valeur du travail scolaire* (r = .36). Cette échelle est composée par des items qui concernent la conscience de la valeur du travail universitaire et les dispositions nécessaires pour se réaliser comme personne à travers l'étude bien fait (intérêt, engagement et plaisir). Il est clair en effet que la valeur qu'on attribue à la vie est étroitement corrélée à la valeur qu'on attribue au travail. Enfin l'item 96 (J'ai des buts qui me donneraient beaucoup de satisfaction personnelle si je pouvais les mener à bien) se corrèle avec l'échelle *Recherche d'aide* (r = .39). Il est évident que dans la mesure dans laquelle les étudiants sont conscients de ne pas être en mesure de rejoindre tout seuls les buts qu'ils se proposent, ils sentent l'exigence de demander de l'aide.

Nous retenons d'avoir fourni une contribution significative, même si pas définitive, à la validité externe de la version italienne des deux instruments d'autoévaluation employés.

Conclusion

Tout ce que le questionnaire utilisé nous fait remarquer peut résulter particulièrement important pour l'étudiant qui peut se rendre compte comment les processus volitifs et le sens qu'il attribue à la vie puissent influencer positivement ou négativement sur son apprentissage.

Avec les instruments employés un étudiant peut se rendre compte si ses résultats académiques peu satisfaisants, en présence d'habiletés intellectuelles qui ne sont pas inférieures à la moyenne, sont imputables à la faiblesse de la volonté, à une perception inadéquate du sens de la vie, à de mauvaises stratégies d'apprentissage, à l'absence d'habiletés métacognitives.

Les aptitudes, les habiletés et les connaissances nécessaires pour l'apprentissage autonome seront en grande partie stimulées chez les étudiants universitaires par la manière dont les enseignants organisent leur discipline. Il est important que les enseignants utilisent de techniques, de stratégies et d'approches pédagogiques basées sur la collaboration entre étudiants universitaires et enseignants, dans le but d'encourager la participation, lorsqu'il s'agira de fixer des objectifs et d'évaluer la démarche d'apprentissage.

Il est important une approche pédagogique qui favorise la confiance en eux-mêmes des étudiants universitaires et leur donne l'opportunité d'avoir la responsabilité de leur propre apprentissage.

Références

BATTISTA, J. - ALMOND, R. (1973), "The development of meaning in life", *Psychiatry*, 36, 409-427.

BOEKAERTS, M. - PINTRICH, P. - ZEIDNER, M. (Eds.) (2000), *Handbook of self-regulation*, Orlando, FL: Academic Press.

BOUFFARD, T. (2006), "Des apprenants autonomes", in BOURGEOIS, E. – CHAPELLE, G., *Apprendre et faire apprendre*, Paris, Puf, 139-149.

CHAMBERLAIN, K. - ZIKA, S. (1988), "Measuring meaning in life: An examination of three scales", *Personality and Individual Differences*, 9, 589-596.

CORNO, L. (1986), "The metacognitive control components of self-regulated learning", *Contemporary Educational Psychology*, 11, 333-346.

CORNO, L. (1993), "The best-laid plan. Modern conceptions of volition and educational research", *Educational Researcher*, 22, 14-22.

CORNO, L. (2001), "Volitional aspects of self-regulated learning", in ZIMMERMAN B.J. - SCHUNK D.H. (Eds.), (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives*, Mahwah, NJ:Erlbaum.

DE VOGLER, K.L. – EBERSOLE, P. (1980), Categorization of college students' meaning in life, Psychological Reports, 46,387–390.

DEBATS, D. L. (1990). The Life Regard Index (LRI): *Reliability and Validity. Psychological Reports*, 67, 27-34.

DEBATS, D. L. (1998). Measurement of personal meaning: The psychometric properties of the Life Regard Index. In WONG, P. - FRY P. (Eds.), *The human quest for meaning: A handbook of psychological research and clinical application*, Hillsdale, NJ: Lawrence Erlbaum, 237-259.

DEBATS, D. L., VAN DER LUBBE, P. M. - Wezeman, F. R. A. (1993), "On the psychometric properties of the Life Regard Index (LRI): A measure of meaningful life", *Personality and Individual Differences*, 14, 337-345.

FRANKL, V. E. (1965), *Man's search for meaning: an introduction to logotherapy*, Boston, Beacon Press.

FRANKL, V. E. (1972), "The feeling of meaninglessness: a challenge to psychotherapy", *American Journal of Psychoanalysis*, 32(1), 85-89.

GALAND, B. (2006), "Avoir confiance en soi". In BOURGEOIS, E. – CHAPELLE, G., *Apprendre et faire apprendre*, Paris, Puf, 247-260.

KUHL, J. – FUHRMANN, A. (1998), "Decomposing self-regulation and self-control: The volitional components inventory", *Motivation and self-regulation across the life span*, 15-49

KUHL, J. (1984), "Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control". In MAHER B.A. - MAHER W.B. (Eds.), *Progress in experimental personality research*, New-York: Academic Press.

KUHL, J. (2000), "A functional-design approach to motivation and self-regulation: the dynamics of personality systems and interaction". In BOEKAERTS M. - PINTRICH P.R. - ZEIDNER M. (Eds.), *Handbook of self-regulation: Theory, research, and applications*, San Diego, CA: Academic Press.

LA MARCA, A. (2002), *Io studio per…imparare a pensare*, Troina, Città Aperta.

NUNNALLY, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

PINTRICH, P.R. - SCHUNK, D.H. (1996), *Motivation in education: theory, research, and applications*, Englewood Cliffs, NJ: Prentice Hall.

POLACEK, K. (2005), *QPA Questionario sui processi di apprendimento*, Firenze, O.S: Organizzazioni Speciali.

ZIKA, S. - CHAMBERLAIN, K. (1992), "On the relation between meaning in life and psychological well-being", *British Journal of Psychology, 83*, 133-145.

ZIMMERMAN, B.J. - SCHUNK, D.H. (2001). (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives*, Mahwah, NJ:Erlbaum.

ZIMMERMAN, B.J. – BONNER, S. – KOVACH, R. (2000), *Des apprenants autonomes, autorégulation des apprentissages*, Bruxelles, De Boech.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

L'EVOLUTION DES QCM ET CRITERES DE QUALITE ETIC –PRAD

Dieudonné Leclercq, Université de Liège

A. Contexte

Comme tout pédagogue qui respecte ses principes, j'ai attribué deux objectifs au présent article. Le premier objectif est que le lecteur, au terme de cet article, ait perdu son innocence[1] quant aux QCM (ne plus croire qu'il existe une et une seule forme de QCM, la seule qu'il connaisse, qu'il pratique...ou qu'il comprenne).

Le second objectif est que le lecteur soit capable de dénouer les brins constituant la « tresse » de la qualité d'un Système d'Evaluation des Performances d'Individus en Apprentissage (SEPIA) incorporant des QCM. Nous parlons d'un système car c'est au regard des diverses composantes d'une évaluation (sa visée, certificative ou formative ; sa précision : diagnostique ou sommative ; ses objets : processus, produits ou les deux ; ses dimensions : unique ou multiples ; ses agents : externes ou internes, etc.) et de leurs interactions que l'on peut juger de la qualité d'une évaluation. Chacun de ces aspects a été décrit et illustré dans notre « Rose des vents des caractéristiques et fonctions de l'évaluation » (Leclercq, 2005). Cette **description** n'étant pas une **évaluation de la qualité** , nous avons par ailleurs développé une série de critères de qualité, série résumée par l'acrostiche ETIC PRAD.

Cette grille ETIC PRAD d'analyse de la qualité est inspirée de la grille VENTURE (1972) développée par le CSE (*Center for the Study of Evaluation)* de UCLA et et de l'association RBS (*Research for Better School*). Cette grille a été appliquée aux USA à plus de 1000 tests à usage scolaire, à l'exclusion de tests de pure connaissance de mémoire. Trois grands domaines ont été investigués : Les Higher Order Cognitive Skills, les Affective Skills et les Relational Skills. Chaque test était évalué selon 7 critères :

Validity (décomposée en construct validity, content validity)

Examinee appropriateness (facilité d'utilisation pour l'élève, adéquation à ses possibilités intellectuelles)

Normed excellence (existence de normes statistiques permettant de situer une performance parmi celles des autres étudiants du même âge ou du même niveau scolaire)

Teaching feedback (intérêt du résultat ou feedback pour l'utilisation en classe, en vue d'améliorer la performance)

Usability (commodité d'administration : rapidité, faible coût, sans matériel spécial, par une seule personne, à correction aisée)

Retest potential (possibilité de réutiliser le même test plusieurs fois, notamment pour mesurer les progrès)

Ethics (absence de problème éthique, notamment de discrimination).

---

[1] Dans le sens que B. Bloom (1975) donnait à cette expression dans son article « L'innocence en pédagogie » : ne plus pouvoir dire « je ne savais pas », parade des incompétents pour justifier tous leurs errements. Désormais, la personne formée sera bien obligée de plaider coupable.

Chacun de ces critères est composé d'indicateurs permettant d'attribuer des points.

Ainsi, pour la validité, les scores peuvent aller de 0 à 13. Par contre, l'Examinee appropriateness ne va que de 0 à 6 et l'Ethics se résume à un Oui/non.

Pour chacun des critères (sauf pour l'Ethics), trois catégories de scores sont déterminées : Poor (P), Fair (F) et Good (G). Ainsi, pour la Validité, un score inférieur à 6 vaut un P, un score entre 6 et 10, un F et un score de 11 à 13 un G. Chacun des 1000 tests reçoit donc un label final en 6 lettres consécutives, comme PPFPGF par exemple, dans l'ordre de VENTURE. En cas de problème éthique, s'ajoute un astérisque. Par exemple **GGFPGF*** Les auteurs ont établi le « portrait robot » du test modal constitué des modes de chaque critère Cela a donné **PGPPGP**

En d'autres termes, en majorité, les tests considérés ne posent pas de problèmes éthiques (Eth), sont bien adaptés aux élèves (Ex), sont faciles à utiliser par les enseignants (U), **mais** on ne sait pas trop bien ce qu'ils mesurent (V), ils ne permettent pas de situer par rapport aux autres élèves (N), ils ne servent à rien pour la formation (T) et ils ne peuvent pas être réutilisés (R).

B. Les 8 qualités ETIC PRAD d'un SEPIA

B1. Validité Ecologique

Cette expression, due à Egon Brunswick (1943), représente la mesure dans laquelle la situation d'évaluation correspond à la situation de la vie réelle qu'elle est sensée représenter ou prédire. Ainsi, pour mesurer la capacité de manœuvrer chez un conducteur de camion, lui demander d'introduire au clavier une grandeur angulaire de braquage ne rencontre pas le critère de validité écologique car, dans la réalité, c'est au moyen d'un volant et à 1,5m du sol qu'il aura à manœuvrer. Certaines personnes ont en effet de grandes capacités de « manœuvres dans l'espace », mais dans certaines situations. Ceci rejoint l'idée de Howard Gardner (1996), le promoteur de l'idée des « intelligences multiples » qui note *« La mesure d'une intelligence donnée … devrait mettre en lumière les problèmes susceptibles d'être résolus dans les données et les outils propres à cette intelligence »*….(1996, 48) et « *Quand les individus sont évalués dans des conditions proches de « véritables situations de travail », il est possible de prédire leur résultat final avec beaucoup plus de précision* » (1996,158). On voit que les qualité de validité écologique et prédictive (pour plus loin) sont souvent associées.

Plus les arguments en faveur de la validité écologique du test seront puisés dans la vie courante, plus l'épreuve aura une « validité apparente » (en anglais ***face validity***).

B2. Validité Théorique

Elle se décompose en validité de **contenu** (ou de « couverture » du contenu : tout ce qu'il faut tester l'est-il et rien que cela ?) et validité de **construct** (le Système d'Evaluation des Performances d'Individus en Apprentissage ou SEPIA est-il fondé sur un modèle crédible, scientifiquement fondé, des **Processus Mentaux** ?). Les auteurs d'un test doivent établir ce type de validité par des arguments empruntés aux grandes théories et par des résultats expérimentaux jugés fiables. La notion de validité

de construct fut introduite par Cronbach et Meehl en 1955. Par exemple, on peut créer un test de compréhension d'une matière en se basant sur la théorie de Bloom (1956)[2], fortement inspirée du behaviorisme des années 50 ou par celle de Smedslund (1997), inspirée, elle, du socio-constructivisme piagétien et du cognitivisme. Elle est d'ailleurs publiée dans les Cahiers de Psychologie Cognitive. (NB les parenthèses sont de nous) :

« Une personne E (comme Elève) comprend correctement ce qui est signifié par une expression X produite par une personne P (comme Professeur) si les deux personnes s'accordent sur ce qui, pour P : (1) est équivalent à X, c-à-d signifie la même chose que X (2) est impliqué par X, c-à-d découle de X, (3) est contredit par X, ou nié par X, (4) est non pertinent par rapport à X, c-à-d n'a rien à voir avec X. Un désaccord sur un seul de ces critères indique une mauvaise compréhension (misunderstanding). Cela suppose que E est logique. On ne peut en effet distinguer un manque de compréhension d' un manque de logique. »

Nous avons montré ailleurs (Leclercq, 2005, chap 4) comment construire des épreuves à QCM sur base de l'une et de l'autre théorie.

[2] Ex : 10 : Transposition : (11) Traduire, (12) transformer, (13)dire avec ses mots, (14) illustrer, (15) préparer, (16) lire, (17) représenter, (18) changer, (19a) réécrire, (19b) redéfinir……(a) des significations, (b) des exemples, (c) des définitions, (d) des abstractions, etc. 20 : Interprétation : (21) Interpréter, (22) réorganiser, (23) réarranger, (24) différencier, (25) distinguer, (26) faire,etc.. 30 : Extrapolation : (31) estimer, (32) inférer, etc

B3. Validité Informative ou Diagnostique

C'est la multiplicité des données et leur distinctivité, leur capacité d'être précises (porter sur une capacité et non sur la voisine). Par exemple, l'examen oral sollicite chez l'étudiant à peu près tous les niveaux de la taxonomie des processus cognitifs de Bloom. Cependant, il est fréquent que la communication vers l'étudiant se résume à un « c'est satisfaisant », sans plus de commentaire. C'est que plusieurs indices concordants (qui se consolident l'un l'autre) confortent l'interrogateur dans son jugement. Par contre, souvent, il ne commente pas chaque dimension, entre autres parce qu'il n'a pu en observer qu'un échantillon…et en combinaison avec d'autres dimensions. Souvent, il ne peut pas, sur la base des données à sa disposition, indiquer la (ou les) causes exactes du manque de qualité d'une performance (manque d'étude ? manque de capacité de mémoriser ? manque de capacité de communiquer ? stress paralysant, etc.).

Un exemple de précision diagnostique est fourni par les feedbacks donnés, tous les 3 mois aux étudiants des 6 années de la Faculté de Médecine de Maastricht qui ont subi le même examen (Progress Test) sur toute la médecine. Ce feedback (Leclercq & Vandervleuten, 1998, 200-201) précise à chaque étudiant pour chacune des 29 disciplines et des 14 catégories, non seulement son nombre de réponses correctes sur le nombre de questions mais sa position (note z) par rapport aux autres étudiants de son année. Et ce 24 fois au cours de toutes les études.

B4. Validité Conséquentielle

La validité conséquentielle (Green, 1998 ; Moss, 1998 ; Reckase, 1998 ; Talepros, 1998 ; Yen, 1998)) d'une évaluation pédagogique s'apprécie aux suites que cette évaluation a sur les représentations, les actes (ex : réviser ou non la matière, changer ou non de méthode d'étude) des apprenants, des formateurs ou d'autres personnes.

Un exemple d'une telle validité est fourni par l'opération RESSAC (Résultats à des Epreuves Standardisées au Service des Apprentissages en Candidatures). Les étudiants d'une première année en psychologie (Leclercq et al., 2003) ont reçu, pour chacune des épreuves de janvier subies dans 4 cours, deux notes : l'une portant sur leur performance de « restitution de Mémoire » et l'autre sur leur Compréhension en profondeur ». Une moitié environ des étudiants en échec ont dit avoir modifié leur méthode d'étude (plus de mémorisation, plus d'étude en profondeur ou les deux) pourpréparer les examens de juin et septembre. Leurs taux de réussite à ces examens a été nettement et systématiquement (dans chacun des 4 cours) nettement supérieur à ceux qui ont déclaré ne pas avoir changé leur stratégie d'étude.

Autre exemple : nous avons démontré à des étudiants de 1° année universitaire la tendance générale à surestimer sa propre compréhension de mots d'un texte technique. Après avoir passé un test qui le montrait sur leurs propres données, nombreux étudiants ont dit avoir, en conséquence, consulté beaucoup plus le dictionnaire. Ce qu'un post-test a confirmé. (Leclercq, Simon, Marotte et Lacaille, 2002.).

B5. Validité Prédictive ou Concurrente

Les mesures obtenues permettent-elles de prédire efficacement (c'est-à-dire avec précision) d'autres mesures souvent ultérieures (par exemple la réussite scolaire ou professionnelle, le rendement à une autre épreuve, etc.). C'est la corrélation entre les mesures prédictives et les mesures critères (ou à prédire) qui permet de répondre à cette préoccupation. Le cas échéant, la validité prédictive peut être établie en l'absence de validité de « construct »; c'est le cas lorsqu'un instrument prédit efficacement sans que l'on comprenne pourquoi. Ce type de situation n'est pas propre à l'éducation.

Un exemple célèbre est celui d'André Inizan (1972) qui a établi une batterie de 8 tests prédictifs passés en dernière année de l'école maternelle[3] . Un score total pour chaque enfant est calculé par la somme de ses scores à ces 8 tests. Par un suivi longitudinal de dizaines d'enfants, il a établi le nombre de mois nécessaires pour savoir lire (atteindre un score de 38 points à son test de lecture) en fonction du score total au test prédictif et de l'âge à l'entrée en primaire.

Un autre exemple est fourni par la Régression Multiple entre les notes obtenues à des tests à l'entrée (prédicteurs) sur la réussite en première année universitaire (critère) dans les diverses facultés des universités de Belgique francophone, dans l'opération MOHICAN (Leclercq et al, 2003), MOnitoring HIstorique des CANdidatures. .

B6. Replicabilité ou stabilité – Reliability (fidélité)

Une formule (Ebel, 1969, 566) précise le nombre de questions d'une épreuve et, pour les QCM le nombre de solutions proposées (distracteurs) nécessaires pour obtenir un niveau de fidélité donné (0,8 par exemple). Des formules répondent à la question de la façon inverse : quel doit être le **coefficient d'allongement n du test** pour atteindre une fidélité donnée (par exemple 0,80 ou 0,90) d'un test qui existe déjà et dont on connaît la fidélité actuelle ?

B7. Acceptabilité - Praticabilité

Pour le professeur, plusieurs composantes sont à envisager : l'adhésion et l'applicabilité. Un exemple relatif à l'adhésion : les évaluations pédagogiques doivent-elles servir plus à sélectionner qu'à former ? Autre exemple : certains enseignants considèrent que la capacité d'évaluation (le niveau le plus élevé de la taxonomie des objectifs cognitifs de Bloom) doit intervenir dans la notation des performances des étudiants, en accordant des points (supplémentaires, donc positifs !) au réalisme dans l'auto-évaluation de leurs compétences. Ce réalisme peut être confronté à la réalité et on peut calculer objectivement la surévaluation et la sous-évaluation . D'autres enseignants, cependant, n'acceptent pas, par principe, de combiner deux mesures dont l'une fait appel à la subjectivité (pourtant mesurée objectivement).

[3] Copie de Figures Géométriques (FG), Répétition de Rythmes sonores (RR), Copie de Rythme écrit (CR), Articulation (A), Mémoire de Dessin (MD), Mémoire de Récit, Copie de lettres ou test de Horst (H), Manipulation de cubes de Kohs (K).

Les critères d'applicabilité sont faciles à énumérer : durée, matériel requis (ordinateur ou seulement papier ?), concentration, précautions (antifraude par ex.), de quel lieu, à quels moments, etc.

Pour l'étudiant aussi plusieurs critères sont à considérer : l'adhésion et la familiarité. Ainsi, fréquemment un nombre élevé d'étudiants n'adhèrent pas au principe de passer des tests (optionnels) formatifs et informatifs : ils ne profitent pas de cette occasion qui leur est offerte.

Quant à la familiarité, il a été démontré (Leclercq, 1986, 108-115) que plus l'étudiant est familier avec les procédures de testing, avec les barèmes de cotation, etc. plus il est « aguerri aux tests » (en anglais « test wiseness »), plus ses chances de réussite sont élevées, tout spécialement avec les QCM.

B8. Validité Déontologique (ou Ethique)

Depuis longtemps la docimologie négative (Piéron, 1963) a montré que les corrections de copies par des juges sont l'objet d'effets regrettables (de contraste, de sévérité du correcteur, de halo, de non concordance interjuges, de non constance intrajuge, d'effet Posthumus, etc.) qui sont largement évités par le recours aux QCM. Par ailleurs, les droits des étudiants étant de plus en plus (et à bon droit) reconnus, les systèmes d'évaluation garantissent de plus en plus la transparence de l'évaluation en termes de recalculabilité de la note à partie de la copie brute, de contrôlabilité du processus, ce que les QCM permettent.

C.L'évolution des QCM à la lumière des critères de validité ETIC PRAD

Nous décrirons ci-après seulement 7 « moments », sur un bien plus grand nombre possibles.

Une mise **en gras** d'une des 8 lettres signifie une amélioration et un **soulignement** un affaiblissement.

0

1C1. La gloire de la consigne classique dès sa naissance

Pressés de sélectionner les officiers parmi les appelés à la guerre de 1914-1918, les Etats Unis font confiance aux Army tests conçus par Otis. Il s'agit de tests constitués de Questions à Choix Multiple fonctionnant avec la consigne classique (QCMC), c-à-d : « Une seule des solutions proposées est correcte ; vous avez droit à une seule réponse ». . Le fait que les USA se sont retrouvés parmi les vainqueurs n'a pas peu fait pour assurer une crédibilité à ce mode de testing (validité **Prédictive**). Au cours des années suivantes, les modalités de testing systématique ont encore accru l'exigence d'efficience (rapport coût/efficacité) tant appréciée par les Américains du Nord (validité d'**Acceptabilité**). Après la guerre 1940-1945, aux USA toujours, l'exigence grandissante d'impartialité (Civil Rights) dans la notation, de non-discrimination, a fait apprécier ce que les américains ont appelé les « objective tests » (validité **Déontologique**), alors que ces tests n'ont d'objectif que la correction. Ces trois types de validité expliquent, à notre avis, le plus grand attachement des Américains aux QCM que ne l'ont été et le sont les Français par exemple.

Pourtant, dès 1963, dans leur livre «La Docimologie », Piéron et ses collaborateurs (Laugier et Weinberg) avaient montré les discordances importantes pouvant exister entre les notes de différents juges d'une même copie « rédigée », et même l'instabilité de la note d'un même juge à une même copie. Au courant de ce problème, les autorités françaises ont cependant maintenu la notation subjective, sur la base du raisonnement selon lequel le correcteur ne connaissant pas l'identité de l'auteur de la copie, les injustices se répartissent au hasard selon un bon vieux principe (français lui aussi) d'égalité. Il se pourrait que la pratique de plus en plus courante de recours en justice (mode venue elle aussi des USA) des étudiants contre la note obtenue amène à reconsidérer la situation. En Tunisie, zone d'influence pédagogique française, certains types de copies du Bac (en philosophie, en français) font déjà l'objet d'une double correction systématique.

C2. Une attaque théorique sur le hasard et sa parade

Tversky (1964) définit la puissance d'un test par « 1 – la probabilité d'atteindre la performance parfaite par chance ». Or on sait qu'à chaque QCMC qui comporte k solutions, l'étudiant a 1/K chances de fournir la solution correcte par chance. Très tôt (Mc Call, 1920), cette objection a été contrée par la « correction for guessing classique : on retire 1/(k-1) pour chaque réponse incorrecte à un QCMC. Tout aussi tôt (West, 1923), cette procédure a été critiquée à son tour, mais ces voix ont été eu entendues (ou peu lues ? ou peu crues ?). Nous prétendons, aujourd'hui encore que cette procédure manque de validité **Déontologique** car elle est injuste : elle pénalise

aveuglément les personnes à qui ont a interdit d'exprimer leur degré de doute. En outre, elle est basée sur un modèle **théorique** dépassé (et faux) comme l'a montré Choppin (1975) ; quelle est basée sur un modèle théorique faux. De plus, elle manque de validité **diagnostique** pour les enseignants puisqu'elle ne leur apprend rien de plus. Enfin, et pour les mêmes raisons, elle n'a pas de validité **conséquentielle** car, à part « omettre plus souvent », elle n'a pas d'effet sur le comportement des étudiants. Cross & Frary (1977) ont démontré pourquoi cette procédure dissuade peu de « deviner ». Les tenants des Degrés de Certitude (voir plus loin) soutiennent que cette procédure répond aux quatre manques signalés ci-dessus.

Une autre parade a consisté à recourir à la consigne des QCRM (Questions de Choix à Réponses Multiples) où l'étudiant est invité à répondre Vrai / Faux / Omission à chacune des solutions proposées, n des k solutions étant correctes et k-n étant incorrectes). Cette procédure, qui corrspond au modèle 2 de l'activité mentale décrit par Choppin) a donné lieu à d'intéressants développements (Coombs, Milholland & Womer, 1966) permettant d'approcher pour la première fois la « connaissance partielle » chère à De Finetti (1965). On attribue alors un point positif par réponse incorrecte et un point négatif par incorrecte. Le score à une question peut donc varier de –k à +k.

C3. Une rafale de critiques théoriques sur les processus mentaux mesurés

C3.1. Les QCM ne mesurent pas l'évocation de mémoire

Il est évident que les QCM ne peuvent prétendre mesurer la capacité d'évoquer des connaissances, mais bien celle de les « reconnaître », ce qui n'est pas la même chose. Depuis longtemps, en effet, on sait (Luh, 1922) que la performance de recognition est bien plus facile (a un taux de réussite plus élevé) que la performance d'évocation. Ces observations ont été maintes fois confirmées dans des contextes aussi différents que l'apprentissage de langues étrangères (Bahrick, 1984) ou de la médecine (Schurwirtz, 2000). Ajouter la solution « Aucune » (ou « Autre ») aux solutions possibles permet de palier, quoiqu' imparfaitement, cette faiblesse et constitue une amélioration de la validité **théorique**.

C3.2. Les QCM invitent au raisonnement à rebours

Même avec la solution «Aucune » ou « Autre », les QCMC induisent un processus mental ne correspondant pas à celui que les étudiants doivent pratiquer dans la vie courante. Avec une QCMC du type « Dans quel pays a été décrété pour la première fois l'Habeas Corpus ? 1. La France 2. L'Italie 3. L'Allemagne 4. Autre. », l'étudiant a tendance à d'abord considérer (et éliminer) les solutions proposées, puis seulement choisir la solution 4 (la réponse correcte est « L'Angleterre »). C'est le modèle 2 de l'activité mentale décrit par Choppin. Or, ce que l'on veut mesurer, c'est sa capacité à EVOQUER la sa solution, puis seulement à la confronter à des solutions possibles. C'est le principe des QCL (Leclercq, 1999) : l'étudiant reçoit une liste de plusieurs centaines de solutions rangées par ordre alphabétique, ce qui rend cette liste proche de l'index d'un livre. Chaque solution possible est affectée d'un numéro d'ordre (par exemple de 1 à 700) et c'est par ce numéro en trois chiffres (lisible par le lecteur optique de marques) que l'étudiant répond. On garde ainsi les avantages de l'automatisation de la correction, en donnant une plus grande validité de construct

dans la mesure où le processus est respecté. Ainsi, à la question «Quel nom donne Gavriel Salomon au processus médiatique qui fait faire le travail par le média au lieu de le laisser faire par le cerveau du spectateur ?…… », l'étudiant doit d'abord penser « supplantation », PUIS consulter la liste alphabétique et noter 482, le code du terme « supplantation ». Il n'aurait pas le temps de reparcourir à chaque fois la totalité des 700 termes pour travailler par « reconnaissance » plutôt que par « évocation ». Cette procédure contribue à la validité **théorique**, de *construct*. L'automatisation de la correction permettant de poser beaucoup de questions (plus d'une par minute, par exemple 100 en une heure) contribue d'une autre façon encore à la validité **théorique** mais dans son aspect « validité de contenu ».

Les QCL seront cependant de plus en plus abandonnées avec le recours aux réponses par clavier. Bien sûr, il suffit, dans ce cas-là de taper le mot (ou l'expresssion courte) correcte. Schurwirtz (1998) a cependant développé un principe de réponse combinant la réponse ouverte et la QCM, voir la QCL. Ce médecin chercheur de l'Université de Maastricht développe, dans le cadre de l'approche PBL (Problem Based Learning) des tests interactifs sur ordinateurs sur base de cas médicaux. Le programme d'ordinateur commence par présenter à l'étudiant testé une « vignette clinique » (un cas de patient), et l'étudiant doit diagnostiquer la maladie. Pour ce faire, il commence à introduire sa réponse au clavier. Sur base des quatre premières lettres, le programme affiche alors la liste des diagnostics possibles (en tout 2500) rangés par ordre alphabétique comme le fait Windows quand on « cherche » sur base d'un mot entré au clavier. Ainsi, si l'étudiant a tapé « Diab », le programme affiche «diabetis (type 1)» et l'étudiant peut « dérouler le menu » accompagnant ce mot pour choisir, parmi toutes les formes de diabète celle qui constitue sa réponse. Ceci constitue une amélioration de la validité d'**Acceptabilité -** Applicabilité de la technique. En effet, quand il existe des milliers de réponses possibles la technique QCL est incommode : chaque testé devrait alors avoir devant lui non pas une feuille A4 (recto-verso), mais un dictionnaire.

2

C3.3. Les QCMC ou QCMR renforcent le curriculum caché de l'école

Le curriculum caché[4] est ce que personne n'enseigne mais que tout le monde apprend à l'école. On y apprend, notamment, que quand une question est posée, il faut y répondre ; or certaines questions, parce qu'elles sont absurdes ou excessivement intrusives, ne doivent ou ne peuvent recevoir de réponse ! On y apprend que quand l'autorité pose une question, elle est forcément pertinente, bien posée, etc. On y apprend que toute question à UNE réponse et que si on ne la connaît pas, on jne peut pas la retrouver par le raisonnement. Bref, le curriculum habituel (il y a de merveilleuses et de plus en plus nombreuses exceptions) n'exerce pas à la vigilance cognitive, à le détection des anomalies, des incohérences, etc. notamment par ses modalités de testing, les QCMC en étant la plus représentative. Grave lacune dans la validité **théorique** de cette technique !

---

[4] Michael Haralambos ("Sociology: Themes and Perspectives", 1991): "*The hidden curriculum consists of those things pupils learn through the experience of attending school rather than the stated educational objectives of such institutions*".

Pour toutes ces raisons, nous avons développé (Leclercq, 1986, 127-144) les QCM à Solutions Génrales Implicites ou QCM SGI. Ces solutions sont au nombre de quatre : Aucune , Toutes, Manque de données dans l'énoncé, Absurdité dans l'énoncé. Elles sont Générales par ce qu'elles sont valables (et identiques) pour toutes les questions d'un test par QCM SGI. Elles sont implicites parce qu'elles ne sont présentées qu'une seule fois (au début du est) et ne sont pas répétées dans chaque question : l'évalué doit y penser tout seul. Du coup, cette procédure a aussi un impact sur la validité **informative** (ou diagnostique) car elle permet de distinguer deux niveaux de la taxonomie de Bloom : la compréhension (sans piège) et l'analyse (avec piège)
Gilles (1999) a montré, que les QCM SGI dont la réponse correcte est une SGI avaient une validité prédictive supérieure à celles dont la réponse correcte est une solution « visible » pour la réussite d'étudiants en médecine.

C4. La rencontre entre QCM et DC (Degrés de Certitude)

Rappelons d'emblée que le recours aux Degrés de Certitude est indépendant des QCM. On peut très bien poser une question ouverte (du genre « En quelle année a eu lieu la bataille de Waterloo ? » et demander à l'étudiant d'accompagner sa réponse d'un degré de certitude. Shufford (1966), Van Naerssen (1965) et De Finetti (1965) ont montré que la consigne ne devait pas être verbale (« peu sûr », « moyennement sûr », « très sûr ») mais probabiliste (en pourcentages de chances). Nous avons en outre montré (Leclercq, 1982, 1993) qu'une précision plus grande que 20% était illusoire, d'où notre consigne en 6 degrés : 0%, 20%, 40%, 60%, 80%, 100%.
Avecles auteurs précités, nous pensons que ce procédé a une plus grande validité **écologique** que le testing habituel qui empêche les étudiants d'exprimer leur doute. Choppin (1975) a décrit ce problème dans ses modèles 1, 2 et 3. Il dénonce la vision manichéenne (tout ou rien) de phrases telles que « répondez uniquement si vous savez ; omettez si vous ne savez pas », alors que nous sommes très souvent (et en particulier lors de situations d'apprentissage) dans des états de <u>connaissance partielle</u>. (DeFinetti, 1965). Des sentiments du genre « j'irai relire le cours, j'irais voir dans le dictionnaire, sont résumés dans le DC), comme nous l'avons montré expérimentalement (Leclercq & Boskin, 1990).
Avec les QCM, les DC résolvent en outre (mais c'est un heureux effet secondaire, pas le but principal) le problème du « guessing », ce qui contribue à la validité d'**acceptabilité** (par les enseignants) des QCM.

Enfin, les DC montrent leur importante contribution à la validité **informative (diagnostique**) des QCM dans des recherches comme celle de Baragabiribije (2005) où les questions qui font l'objet de conceptions erronées (*misconceptions*), ici en physique, présentent des solutions erronées reçoivent une certitude moyenne plus élevée que les réponses correctes, ce qui est anormal.

La place étant limitée, nous arrêterons ici cette dialectique entre les améliorations apportées aux QCM et les critiques qui continuent à leur être faites, les deux contribuant à améliorer divers aspects de la validité des mesures. L'histoire n'est as finie. Nous invitons les lecteurs qui s'en sentent le désir...et le courage, d'écrire quelques pages de cette passionnante histoire.

Références

Alkin, M. , Thomas, J, Hill, R., Levin, J., Scanlo,n, R., Boyer, G & Simon, A. (1972) VENTURE. Los Angeles : Center for the Study of Evaluation at UCLA and Research for Better Schools, Inc..

Baragabirije, D. (2003) Testing des représentations erronées de la physique à l'université. DES en Technologie de l'Education. Universités de Liège et Namur.

Barhrick, H.P.(1984). Semantic memory content in permastore : 50 years of memory for Spanish learned in school. Journal of Experimental Psychology : Genereal, 120, 1-29.

Bloom, B. et al. (1956), Taxonomy of educational objectives. Handbook 1 : Cognitive domain, New-York, McKay, traduit par M. Lavallée sous le titre « Taxonomie des objectifs pédagogiques », Montréal, Education nouvelle (1969).

Bloom, B.S. (1972). L'innocence en pédagogie, Education - Tribune Libre, 135, 14-20.

Brunswick, E. (1943) Organismic achievement and Environment Probability, Psychological Review, 50, 255-272.

Choppin, B.H. (1975), Guessing the answer on objective tests, British Journal of Educational Psychology, 45, 206-213.

Cronbach, L. & Meehl, P. (1955) Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Cross, L. & Frary, (1977), An empirical test of Lord's theoretical results regarding formula scoring of multiple choice tests, Journal of Educational Measurement, vol. 14, 313-321.

De Finetti, B. (1965), Methods for discriminating levels of partial knowledge concerning a test item, British Journal of Mathematical and Statistical Psychology, 18, 87-123.

Ebel, R.L. (1969), Expected reliability as a function of choices per item, Educational and Psychological Measurement, 29, 565-570.

Gardner, H. (1996) Les Intelligences Multiples (traduit de Multiple Intelligences, 1993). Paris : Retz.

Gilles, J.L. (1999), Apports des mesures métacognitives lors d'un test de compréhension d'un article scientifique, in C. Depover & B. Noël (Eds), Approches plurielles de l'évaluation des compétences et des processus cognitifs, Proceedings of the 12th ADMEE Conference, Mons; UMH-FUCAM, 19-30.

Green, D. R. (1998). Consequential aspects of the validity of achievement tests: A publisher's point of view. *Educational Measurement, 17*, 16-19, 34.

Haralambos, M. (1991). Sociology: Themes and Perspectives.

Inizan, A., Le temps d'apprendre à lire, Paris : Bourrelier, 1963.

Leclercq, D. (1986) La conception des QCM. Bruxelles : Labor.

Leclercq D. & Boskin A.(1990), Note taking behavior studied with the help of hypermedia, in Estes, Heene & Leclercq (Eds), Proceedings of the 7th ICTE , Brussels, vol 2, 16-19, Edimburgh : CEP Consultants.

Leclercq, D., Simon, F., Marotte, P., Lacaille, C. (2002). Former des étudiants de première candidature universitaire à des compétences transversales : lesquelles et comment ?, 2e Congrès des chercheurs en éducation, Louvain-la-Neuve, Mars.

Leclercq, D. (Ed.) (2003), Diagnostic cognitif et métacognitif au seuil de l'université. Le projet MOHICAN mené par les 9 universités de la Communauté française Wallonie Bruxelles. Liège : Editions de l'Université de Liège

Leclercq, D., Detroz, P., Dupont, C., Gilles, J-L. (2003). Changer de méthode d'étude suite aux feedbacks : l'opréation RESSAC, in Leclercq, D. (Ed.) (2003), Diagnostic cognitif et métacognitif au seuil de l'université. Liège : Editions de l'Université de Liège, 155-170.

Leclercq, D. (2005). Edumétrie et docimologie pourpraticiens chercheurs. Editions de l'université de Liège.

Linn, R. L. (1998). Partitioning responsibility for the evaluation of the consequences of assessment programs. *Educational Measurement, 17,* 28-30

Luh, C.W. (1922). The conditions of retention. Psychol. Monograph, 31, 142, 401-410.

Lune, S., Parke, C. S., & Stone, C. A. (1998). A framework for evaluating the consequences of assessment programs. *Educational Measurement, 17,* 24-28.

Messick, S. (1988, 3° edition). Validity. In Linn R. (Ed), Educational Measurement. NY : Macmillan

Moss, P. A. (1998). The role of consequences in validity theory. *Educational Measurement, 17,* 6-12.

Pieron, H. (1963), Examens et docimologie, Paris : Presses Universitaires de France.

Reckase, M. D. (1998). Consequential validity from the test developer's perspective. *Educational Measurement, 17,* 13-16.

Schurwirth, L.,(1998) An approach to the assessment of medical problem solving : Computerised Case-based Testing, Ph. D., Rijksuniversiteit Maastricht : Datawyse Universitaire Pres, 1998.

Shufford, E., Albert, A. & Massengill, N.E. (1966), Admissible probability measurement procedures, Psychometrika, 31, 125-145.

Smedslund, J. (1997), The forgotten variable of understanding . Cahiers de Psychologie Cognitive – Current Psychology of Cognition, 16 (1-2), 217-221.

Taleporos, E. (1998). Consequential validity: A practitioner's perspective. *Educational Measurement, 17,* 20-23, 34.

Tversky, A. (1964) On the Optimal Number of Alternatives at a Choice Point', Journal of Mathematical Psychology 1(2): 386-391.

Van Naerssen, R.F (1962), A scale for the measurement of subjective probability, Acta Psychologica, 20, 2, 159-166.

Yen, W. M. (1998). Investigating the consequential aspects of validity: Who is responsible and what should they do? *Educational Measurement, 17,* 5.

DE LA NECESSITE D'UN DOUBLE REGARD. LE CAS DE L'EVALUATION D'UN MOYEN D'ENSEIGNEMENT DES MATHEMATIQUES EN SUISSE ROMANDE.

Dr. Chantal Tièche Christinat, IRDP Neuchâtel et HEP, Lausanne.

Mots clés : double approche, complémentarité, moyen d'enseignement.

Résumé: Les moyens d'enseignement romand des mathématiques introduits progressivement dès 1997 dans les classes ont fait l'objet de deux évaluations portant sur les degrés primaires, répondant ainsi au besoin d'un suivi scientifique de cet enseignement de 1P à 4P et à la demande d'évaluation des compétences mathématiques des élèves. La première recherche consiste en une approche qualitative et longitudinale de l'enseignement des mathématiques. Basée sur l'observation de classes, le recueil d'entretiens et de questionnaires, elle permet d'appréhender la réalité scolaire, les gestes et attitudes des enseignants ainsi que les adaptations faites aux propos didactiques qui fondent ces moyens d'enseignement. Une deuxième recherche de type processus-produit (Mathéval) cible, par un survey effectué à la fin du 2ème et du 4ème degré primaire, les compétences des élèves et parvient à établir la réussite des élèves à différents items relevant du domaine numérique et géométrique. Un questionnaire d'accompagnement destiné aux maîtres permet de compléter les données obtenues.

Afin d'illustrer la complémentarité de notre approche, nous exposerons quelques résultats remarquables de Mathéval et croiserons ces résultats avec ceux obtenus soit lors des entretiens ou questionnaires, soit dans les observations des leçons. La complémentarité illustrée de la sorte recouvre différents caractères : parfois les données sont homogènes et permettent d'affiner la compréhension du lien entre le moyen d'enseignement et les compétences des élèves. Dans d'autres cas, la complémentarité est divergente et les pratiques déclarées ou les pratiques effectives ne parviennent pas à expliquer les résultats des compétences. La perspective d'un pilotage scientifique semble plus clairement dévolue à l'étude longitudinale, puisque les résultats de Mathéval permettent de questionner les moyens, mais ne fournit guère d'indications pour améliorer la mise en œuvre. La complémentarité des deux études montre également l'importance à étudier les pratiques enseignantes et les transpositions dont elles sont issues.

Introduction

En 1997, l'école suisse romande a introduit progressivement des nouveaux moyens d'enseignement des mathématiques dans tous les degrés. Les sept cantons concernés appelaient cette innovation, rendue nécessaire en raison de l'obsolescence des anciens moyens conçus dans les années 1970-80. Souhaitée par les enseignants, mais également par les mathématiciens et les didacticiens, cette innovation concerne à la fois le degré primaire et secondaire jusqu'à la fin de la scolarité obligatoire.

Pour l'ensemble du corps enseignant primaire 1P-4P, ce moyen est officiel et constitue un matériel scolaire obligatoire qui est à disposition des élèves et des enseignants. Chaque enseignant reçoit un matériel destiné à sa pratique professionnelle de l'enseignement des mathématiques et ses élèves disposent gratuitement de livre et fichiers. Des commentaires didactiques explicitant les fondements épistémologiques de l'innovation complètent ce matériel.

De fait, l'innovation s'effectue sur trois plans : le contenu mathématique, la théorie de l'apprentissage et la didactique des mathématiques. Le constat d'un processus transpositif trop important entre les pratiques mathématiciennes et enseignantes (Jaquet, 2000) et d'un vieillissement culturel (Joshua et Dupin (1993) entraîne l'abandon de l'enseignement des mathématiques ensemblistes et d'un chapitre très controversé portant sur la numération en bases différentes de 10. Cependant le contenu mathématique curriculaire est quant à lui presque inchangé. Les aspects numériques (étude du nombre, du champ conceptuel de l'addition et de la multiplication) restent centraux, alors que les aspects liés à la mesure, à l'espace et aux formes géométriques acquièrent un poids plus important qu'auparavant. Inscrits dans l'évolution des sciences et du développement des théories d'apprentissage/enseignement, ces moyens satisfont aux exigences de la recherche de sens des algorithmes et autres connaissances mathématiques enseignées et instaurent les démarches de recherche et les inévitables tâtonnements et erreurs qui en découlent comme un principe essentiel. Les lignes directrices de cette nouvelle approche de l'enseignement/apprentissage des mathématiques se réfèrent aux travaux récents des didacticiens français (Charnay et Mante, 1996 ; Vergnaud, 1991) et s'opérationnalisent dans la résolution de problèmes. Les principes qui constituent la méthodologie articulent les finalités de l'enseignement mathématique (acquérir des démarches de pensée), les démarches d'apprentissage (confronter l'élève à des situations-problèmes où il rencontre de véritables obstacles à sa mesure), les concepts de développement (la pensée mathématique s'élabore longuement) et les techniques de résolution. L'ensemble de ces propositions recouvrent les recommandations officielles des cantons (CDIP ; 1998) en matière d'enseignement des mathématiques. Les options pédagogiques s'inscrivent dans la continuité des apprentissages constructivistes précédemment mis en place autour de l'action, de l'erreur et de la différenciation, et s'enrichissent d'une orientation socioconstructiviste d'inspiration vygostkyenne soulignant l'importance de la médiation verbale et de la construction

sociale des connaissances. Leur application en classe nécessite un travail de groupes et donne à la parole un aspect primordial attribuant de fait une part importante aux interactions langagières entre pairs,

La théorie des situations didactiques (Brousseau ; 1996) constitue les racines profondes de l'innovation et est instituée sous des formes diverses dans les nouveaux moyens romands d'enseignement. La situation didactique est dès lors conçue comme la situation d'enseignement, qui voit le couple enseignant/enseigné relié par un contrat qui stipule que l'enseignant donne une tâche à l'élève dans le but de tester ou accroître ses connaissances. Cette situation didactique inclut la situation adidactique qui prend forme lorsque le problème soumis par le professeur devient le problème de l'élève et n'appartient plus dès lors au professeur. Le principe de dévolution est extrêmement important dans la situation didactique, car il génère une série de règles dont les commentaires méthodologiques qui accompagnent les moyens d'enseignement font partiellement écho. Afin de construire de nouvelles connaissances, le problème donné à l'élève doit à la fois mobiliser ses connaissances et lui résister (Vergnaud, 1991). Le respect dans la classe de cette condition première entraîne une application de la différentiation pédagogique et vient ainsi renforcer les aspects constructivistes et socioconstructiviste de l'apprentissage que l'école entend instaurer. Par ailleurs, la situation adidactique modifie le contrat pédagogique en cours. Le professeur n'a pas à transmettre de nouvelles connaissances, mais doit créer ou choisir des situations pertinentes selon le degré de connaissances des élèves, le curriculum scolaire et les savoirs mathématiques en jeu.

Évaluation des moyens d'enseignements pour le degré primaire

Afin d'évaluer à sa juste mesure l'introduction de ces moyens d'enseignement et leurs effets sur la réussite des élèves en mathématiques, l'Institut de Recherche et de Documentation Pédagogique (IRDP) met en place un suivi de cette innovation pour le degré primaire (1-4) au moyen de deux études dont nous chercherons à démontrer la nécessaire complémentarité.

1. Un suivi longitudinal qui démarre dès l'entrée en vigueur ayant pour fonction de piloter et d'appréhender les adaptations entreprises par les enseignants dans leurs classes. (Tièche Christinat et Delémont, 2006)

2. Un bilan des compétences en mathématiques des élèves après 4 années d'utilisation des moyens (2P : passation mai 2002; 4P passation mai 2004) nommé Mathéval (Antonietti, 2003 et 2005)

Les deux études ont pour mission de décrire et d'évaluer les moyens mathématiques 1P-4P utilisés dans les classes romandes. La première pourrait être qualifiée du titre de recherche de type processus et la seconde de type produit.

Le suivi scientifique a pour objet d'observer ce qui se passe au sein même de la classe afin de pouvoir appréhender la réalité scolaire de l'innovation et les adaptations dont les moyens font l'objet. De type longitudinal et qualitatif, cette étude se décline en deux parties: une observation biannuelle de leçons de mathématiques dans quelques classes de la Suisse Romande et une récolte diversifiée de propos portant sur l'innovation et ses conséquences (entretiens réguliers et questionnaires remis aux enseignants visités, discussions orientées entre collègues de même établissement). La recherche longitudinale par l'observation en classe qu'elle effectue cherche à prendre une mesure de l'input afin de dégager

- Les gestes professionnels et les caractéristiques des adaptations introduites dans la pratique: activités mathématiques, déroulement des leçons, organisation générale du travail.

- Le profil des enseignants et son évolution à partir de leurs attitudes à l'égard du nouvel enseignement

- L'évolution et l'adaptation des moyens d'enseignement.

Mathéval se propose d'étudier les effets de l'enseignement introduit en 1997 sur les compétences des élèves, et des liens qui peuvent exister entre ces dernières et les pratiques décrites par les enseignants. Cette enquête comporte deux volets: des épreuves mathématiques établies pour des élèves de 2P et 4P et un questionnaire destiné aux maîtres des classes testées.

Mathéval pose trois questions essentielles qui permettent une mesure de l'output.

- Les objectifs du plan d'étude sont-ils atteints ?

- Les compétences mathématiques construites et mobilisées sont-elles transférables à d'autres situations-problèmes ?

- Quels sont les paramètres (élève, enseignant) qui influencent les résultats obtenus ?

Afin de pouvoir évaluer l'atteinte des objectifs inscrits dans le plan d'étude, et à défaut d'une comparaison fiable avec d'anciennes données, Mathéval a opté lors de l'épreuve destinée aux élèves de 4P d'établir un seuil de compétence. Ainsi un caractère normatif est conféré à cette étude. De plus, une analyse détaillée des résolutions de problèmes lui attribue également une valeur descriptive et compréhensive, qui pourrait se révéler utile dans le cadre d'une évaluation diagnostique.

Données et échantillon

Le suivi longitudinal a permis l'observation de 27 cohortes d'élèves durant 4 ans à raison de deux fois l'an pour la majorité des cas. 197 observations de 45 minutes ou plus ont pu être analysées, dont 43 en 1P, 52 en 2P et 51 en 3P et en 4P également. 95 entretiens récoltés auprès de 58 enseignants et 82 questionnaires constituent notre base concernant les pratiques déclarées et les attitudes des enseignants.

Tableau 1 : type et nombre de données récoltées par l'étude du suivi longitudinal dans les classes

	1P	2P	3P	4P	total
Observations	43	52	51	51	197
Cohortes d'élèves	27	27	27	27	27
Questionnaires	22	15	23	22	82
Enseignants interviewés • 1 fois • 2 fois ou plus	14 17		13 14		27 31

L'échantillon de Mathéval est constitué de 140 classes répartis dans les 7 cantons romands. Au total, 1912 et 2252 épreuves ont été analysées. En 2P, les élèves avaient 7 problèmes individuels et 4 problèmes en duos à résoudre et en 4P, les problèmes étaient au nombre de 14 par élève. Chaque titulaire des classes s'est vu remettre un questionnaire. Certains maîtres fonctionnant en duo pédagogique ont rempli séparément un questionnaire, d'où le nombre plus élevé de questionnaires reçus que de classes testées.

Tableau 2 : type et nombre de données récoltées par l'étude Mathéval

	Mathéval 2P	Mathéval 4P
Classes	140	140
Élèves	1912	2252
Problèmes testés	22	56
Problèmes par élève • résolution individuelle • résolution collective (2-3 élèves)	7 4	14
Enseignants interrogés	145	151

Nous aimerions souligner de prime abord l'indépendance de chaque recherche. En effet, les méthodologies choisies pour chacune sont fonction des questions de recherches posées. Les populations étudiées sont différentes, aucun enseignant n'a participé aux deux recherches. De fait chaque étude fait l'objet de rapports scientifiques séparés et constitue une approche spécifique de l'évaluation d'un moyen d'enseignement. Chaque enquête comporte plusieurs volets conçus pour fournir des éléments explicatifs internes à chaque recherche. Ainsi dans les deux études, les questionnaires permettent de dégager quelques facteurs incidents sur les résultats des élèves et sur les gestes enseignants.

Les deux recherches ont toutefois adopté un principe épistémologique commun: aussi bien l'évaluation des compétences des élèves, que les questionnaires soumis aux enseignants et les observations prennent en compte les changements didactiques intervenus dans l'enseignement/apprentissage et les interrogent. Il s'ensuit des items mathématiques basés sur la résolution de problèmes, mettant en œuvre des champs étudiés et inscrits au plan d'étude ainsi que des compétences transversales. Les questionnaires des deux recherches sont composés de thématiques communes autour de la mise en œuvre des moyens d'enseignement et l'observation est analysée en fonction du cadre didactique théorique proposé aux enseignants.

Analyse des données et complémentarité

Partant des résultats et les analyses des données de l'étude Mathéval, qui ont pour caractéristiques d'être statistiquement validées et bénéficiant déjà d'un premier jeu de données que nous pouvons croiser, nous n'esquisserons dans les lignes ci-dessous quelques résultats remarquables et les articulerons à différentes données recueillies soit au moyen de questionnaires ou entretiens, soit lors de l'observation des leçons de mathématiques.

Quelques résultats de Mathéval

En 2P, selon Mathéval la quasi totalité des élèves (3% seulement d'échec) peut résoudre au moins partiellement les problèmes additifs. La construction du nombre semble également être satisfaisante. Par contre, les résultats obtenus en 2P dans le domaine de l'espace sont nettement moins satisfaisants que ceux du domaine numérique. En 4P, la tendance s'inverse et les objectifs minimaux établis par des experts sont atteints par seulement trois élèves sur dix. Le taux de réussite fixé en référence au moyen des documents officiels par les experts est ainsi plus bas qu'espéré. Si ces résultats faibles déconcertent, l'analyse détaillée des copies révèle comme en 2P des élèves astucieux pleins d'intuition et de créativité. Face à des situations complexes, les élèves de 2P et de 4P parviennent à saisir les données pertinentes du problème et tentent des démarches de résolution très diversifiées qui

malheureusement n'aboutissent pas toujours à une solution adéquate, souvent par manque de contrôle de la réponse ou abandon en cours de recherche. Nous relevons également en 4P des difficultés importantes à conduire des raisonnements logiques.

Croisement des données et mise en perspective

Les analyses des pratiques déclarées dans Mathéval confortent et expliquent partiellement de tels résultats. Les enseignants de 2P et 4P disent accueillir les moyens avec un certain enthousiasme, voire pour le moins une certaine satisfaction. L'obsolescence des anciens moyens est soulignée et le modèle constructiviste de l'apprentissage est approuvé. Par contre, plusieurs enseignants font le reproche "d'un manque de structure et de précision et regrettent qu'une progression explicite ou qu'un ordre chronologique des activités ne soient pas offerts" (Antonietti, 2003, p. 128). Il se pourrait que cette absence de guide les ait conduits à omettre ou à négliger les activités liés à l'espace et aux figures géométriques, qui forment les derniers modules du livre et qui comportent par ailleurs en 1P-2P relativement peu d'activités. Les entretiens et questionnaires recueillis lors de l'étude qualitative relèvent également un inconfort des maîtres du premier cycle quant à la mise en œuvre des modules portant sur l'espace et les figures géométriques. En deuxième cycle, ce malaise semble s'étendre également au module de la mesure ainsi qu'aux activités de raisonnement, mettant en œuvre des compétences transversales.

La concordance des deux études et des pratiques déclarées nous incite à interroger les contenus mathématiques qui s'avèrent moins bien notés que la modification du paradigme d'apprentissage. Les activités numériques demeurent très clairement les objectifs prioritaires et centraux de l'enseignement/apprentissage des mathématiques au niveau primaire. En comparant le nombre d'activités observées au nombre d'activités proposées par l'étude des livres du maître, nous avons dégagé une surreprésentation d'activités qui touchent aux problèmes additifs, dans le premier degré au détriment des activités relevant de la construction du nombre et au deuxième degré une surreprésentation des activités du module portant sur les problèmes multiplicatifs et sur la construction du nombre. Ces résultats soulignent l'attention que les enseignants 1P-2P portent aux nombres naturels et aux opérations. Les activités portant sur l'enseignement/apprentissage de l'espace et des figures géométriques sont peu fréquentes, mais leur taux semble toutefois conforme à celui relevé dans les livres du maître. Au deuxième cycle, les enseignants ont en commun d'offrir peu d'activités participant des compétences transversales lors de nos visites, tandis que les activités portant sur l'étude du champ conceptuel de la multiplication et de l'espace sont surreprésentées.

L'infériorité des compétences des élèves dans le domaine numérique par rapport aux compétences attendues et une faible progression des élèves dans ce domaine interroge. Les propos des enseignants dans les deux études se montrent très critiques à

l'égard des activités portant sur ce domaine. Ils regrettent l'absence de calculs numériques simples hors problèmes et soulignent l'absence d'exercices de drill et de réinvestissement de techniques opératoires. Le recours à des moyens complémentaire compense en partie ce manque ; 51% des enseignants avouent utiliser des fiches de calculs et d'opérations et 28% citent les carences de la méthodologie au niveau du calcul et des algorithmes comme raisons d'employer des moyens complémentaires. Ainsi pouvons-nous prétendre que les pratiques enseignantes compensent la faiblesse évoquée des moyens. Les observations en classes confirment l'usage de fiches d'entraînement parallèlement à l'utilisation des moyens officiels. En effet 10% des activités observées proviennent d'autres sources et poursuivent pour la plupart un objectif d'exercisation ou de réinvestissement des notions étudiées (algorithmes, table d'addition et de multiplication). La faiblesse avéré ne trouve guère d'explication dans ce premier croisement de données, mais renvoie à la nécessité de chercher d'autres perspectives.

Les résultats plus descriptifs de Mathéval laissent entrevoir une absence de contrôle de la réponse qui trouve une explication dans la gestion didactique de la leçon observée par le suivi longitudinal. En effet, les phases d'action qui permettent aux élèves de se confronter seuls, sans relance du maître, se réalisent durant un bref moment de l'activité de résolution. Les enseignants très rapidement soutiennent et appuient la résolution en cours par des relances et des validations, et les élèves attendent voire demandent cet appui. L'analyse qualitative est dans ce cas contrastante, puisque l'efficacité des moyens remise en cause par la faible compétence des élèves renvoie non pas aux contenus enseignés et incriminés, mais aux gestes didactiques que les enseignants mettent en place.

Les enseignants soulignent dans le questionnaire Mathéval leur credo à l'apprentissage constructiviste. Ils accordent aux activités proposées dans les moyens un statut didactique central, et considèrent qu'elles peuvent engendrer chez les élèves du sens aux connaissances qu'ils manipulent. Le respect de ce paradigme d'enseignement/apprentissage questionne la posture enseignante et les entretiens lors du suivi esquissent chez certains un certain malaise professionnel. En effet, quelques maîtres se considèrent transformés en animateurs d'activités et de jeux et non en enseignants et perçoivent un changement conceptuel du métier. Les observations en classe confirment la gêne énoncée et montrent que la phase d'institutionnalisation, qui bien que plus visible en 4P qu'en 2P apparaît peu investie en fréquence et en temps par les enseignants. Le suivi des pratiques enseignantes permet en partie d'éclairer ce point. L'analyse du déroulement de la leçon souligne la difficulté exprimée lors des entretiens à mettre en pratique les phrases de formulation et d'institutionnalisation des savoirs et à donner une place réelle à la transmission des savoirs. La difficulté de la gestion du temps didactique semble être un facteur explicatif, mais sont évocables également l'amalgame entre phase de formulation et institutionnalisation, l'assujettissement voire la réduction de l'institutionnalisation des savoirs à l'évaluation de la réponse en terme de juste ou faux, la compréhension peut-être partielle ou partiale des commentaires didactiques accompagnant les moyens d'enseignement, l'importance du facteur horaire, etc. Les types d'organisation générale du travail des

élèves rencontrées lors des visites de classe ne facilitent pas tous par ailleurs la mise en place de ces deux phases propices à l'émergence et à la transmission des savoirs culturels reconnus. L'organisation en plan de travail bimensuel ou mensuel par exemple contribue à éloigner le processus de résolution du processus de formulation collective et d'institutionnalisation.

Parmi les points positifs des nouveaux moyens, l'aspect ludique, la variété des activités, la possibilité de développer des facultés de recherche et la collaboration et le travail en groupe sont cités. Les discours à ce sujet sont homogènes et divers liens peuvent être établis entre les questions. Ainsi les enseignants disent apprécier le module "apprendre à conduire raisonnement", puisque celui-ci met en jeu des activités de recherche et obligent les élèves à conduire des raisonnements logiques et à faire appel à des compétences transversales. Toutefois, les résultats des élèves nuancent ou interrogent ces résultats, et mettent en évidence un certain nombre de décalages entre ce qui est affirmé être fait et ce qui est réalisé. En effet, les compétences transversales en 4P sont peu élevées. Très peu d'élèves sont capables de recherche exhaustive. La plupart donnent l'impression de ne savoir quoi faire de leurs connaissances et savoirs. La difficulté majeure vient d'un manque de représentation du problème. De plus, une situation-problème nécessitant l'organisation d'informations pour trouver de nouvelles données dans un contexte non numérique (relations d'ordre) n'est pas mieux réussie en 4e qu'en 2e. Le suivi en classe apporte quelques éléments explicatifs qui contredisent les réponses aux questionnaires. En effet, le module "apprendre à conduire raisonnement" et les problèmes de type non numérique et/ou nécessitant des compétences transversales qu'il propose est peu à peu abandonné au profit des modules relatifs aux opérations. Par ailleurs, la répartition des activités selon la modalité "élèves" indique que les enseignants de 1P-2P proposent moins d'activités pour un élève que ne le suggère le livre du maître, alors que ceux de 3P-4P en propose plus. Ce renversement de tendance affecte également les activités conçues pour deux élèves. Ceci laisse supposer que si les bases constructivistes sont reconnues et admises par les enseignants, le travail en duo est de moins en moins privilégié et la construction des connaissances par la médiation des pairs ne constitue pas une base d'enseignement privilégié par les enseignants du deuxième cycle.

Discussion

Les deux études ont pour but d'évaluer les moyens d'enseignement des mathématiques. Leur mise en perspective et leur articulation permet de répondre en partie à la réflexion critique émise par Bru (2002), en ce sens que les deux études s'articulent et fournissent des dimensions explicatives supplémentaires. En effet, la démarche qualitative appuie dans certains cas la relation établie entre la modalité d'enseignement et les résultats des élèves. Dans certains cas de la complémentarité pourrait être qualifiée de divergente ou contrastante, dans le sens où les informations fournies par les études ne confortent pas les résultats obtenus ou les liens établis entre les différentes études. Cette complémentarité divergente met très souvent dans notre cas en opposition les données de l'observation à celles provenant de l'étude des compétences, mais nous constatons également l'existence de décalage entre les discours des enseignants et les pratiques enseignantes.

L'étude Mathéval questionne les moyens d'enseignement en s'intéressant aux points de vue des enseignants et aux compétences des élèves. Les pratiques déclarées et les effets mesurés des moyens d'enseignement sur les élèves en terme de compétence sont analysés par Mathéval. D'une certaine manière, nous pouvons dire que l'impact de l'innovation est perçu globalement. Les données du suivi relèvent quant à eux d'un dispositif d'observations des pratiques pédagogiques. Si ce type d'analyse est relativement lourd, comme le signale Genelot et Suchat (2000), les données recueillies auprès des deux populations et leur articulation permettent de questionner la pertinence des moyens. Ainsi, nous l'avons vu plus haut, les données du suivi peuvent également jouer ce rôle et étayer, affiner et préciser les conclusions de l'étude statistique et les croisements des données effectuées au sein de Mathéval.

Le suivi de l'enseignement par l'éclairage qu'il donne des pratiques enseignantes dans la classe permet de saisir l'importance des gestes professionnels et parallèlement les effets de la transposition. La complémentarité par ailleurs nous semble également se jouer aussi à un autre niveau. Les recherches entreprises par l'IRDP font parties des évaluations du système éducatif et ont pour fonction non seulement de renseigner le système éducatif et politique à propos d'un objet précis, mais également de lui permettre un pilotage scientifique. Les données obtenues par la recherche Mathéval permettent de questionner les moyens d'enseignement, voire même de les rendre responsables de la faible performance des élèves. Cependant, de tels résultats n'apportent guère de réponses aux adaptations à fournir pour améliorer si nécessaire la mise en œuvre. Dès lors, il nous semble indispensable pour piloter le système éducatif de favoriser des recherches qualitatives portant sur les gestes enseignants parallèlement aux recherches quantitatives portant sur l'appréhension et l'effet du système chez les maîtres et les élèves.

La recherche de type input par ses caractères cliniques offre des facteurs explicatifs permettant un pilotage du système. Les données que nous avons présentée confirment

les résultats d'autres études, dont Delémont (2006) rappelle les caractéristiques« C'est en fonction de son expérience, de ses compétences, de ses intérêts, de sa formation et de sa conception des mathématiques et des processus d'apprentissage, que le maître sélectionne les activités qui lui conviennent et qui font sens pour lui. Chacun n'enseigne ainsi pas de la même manière, et ces différences peuvent se traduire, entre autres, dans les performances inégales des classes. »

Références

Antonietti, J.-P. (éd.). (2003). *Évaluation des compétences en mathématiques en fin de 2ème année primaire: résultats de la première phase de l'enquête Mathéval.* Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP) (03.2).

Antonietti, J.-P. (éd.). (2005). *Évaluation des compétences en mathématiques en fin de 4ème année primaire: résultats de la deuxième phase de l'enquête Mathéval.* Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP) (05.3).

Bru, M. (2002). Pratiques enseignantes : des recherches à conforter et à développer. Revue Française de Pédagogie, 138, 63-73.

Brousseau, G. (1996). *Fondements et méthodes de la didactique des mathématiques. In: J. Brun (Ed.) : Didactiques des mathématiques*. Lausanne: Delachaux et Niestlé.

Charnay, R. & Mante, M. (1996). *Préparation à l'épreuve de mathématiques du concours de professeur des écoles.* Tome 1. Paris : Hatier.

Delémont, M. (2006). *L'épistémologie des enseignants : quel impact sur les procédures des élèves en mathématiques.* Neuchâtel : Institut de recherche et de documentation pédagogique (IRDP) (06.1).

CDIP (Conférence des directeurs de l'instruction publique). (1988). Freiräume - Richtlinien - Treffpunkte: Mathematikunterricht während der obligatorischen Schulzeit / Espaces de liberté - lignes directrices - points de convergence: l'enseignement des mathématiques durant la scolarité obligatoire. *Dossier 49*. Berne : EDK.

Genelot, S. & Suchat, B. (2000)

Jaquet, F. (2000). Éditorial : Moyens d'enseignement de mathématiques de Suisse romande: défis et nécessités. *Maths-École*, 190, 1-4.

Tièche Christinat, C. & Delémont, M. (2006). P*ratiques et discours. Le nouvel enseignement des mathématiques 1P-4P sous la loupe*. Neuchâtel : IRDP:

Vergnaud, G. (1991). La théorie des champs conceptuels. *Recherches en didactiques des mathématiques*, 10/2.3, 133-170.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

DEVELOPPEMENT D'UN OUTIL POUR LA CONSTRUCTION ET L'EVALUATION DE STRATEGIES DE RESOLUTION DE PROBLEMES EN MATHEMATIQUES A L'AIDE DU TESTING ASSISTE PAR ORDINATEUR.

Gilbert Busana, Université du Luxembourg
Romain Martin, Université du Luxembourg

Résumé :

Depuis plusieurs années, différents acteurs du système éducatif luxembourgeois (Task-Force *e*Luxembourg, Ministère de l'Education Nationale, Centre de Technologies de l'Education) augmentent les efforts envers une intégration des technologies de l'information et de la communication (TIC) dans les apprentissages scolaires. On peut constater que bien des écoles disposent de matériel informatique pour utiliser les TIC, mais les enseignants relèvent un grand besoin en concepts d'intégration des TIC (Conrad, 2002). Le présent projet vise à développer (et à évaluer) un tel concept à travers un outil informatisé pour la compréhension et la résolution de problèmes en mathématiques. Cet environnement, nommé CAMPUS (Computer Assisted Mathematical Problem Understanding and Solving), est basé sur le mode de représentation de la ligne numérique (Klein, Beishuizen, & Treffers, 1998; Petitto, 1990) et de graphes (diagramme des flèches). Il est déployé par une plateforme de Testing Assisté par Ordinateur (TAO, http://www.tao.lu). À travers les fonctionnalités de cette plateforme, CAMPUS peut enregistrer une multitude d'indicateurs comportementaux de l'apprenant et permet ainsi une meilleure analyse des stratégies de résolution mises en place voire une évaluation des compétences sous-jacentes.

Présentation

Les caractéristiques principales de cet environnement sont donc l'utilisation de l'ordinateur en tant qu'environnement (1) pour le développement de compétences de résolution de problèmes en mathématiques ainsi que (2) pour l'évaluation des compétences de résolution de problèmes.

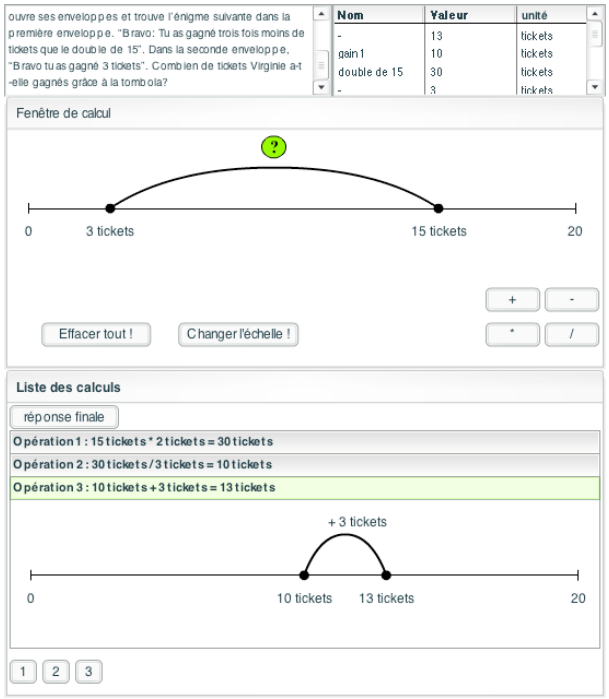


Figure 1 : Copie d'écran de CAMPUS

Le recours au mode de représentation de la ligne numérique (cf. figure 1, *fenêtre de calcul*) et de graphes permet aux apprenants de construire des représentations analogiques (Shepard, 1978) des différentes opérations à effectuer, ce qui devrait les aider à développer l'intuition numérique ou *number sense* (Dehaene, 1997) dans des situations concrètes. Un graphe (*liste des calculs*) résume toutes les étapes de calcul. Ainsi, l'apprenant dispose à tout moment d'une "illustration" des calculs qu'il est en train d'effectuer. Cet outil est intégré dans la démarche de l'enseignant et ne constitue donc pas un outil à part. L'environnement d'apprentissage CAMPUS permet à l'enseignant d'analyser plus efficacement les différentes étapes du développement de stratégies de résolution faites par ses élèves et lui offre la possibilité de donner des feedbacks plus adaptés afin de pouvoir mieux guider le processus d'apprentissage. Les feedbacks donnés par l'outil et par l'enseignant dans une telle situation d'apprentissage incitent chez l'apprenant une réflexion méta-cognitive sur sa démarche et ses stratégies et devraient l'amener à construire des modèles mentaux complexes. La conception de l'outil s'inspire donc explicitement de travaux de

recherche concernant l'apprentissage qui ont été réalisés dans le domaine de la psychologie cognitive et qui sont combinés d'une manière novatrice dans le présent projet.

L'environnement d'apprentissage CAMPUS n'impose aucune restriction dans le processus de résolution de l'apprenant, mais évite que celui-ci se perde dans des chemins de résolution manifestement faux (par exemple : addition d'objets de classes différentes). De cette façon, l'architecture CAMPUS est un outil cognitif qui aide l'apprenant dans la structuration de ses pensées en lui communiquant quels calculs ne sont mathématiquement ou logiquement pas permis. Or, l'outil ne suggère aucune voie précise vers la solution. Au contraire, chaque action logiquement ou mathématiquement correcte sera acceptée par le système. Par conséquent, CAMPUS n'est pas un outil d'exercisation (drill and practice), mais propose un environnement d'apprentissage à l'élève où celui-ci peut développer des stratégies de résolution de problèmes dans des situations complexes. CAMPUS se base sur une approche socio-constructiviste. En effet, l'outil permet de résoudre les problèmes de manière individuelle ou en groupe, bien que le travail en équipe soit désirable dans une telle approche pédagogique (Webb, 1994; Yadrick, Regian, Connolly Gomez, & Robertson Schule, 1997). Le processus d'apprentissage des élèves est supporté d'un côté par l'outil et de l'autre côté par l'enseignant. L'élève se trouve donc constamment en situation d'évaluation formative. L'apprenant et son processus d'apprentissage sont au centre des activités scolaires (Tardif, 1998) et le rôle de l'enseignant est de superviser et supporter l'élève dans un tel environnement d'apprentissage (Hudson, 1997; Tardif, 1998).

De plus, l'utilisation de la ligne numérique (pour les opérations à faire) et de graphes (pour représenter les étapes de résolution) devrait favoriser le développement de modèles mentaux qui incluent des représentations analogiques et supportent le processus de résolution de ce type de problèmes, même en dehors de l'utilisation de la plateforme informatisée. Nous supposons (hypothèse reste à être vérifiée lors de l'expérimentation) que l'apprenant va développer, de par cette expérience avec la plateforme (qui évite l'expérimentation avec des calculs qui ne font pas sens), de solides compétences dans le domaine de la résolution de problèmes mathématiques et ceci, au moins partiellement, sur base de processus d'apprentissage non-explicites, tels que décrits par les modèles connexionnistes (Spitzer, 2000). Ces modèles connexionnistes montrent qu'un apprentissage significatif peut être acquis d'une manière non-explicite si l'apprenant est placé dans un environnement qui structure son comportement selon certaines régularités (même si ces régularités ne sont pas enseignées de manière explicite).

La présentation inclura une démonstration de CAMPUS avec plusieurs exemples de problèmes, l'analyse des premiers résultats empiriques de l'utilisation de CAMPUS avec des élèves, ainsi qu'une discussion sur la pertinence des différents indicateurs comportementaux.

Bibliographie

Conrad, C. (2002). *Der Computereinsatz in der Grundschule in Luxemburg*. ISERP, Walferdange.

Dehaene, S. (1997). *La bosse des maths*. Paris: édition Odile Jacob.

Hudson, B. (1997). Group work with Multimedia in Mathematics: the role of the technology and of the teacher. *British Journal of Educational Technology, 28*(4), 257-270.

Klein, A. S., Beishuizen, M., & Treffers, A. (1998). The empty number line in Dutch second grades: Realistic versus gradual program design. *Journal for Research in Mathematics Education, 29*(4), 443-464.

Petitto, A. L. (1990). Development of numberline and measurement concepts. *Cognition and Instruction, 7*(1), 55-78.

Shepard, R. N. (1978). The Mental Image. *American Psychologist, 33*, 125-137.

Spitzer, M. (2000). *Geist im Netz. Modelle für Lernen, Denken und Handeln*. Heidelberg: Spektrum Akademischer Verlag.

Tardif, J. (1998). *Intégrer les nouvelles technologies de l'information: Quel cadre pédagogique?* Issy-les-Moulineaux: ESF éditeur.

Webb, M. E. (1994). Beginning computer-based modelling in primary schools. *Computers Educ., 22*(1/2), 129-144.

Yadrick, R. M., Regian, J. W., Connolly Gomez, C., & Robertson Schule, L. (1997). Dyadic vs. individual practice with exploratory and directive mathematics tutors. *Journal of Educational Computing Research, 17*(2), 165-186.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

L'ÉVALUATION EN MATHÉMATIQUE : À LA RECHERCHE D'UNE CONCEPTUALISATION DE L'ACTIVITÉ ÉVALUATIVE DES ENSEIGNANTS

Maria Palmira Alves, José António Fernandes & Eusébio André Machado,
Université du Minho, Portugal

Mots-clés : Perspectives et pratiques d'évaluation, enseignement élémentaire et secondaire, Mathématique.

Résumé

Depuis les deux dernières décades, le scénario de l'évaluation éducationnel a beaucoup changé. De plus en plus, la société exige l'efficience et l'excellence scolaires et, en même temps, que l'école soit juste et efficace. Au Portugal, l'encadrement normatif, en ce qui concerne l'évaluation a beaucoup changé, surtout depuis 1986[1]. Ce nouveau contexte légitime une investigation aux pratiques d'évaluation des enseignants, pour en évaluer l'impact.

Cette communication présente une étude sur les conceptions et les pratiques d'évaluation des enseignants de Mathématique, dont les questions de recherche: 1- quelles perspectives ont et quelles pratiques d'évaluation mettent en place les enseignants de Mathématique dans leur activité professionnelle?; 2. les perspectives et les pratiques d'évaluation des enseignants de Mathématique sont influencées par *leur expérience*, par *le niveau qu'ils enseignent* et par leur rapport avec le *stage pédagogique*? 177 enseignants de Mathématique du 2.ième[2] et du 3.ième[3] cycle de l'enseignement élémentaire et du secondaire[4], se sont portés volontaires pour remplir un questionnaire sur neuf dimensions de l'évaluation. Les résultats, concernant la première question de recherche, qui nous présenterons, montrent qu'il y a un très fort décalage entre les orientations législatives et programmatiques et les pratiques d'évaluation des enseignants. Même si les programmes de Mathématique demandent clairement que les instruments à utiliser pour évaluer les apprentissages des élèves soient diversifiés, les réponses des enseignants montrent que les tests à papier et crayon sont l'instrument le plus important ou le seul qu'ils utilisent pour évaluer l'apprentissage.

La définition, la nature et les fonctions de l'évaluation

Depuis vingt ans, l'évaluation devient un objet de changements légaux soit au niveau de l'enseignement élémentaire, soit au niveau du secondaire que, d'une façon générale, configurent un nouveau paradigme d'enseignement et d'apprentissage. Un des plus forts éléments de ces dernières orientations est, sans contredit, la rénovation

[1] C'est l'année de la réforme curriculaire au Portugal.
[2] L'âge des élèves: 10-12
[3] L'âge des élèves: 13-15
[4] L'âge des élèves: 16-18

du curriculum (Loi n.º6/2001) qui comprend, entre autres, l'évaluation des apprentissages surtout le caractère formatif, l'importance de l'auto-évaluation régulée, la nécessité de diversifier les instruments et les participants dans le processus de l'évaluation.

Dans ce sens on définit l'évaluation des apprentissages à la suite de Barbier (1990) comme une démarche qui permet de porter un jugement sur les compétences développées et les connaissances acquises par l'élève en vue de prendre des décisions et d'agir. Ce jugement doit s'appuyer sur des informations pertinentes et suffisantes qui donnent un sens à la décision et respecter les valeurs qui orientent l'évaluation des apprentissages : la justice, l'égalité et l'équité ainsi que la cohérence, la rigueur et la transparence (Bélair, 1999).

Dans la mesure où l'évaluation doit être intégrée au processus d'enseignement et d'apprentissage et jouer un rôle de régulation, les situations d'évaluation sont accordées de favoriser le développement des compétences.[5] La régulation faite par l'enseignant (celle que nous intéresse dans cette étude) peut porter sur différents objets d'apprentissage comme la démarche utilisée, le résultat final ou les connaissances antérieures et, selon Depover et Noel (1999), doit favoriser le progrès de l'élève. Pour rendre compte du développement des compétences de l'élève, les enseignants réalisent le bilan de ses apprentissages en considérant l'ensemble des informations pertinentes. Ils interprètent ces dernières à la lumière des attentes de fin du programme de formation (Allal, 1999).

Les actuels programmes de Mathématique sont en accord et conséquents avec l'encadrement normatif. Dans le programme de l'enseignement élémentaire le caractère formatif est valorisé, l'engagement de l'élève dans le processus, l'élargissement du domaine conceptuel, les aptitudes et les attitudes, la diversité des instruments de l'évaluation et la participation de l'élève. Au secondaire, les programmes présentent les mêmes orientations. En effet, il est prévu que l'évaluation ne se limite pas à évaluer le produit, mais aussi le processus de l'apprentissage, en garantissant que l'élève soit le protagoniste en tant qu'élément actif, réflexif et responsable de son apprentissage, et l'importance de la diversification des instruments, en particulier la réalisation obligatoire de rédactions mathématiques (APM, 1998). Il y a donc, dans les programmes de Mathématique, une conception de l'évaluation centrée sur le processus, sur l'élève et sur l'évaluation formative (Abrantes, 2001; Allal, 2000 ; Alves, 2002 ; Hadji, 1994) et non pas sur le produit, sur l'enseignant et sur l'évaluation sumative, interne ou externe. Il s'agit donc, d'un paradigme qui valorise l'évaluation interne et contextuelle, selon les conceptions cognitivistes de l'évaluation (Cardinet, 1994) ou encore, selon Alves e Machado, (2003) d'un paradigme émancipatoire.

Par contre, plusieurs recherches montrent que les pratiques d'évaluation des enseignants de Mathématique se sont attardées au produit (Ponte, Matos, & Abrantes, 1998 ; Ramalho, 2002) et, alors, inscrites dans le paradigme de la performativité (Alves & Machado, 2003), dont l'évaluation est un processus rigide et linéaire et les outils d'évaluation ne favorisent pas la participation active de l'élève à l'évaluation.

---

[5] On entend par compétence un savoir-agir fondé sur la mobilisation et l'utilisation efficaces d'un ensemble de ressources. La compétence est indissociable des situations et des contextes dans lesquels elle est appelée à se manifester. Elle est complexe et évolutive.

Par contre, plusieurs recherches montrent que les pratiques d'évaluation des enseignants de Mathématique se sont attardées au produit (Ponte, Matos & Abrantes, 1998 ; Ramalho, 2002) et, alors, inscrites dans le paradigme de la performativité (Alves & Machado, 2003), dont l'évaluation est un processus rigide et linéaire et les outils d'évaluation ne favorisent pas la participation active de l'élève à l'évaluation.

Ces présumés nous ont permis d'entamer une recherche, centrée sur les perspectives et sur les pratiques de l'évaluation en Mathématique, a partir de deux questions de recherche:

1- Quelles conceptions ont-ils et quelles pratiques d'évaluation mettent en place les enseignants de Mathématique dans leur activité professionnelle?,

2- Les conceptions et les pratiques sont-elles influencées par l'expérience de l'enseignant, par le niveau de scolarité qu'ils enseignent et en rapport avec le stage pédagogique?

Il s'agit d'une étude quantitative à caractère simultanément descriptif et comparatif causal. Les participants ont été 177 enseignants de Mathématique de la région Nord du Portugal, ainsi distribués : 32 pour le 2ème cycle, 93 pour le 3ème cycle et 51 pour le secondaire. Les donnés ont été recueillies par un questionnaire, constitué par 23 items, dont quelques-uns avec plusieurs items ouverts et/ou fermés, sur neuf dimensions de l'évaluation: fonctions, intervenants, formes, préparation, périodicité, conséquences, adaptation et difficultés ressenties et comment améliorer l'évaluation.

En termes d'analyse des donnés, pour la première question de recherche, on a utilisé la statistique descriptive pour les items fermés et l'analyse de contenu pour les items ouverts ; pour la deuxième question de recherche nous n'avons utilisé que des donnés des items fermés avec la statistique descriptive.

Perspectives d'évaluation en mathématiques

Dans cette section, nous présentons les perspectives et les pratiques d'évaluation des enseignants de mathématique impliqués dans cette étude, à partir des nouvelles dimensions d'évaluation contemplées dans le questionnaire : les fonctions, les intervenants, les formes, la préparation, la périodicité et le processus, les conséquences, l'adaptation, les échelles et les difficultés ressentis et comment améliorer l'évaluation.

Fonctions

La majeure partie des enseignants affirme évaluer en mathématique avec les finalités de réguler le processus de l'enseignement (86%) et l'apprentissage des élèves (81%), attribuer des classifications aux élèves (77%) et développer l'élève comme une personne (62%). Toute la finalité de l'évaluation pour orienter les élèves dans leurs futurs choix a été référée par peu d'enseignants (18%). Ces résultats montrent la valorisation de l'intégration de l'évaluation dans le processus de l'enseignement-apprentissage, allant à la rencontre des orientations actuelles sur l'évaluation des élèves.

Intervenants

L'enseignant (99%), les élèves (96%) et le conseil de classe (63%) sont les intervenants dans l'évaluation référés par la majeure partie des enseignants, suivit avec un plus faible pourcentage par le Ministère de l'Éducation ou ses organismes (44%), l'école (40%) et les parents des élèves et les tuteurs (19%). Ces résultats manifestent l'évaluation comme quelque chose qui concerne les enseignants ou les organismes de l'école et très peu d'enseignants reconnaissent aux parents d'élèves et aux tuteurs le rôle qui leur appartient dans l'évaluation. Cependant, quand les enseignants sont questionnés sur les intervenants les plus influents dans l'évaluation, seulement l'enseignant (90%) et les élèves (84%) continuent à être référés par la grande majorité, ce que restreint l'évaluation aux intervenants dans la salle de classe.

En ce qui concerne la participation des élèves dans l'évaluation, il faut distinguer une intervention ponctuelle, surtout à la fin de chaque trimestre scolaire, à laquelle ils appellent l'autoévaluation de l'élève.

Formes

Dans l'attribution des classifications à leurs élèves, presque tous les enseignants réfèrent l'utilisation de testes (99%), de devoirs (94%) et le comportement des élèves dans la classe (89%) ; la majeure partie des enseignants font référence à l' assiduité et à la ponctualité des élèves (79%), aux travaux individuels (71%) et aux travaux de groupe (67%) ; et avec un pourcentage clairement inférieur, ils réfèrent les compositions mathématiques (35%) et les travaux de projets (19%).

Quand les enseignants sont questionnés sur les formes d'évaluation les plus valorisées dans l'attribution des classifications à leurs élèves, seulement les testes continuent à être cités par presque tous les enseignants (96%) et la participation ainsi que l'engagement dans la salle de classe (73%). Nous pouvons, alors, conclure qu'une évaluation individuelle dans l'attribution des classifications aux élèves est la plus valorisée. Les compositions mathématiques assument soit un caractère subsidiaire, s'intégrant dans les testes, dans les travaux individuels et de groupes, soit un caractère autonome, étant considéré comme un instrument qui fonctionne *per se*. Cette forme d'évaluation est pratiquement limitée à l'enseignement secondaire, ce qui est expliqué par son inclusion dans le respectif programme scolaire (Ministère de l'Education, 2002a) et dans l'Examen National de Mathématique de la terminale.

Les travaux de groupe sont implémentés sous le contrôle de l'enseignant, qui définie la composition des groupes, choisit et propose les thèmes et l'évaluation aussi que la présentation orale. Ils sont réalisés à l'intérieur et à l'extérieur de la salle de classe, selon le temps disponible. Aussi sous le contrôle de l'enseignant, les travaux individuels s'intègrent dans les tâches individualisées les plus variées, tels que la résolution d'exercices dans la salle de classe, les travaux de recherche et la correction de testes.

Différemment des formes d'évaluation cités précédemment, les devoirs ont l'impression de valoir *per se*, ayant comme objectif consolider et renforcer les apprentissages et promouvoir les habitudes et les méthodes d'étude. Ces tâches, avec un caractère universel et retirées des livres de texte, s'insèrent plus dans une logique de vérification que dans une logique de régulation des apprentissages, elles se centrent plus dans une différenciation négative (surpasser les lacunes, les déficiences, les limitations et les difficultés) et elles ne constituent pas d'opportunités d'enrichissement, de développement et de différenciation positive.

Finalement, dans le cas de la participation et de l'engagement, du comportement et la assiduité et la ponctualité, les enseignants fréquemment réfèrent leur importance dans la pondération des classifications à attribuer dans le final des trimestres scolaires et réfèrent le recours à des pratiques d'observation en salle de classe à travers du recours à des instruments de registre, nommément les grilles.

Préparation

La plus grande partie des enseignants élabore les différentes formes d'évaluation individuellement (77%) et attribue une plus grande importance au programme de la discipline de mathématique (89%) et à la planification de son Département/groupe Disciplinaire (88%). Ces résultats paraissent détacher le domaine disciplinaire de l'évaluation, centré en chaque enseignant.

L'influence des testes dans les classifications à attribuer aux élèves excède clairement l'importance de toutes les autres formes d'évaluation (66%). La majeure partie des fois, le pourcentage attribué aux testes a été établis par le département/Groupe Disciplinaire (56%) et présentent des avantages au niveau de la validité et de la fiabilité, permettent recueillir plus d'information et individualiser la classification des élèves et constituent un stimulant à l'étude et une meilleur propédeutique pour les examens. Comme désavantages, les enseignants réfèrent les effets psychologiques inhérentes à la réalisation des testes (e.g., stress, anxiété et nervosité) le fait des testes ne permettent pas d'évaluer toutes les capacités/compétences mathématiques considérés importantes (e.g., interaction, oralité et recherche d'information) et incitent l'étude dans les moments qui précèdent les testes.

Relativement aux autres formes d'évaluation, ces avantages se centrent dans le plus grand choix de capacités/compétences qui permettent d'évaluer, stimulent la motivation et le goût pour les mathématiques et, par conséquent, le succès éducatif, promouvant une plus grande confiance dans l'évaluation, car en englobant tout le travail développé et toute la connaissance de l'élève s'assume comme la plus correcte et la plus juste, et sont plus adéquates au promotion d'une évaluation plus régulatrice et formatrice. Du côté des inconvénients, il est référé sa faible mensurabilité, ce qui augmente la subjectivé, rend difficile la quantification et succincte une plus grande discussion entre les enseignants parents d'élèves, de plus grandes exigences de travail et de temps de la part de l'enseignant, une plus grande diversité entre ces pratiques d'évaluation et les Examens Nationaux et plus faibles garanties de légitimité et d'authenticité du travail réalisé par les élèves.

Périodicité et processus

La majeure partie des enseignants fournissent des informations d'évaluation à leurs élèves une ou deux fois par période scolaire, à la fin de chaque unité d'enseignement ou mensuellement (71 %), démontrant se traiter d'une information relative à des testes écrits. A peine un sur quatre enseignants fournissent des informations plus fréquemment. Cette information d'évaluation, pour la majeure partie des enseignants (70%), consiste dans la classification totale, dans les ponctuations parcellaires et en commentaires. Finalement, les restants enseignants (24%) remettent la correction aux

élèves avec des commentaires pour qu'ils corrigent, changeant, de cette façon, son évaluation.

Conséquences

À peine deux sur cinq enseignants (38%) attribuent aussi l'échec scolaire dans la discipline de mathématique au type d'évaluation pratiquée. Dans le cas des enseignants qui n'attribuent pas l'échec scolaire au type d'évaluation pratiquée, qui sont la majeure partie, ils attribuent l'échec scolaire au manque de bases et d'étude et à la vision sociale négative de la discipline ; dans le cas des autres enseignants, on distingue la non correspondance entre ce qui est fait dans les cours et ce qui est requis dans l'évaluation et non la diversification des formes d'évaluation, basée exclusivement dans les testes et centrée dans la classification.

La plus grande partie des enseignants (84%) retire la conséquence de l'évaluation pour ses pratiques pédagogiques, spécialement en cherchant motiver plus les élèves, diminuant le rythme de l'enseignement, individualisant plus le travail avec les élèves, proposant des tâches plus abstracts, débattant les critères d'évaluation et diminuant le degré d'exigences des testes. Dans ces réponses, nous pouvons identifier une présupposition commune : toutes les adaptations surgissent, presque tout le temps, dans la séquence de mauvais résultats et dans un tableau de surpassement de situations d'échec.

Adaptation

Approximativement trois en chaque cinq enseignant (63%) affirment adapter l'évaluation qu'ils effectuent au niveau du dégagement de la classe, ce qui arrive dans les situations d'échec scolaire. Ainsi, ils cherchent adapter les instruments d'évaluation, ils réalisent des testes indifférenciés, changent les critères d'évaluation et diminuent le degré d'exigence. Près de trois sur cinq enseignants (62%) affirment avoir en considération dans l'évaluation les caractéristiques individuelles de ses élèves, stimulant les élèves avec des doutes, élaborant des matériaux et réalisant des testes adaptés. Dans ce cas, il semble être implicite dans les réponses de la majeure partie des enseignants que les caractéristiques individuelles se rapportent vers des aspects plus ou moins pathologiques, faisant quelques uns référence aux élèves avec des nécessités éducatives spéciales.

Échelles

L'échelle de 1 à 5, utilisée dans l'enseignement élémentaire et secondaire, n'est pas adéquate pour plus de deux sur trois enseignants (72%), car elle ne permet pas de différencier les élèves et elle favorise la médiocrité, puisqu'elle ne favorise pas la compétition et ne motive pas suffisamment les élèves. Dans le cas de l'échelle de 0 à 20, utilisé dans l'enseignement secondaire, environ six sur sept des enseignants (88%) la considère adéquate, car elle permet une plus grande différenciation et une plus grande objectivité dans l'évaluation et elle est aussi l'échelle utilisée dans les Examens Nationaux.

Difficultés ressentis et comment améliorer l'évaluation

Les difficultés des enseignants dans l'évaluation se situent au niveau des dispositifs de l'évaluation, des difficultés relatives aux principes de l'ordre éthique et des difficultés en relation avec l'évaluation des aspects comportementaux et attitudinaux. Pour améliorer la manière d'évaluer les élèves, les enseignants suggèrent informer à temps les élèves des critères d'évaluation, implémenter l'évaluation diagnostique et continue, diversifier les formes d'évaluation, approfondir sa formation et le dialogue entre les enseignants dans ce domaine, individualiser davantage l'évaluation, introduire des Examens Nationaux au final de tous les cycles et finir avec les examens. Au delà de ces suggestions spécifiques pour l'évaluation, les enseignants proposent d'autres suggestions plus générales, comme la réduction/changement des programmes scolaires, l'adaptation de l'enseignement des Mathématiques dans la vie quotidienne, une plus grande utilisation des nouvelles technologies et responsabiliser les parents d'élèves.

Synthèse

Pour la majeure partie des enseignants de Mathématique participants dans cette étude: i) prendre des informations pour décider sur l'enseignement, attribuer des classifications et développer l'élève en tant que personne, sont les fonctions de l'évaluation les plus valorisées par les enseignants; ii) l'enseignant et les élèves sont les principaux intervenants dans l'évaluation en mathématique; les parents n'ont presque aucune influence; iii) les instruments les plus valorisés par les enseignants sont les tests à papier et crayon suite à la participation et l'intérêt démontrés par l'élève dans la classe; iv) pour élaborer les instruments d'évaluation, les enseignants comptent surtout sur le programme et la planification de leur département; v) la majorité des enseignants élabore les formes d'évaluation individuellement; un peu plus que la moitié donne aux élèves des informations sur les résultats de l'évaluation une ou deux fois par trimestre ou à la fin de chaque unité d'enseignement; vi) l'échec dans la discipline de mathématique est attribué au manque de connaissances des élèves plutôt qu'à l'évaluation qui est faite par les enseignants; vii) les résultats de l'évaluation des élèves influencent les pratiques pédagogiques ultérieures des enseignants, puisque la majorité adapte l'évaluation au niveau de la classe et prend en considération les caractéristiques individuelles, mais seulement s'il y a des situations d'échec; viii) concernant les difficultés ressenties et la façon d'améliorer l'évaluation, les enseignants identifient des difficultés liées aux dispositifs d'évaluation, surtout la façon de définir les meilleurs instruments d'évaluation, avec l'amplitude adaptée à la diversité des compétences et au nombre d'élèves qu'ils sont obligés à évaluer; difficultés liées à des principes d'ordre éthique (comment être juste? comment dépasser l'affectivité?); difficultés liées à l'évaluation du comportement et des attitudes (comment évaluer les élèves qu'ont un mauvais comportement, mais des bons résultats?); ix) les suggestions présentées par les participants pour améliorer la façon d'évaluer les élèves sont les suivantes: informer les élèves sur les critères d'évaluation dès la première session du cours; faire l'évaluation diagnostique; diversifier les instruments d'évaluation; dialogue entre les enseignants surtout pendant les actions de formation; réduction du nombre d'élèves par classe; examens nationaux à la fin de chaque cycle; diminution des contenus ou leur altération, adaptation de l'enseignement de chaque discipline à des problèmes de la vie pratique, adapter les salles, surtout avec des ordinateurs et, finalement, responsabiliser les parents.

Conclusion

Ces résultats nous amènent à affirmer que les pratiques des enseignants sont plutôt de l'ordre de la mesure, du control, de l'efficacité et de la performativité (Alves & Machado, 2003 ; Bonniol & Vial, 1997) que de l'ordre de la régulation, de la transformation, de la participation et de l'émancipation (Alves & Machado, 2003 ; Bonniol & Vial, 1997).

Le tableau 1 ci-dessous fait la synthèse des résultats, selon les différentes dimensions étudiées dans l'étude et les paradigmes présentés.

Tableau 1 - Inscription des pratiques dans les paradigmes de l'évaluation.

Paradigme de la performativité (mesure, control, efficacité)	**Paradigme de l'émancipation** (régulation, transformation, participation)
Fonctions de l'évaluation	
Attribuer des classifications.	Décider sur l'enseignement et développer l'élève en tant que personne.
Principaux intervenants dans l'évaluation	
L'enseignant et les élèves.	Les parents n'ont presque aucune influence.
Instruments les plus valorisés	
Les tests à papier et crayon.	La «participation» et l'«intérêt» démontrés par l'élève dans la classe.
Elaboration des instruments d'évaluation	
Les enseignants comptent surtout sur le programme et la planification de leur département. La majorité des enseignants élabore les formes d'évaluation individuellement.	
Échelles	
L'échelle de 1 à 5 (l'enseignement élémentaire) n'est pas adéquat pour plus de deux sur trois enseignants (72%), ne permet pas de différencier les élèves (elle favorise la «médiocrité», ne promeuve pas la compétition et ne motive pas suffisamment les élèves). Dans l'échelle de 0 à 20 (enseignement secondaire) environ six sur sept des enseignants (88%) la considère adéquate, permet une plus grande différenciation et une plus grande objectivité dans l'évaluation, et est aussi l'échelle utilisée dans les Examens Nationaux.	
Communication des résultats	

Un peu plus que la moitié donne aux élèves des informations sur les résultats de l'évaluation une ou deux fois par période ou à la fin de chaque unité d'enseignement.	
L'échec dans la discipline de mathématique	
Manque de connaissances des élèves plutôt qu'à l'évaluation qui est faite par les enseignants.	
L'influence des résultats de l'évaluation dans les pratiques pédagogiques	
	La majorité adapte l'évaluation au niveau de la classe et prend en considération les caractéristiques individuelles, mais seulement s'il y a des situations d'échec.
Difficultés senties et la façon d'améliorer l'évaluation	
Définir les meilleurs instruments d'évaluation, avec l'amplitude adaptée à la diversité des compétences et au numéro d'élèves.	Liées à des principes d'ordre éthique (comment être juste? comment dépasser l'affectivité?) et au comportement et aux attitudes (comment évaluer les élèves qu'ont un mauvais comportement, mais des bons résultats?).
Suggestions présentées	
Examens nationaux à la fin de chaque cycle; diminution des contenus ou leur altération, adaptation de l'enseignement de chaque discipline à des problèmes de la vie pratique, adapter les salles, surtout avec des ordinateurs.	Informer les élèves sur les critères d'évaluation dès la première session du cours, faire l'évaluation diagnostique, diversifier les instruments d'évaluation, dialogue entre les enseignants surtout pendant les actions de formation, réduction du nombre d'élèves par classe et responsabiliser les parents.

Si nous essayons une conceptualisation de l'activité évaluative des enseignants de Mathématique, il ressort évident:

- Un changement hybride: il y a besoin de répondre, à la fois, au paradoxe de l'exigence de la performativité et les ambitions de l'émancipation et ça se fait plutôt par un engagement entre les nouvelles exigences et les contraintes des pratiques, que par une innovation effective et holistique,
- Un processus d'innovation «hypocrite» (Brusson, 2006): pour survivre à une pression dilacérant les enseignants adaptent pragmatiquement les propositions d'innovation aux pratiques qui leurs assurent «tranquillité» sans l'obligation de prendre des risques individuellement,
- La résistance aux changements dans l'évaluation ne se situe pas au niveau de la pensée de l'enseignant : ses représentations, ses conceptions, etc., (Garcia, 1999); en revanche, il y a une adhésion conceptuel au changement sans une totale correspondance dans les pratiques, tel que Louis (1999), l'avait déjà observé.

Références

Abrantes, P. (2001). *Reorganização curricular do ensino básico - Princípios, medidas e implicações.* Lisboa : ME/DEB.

Allal, L. (2000). Et l'architecture des savoirs? *Revue suisse des sciences de l'éducation, 22*(1), 165-172.

Alves, M. P. & Machado, E. A. (2003). Sentido da Escola e os sentidos da avaliação. *Revista de Estudos Curriculares, 1*(1), 79-92.

Alves, M. P. (2002). A avaliação e o desenvolvimento profissional do professor. In A. F. Moreira & Elisabete M. (Orgs.), *Currículo, práticas pedagógicas e identidades* (pp. 138-159). Porto : Porto Editora.

APM (1998). Matemática 2001. *Diagnóstico e recomendações para o ensino e aprendizagem da matemática. Relatório preliminar*. Lisboa : Autor.

Barbier, J. –M. (1990). *A Avaliação em Formação*. Porto : Edições Afrontamento.

Belair, L. (1999). *L'évaluation dans l'école.* Paris : ESF.

Bonniol, J.-J. & Vial, M. (1997). *Les modèles de l'évaluation : textes fondateurs avec commentaires.* Bruxelles : De Boeck Université.

Brusson, N. (2006). *A Organização da hipocrisia*. Porto : Edições ASA.

Cardinet, J. (1994). *Avaliar é medir?.* Rio Tinto : Edições ASA.

Depover, C. & Noel, B. (1999). *L'évaluation des compétences et des processus cognitifs.* Bruxelles : De Boeck Université.

García, C. M. (1999). *Formação de professores. Para uma mudança educativa*. Porto : Porto Editora.

Hadji, C. (1994). *A avaliaçäo, regras do jogo : Das intenções aos instrumentos*. Porto : Porto Editora.

Louis, R. (1999). *L'évaluation des apprentissages en classe : théorie et pratique*. Laval : Éditions Études Vivantes.

Ponte, J. P., Matos, J. M. & Abrantes, P. (1998). *Investigação em educação matemática : implicações curriculares*. Lisboa : Instituto de Inovação Educacional.

Ramalho, G. (2002). *PISA 2000 : Segundo relatório nacional*. Lisboa : ME/GAVE.

PRECODEM, UN EXEMPLE DE TAO MIS AU SERVICE DE L'EMPLOI

JADOUL Raynald, Centre de Recherche Public Henri Tudor
MIZOHATA Sachié, Université de Paris V, Sorbonne

Mots-clés : employabilité, compétence, TAO.

Résumé :
Dans ce secteur d'activité florissant qu'est l'IT, où la demande en ressources humaines est grande, il y a pourtant du chômage. Pourquoi une telle situation paradoxale ? Comment en sortir ? En première approche, une explication simple semble que la maîtrise des savoirs devrait évoluer à la cadence effrénée des avancées technologiques mais que pour la majorité des travailleurs ce soit au-delà de l'humainement réalisable. Ainsi, les compétences « passent » et perdent de leur valeur sur le marché de l'emploi. Alors, que reste-t-il à faire ? Se former. Pour ce faire, les moyens sont mis en oeuvre et les initiatives se multiplient. PRECODEM (« PRÉparation aux COmpétences de DEMain ») est un projet en partenariat entre le Centre de Recherche Public Henri Tudor et l'Agence pour l'Emploi du Luxembourg (ADEM) dont l'objectif est de concevoir, d'expérimenter et d'évaluer de nouvelles approches en matière de formation et de lutte contre le chômage. Cette démarche prévoit une phase de veille métier permettant l'élaboration de référentiels de compétences qui seront évaluées chez 30 candidats demandeurs d'emploi, une première fois avant qu'ils ne reçoivent un ensemble de formations et, une deuxième fois après lesdites formations. La comparaison des résultats de ces tests combinée à une série d'informations liées aux profils des candidats devrait notamment permettre d'évaluer la progression du niveau de compétence desdits candidats et de déterminer la qualité de l'ensemble de la démarche, des formations et leur degré d'adéquation vis-à-vis des objectifs fixés. Les résultats consolidés du projet seront disponibles à la fin du mois d'août 2006 et formeront la base de recommandations destinées aux instances nationales compétentes.

Introduction

Le domaine des technologies de l'information (« information technology » en anglais ou IT) est vaste, riche en potentialités et, malgré de légères stagnations sporadiques, en expansion depuis plusieurs dizaines d'années. La demande en ressources humaines y est restée globalement forte (Gille & Marti, 2000 ; Ministère de l'Education nationale française, 2006).

Pourtant, de nombreux informaticiens sont sans emploi. Or, comme l'avance Amartya Sen, Prix Nobel d'Economie, les preuves que la perte d'emploi a des implications qui vont bien au-delà du déboire financier sont nombreuses ; citons, par exemple, la perte de compétences (Sen, 2000). De plus, être sans emploi alors qu'il y

a de la demande doit être difficile à vivre, mais ne pas savoir comment sortir de cette impasse l'est certainement plus.

Partant de ce constat, on peut s'interroger, d'abord, sur les raisons qui font qu'un demandeur d'emploi, motivé à retravailler, ne puisse « coller » à la demande du marché et, ensuite, sur les remèdes à ce « problème ».

Dans ce document, nous aborderons d'abord, sans prétention d'exhaustivité, la problématique de la perte d'emploi dans le secteur de l'IT, ensuite, nous envisagerons différentes pistes de solution et, enfin, dans cette perspective, nous présenterons le projet PRECODEM - initiative de l'Agence pour l'Emploi du Luxembourg (ADEM) cofinancée par l'Union Européenne. Nous conclurons sur quelques recommandations de déploiement d'un outillage de type TAO* comme soutien d'une politique réactive et pérenne de maintien de l'employabilité.

Pourquoi une crise de l'employabilité dans le secteur de l'IT ?

Les motifs pouvant empêcher une personne de trouver un emploi sont très nombreux : sociaux, culturels, structurels, psychologiques, etc. (Sen, 2000). Néanmoins, pour le secteur IT, c'est l'évolution effrénée des connaissances, du savoir-faire et des outils qui revient parmi les facteurs les plus souvent cités dans les enquêtes.

La puissance des machines double quasi tous les deux ans (Tuomi, 2002), ouvrant ainsi la porte aux besoins toujours plus complexes tandis que le professionnel de l'IT s'échine pour se mettre à la page des nouvelles architectures matérielles et logicielles, des derniers langages et paradigmes de développement, des redécoupages du travail, etc.

Cette évolution galopante a donc une influence directe sur l'« employabilité » de l'individu. Pour rappel, l'employabilité peut se définir comme l'ensemble des préalables à posséder à un niveau minimal afin de trouver et de conserver un emploi (Bolton, 1981). Cette définition de Bolton ajustée par les études récentes (INTEREF, 2001), donne une vue de l'employabilité déclinée en huit pôles (voir « Figure 1 - Les huit pôles de l'employabilité »).

* TAO : Testing Assisté par Ordinateur

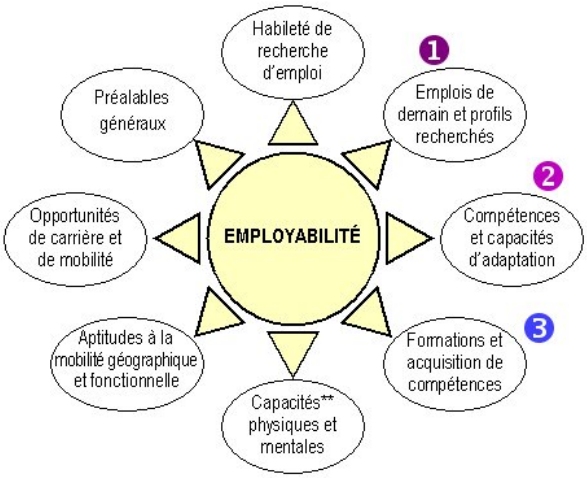


Figure 1 – Les huit pôles** de l'employabilité**.

Nous n'abordons dans les paragraphes qui suivent que les trois premiers pôles du diagramme de l'employabilité car ceux-ci font directement sens dans le cadre du projet PRECODEM.

D'abord, l'axe ❶ « Emploi de demain et profils recherchés » suggère une veille économique et technologique qui identifierait les secteurs porteurs et les formations cadrant avec les profils recherchés. Hélas, cette information est souvent peu visible et ne sert donc pas le chercheur d'emploi.

Le point ❷ cerne les « Compétences et capacités d'adaptation » liées aux nouveaux savoirs, savoir-faire, modes organisationnels et outillages. De précédentes études montrent que la « demi-vie professionnelle » (ou délai nécessaire pour que les compétences d'un travailleur se dévalorisent de moitié sur le marché suite à l'évolution des savoirs et technologies) était, dans les années 60, située entre 7 et 14 ans (Kearsley, 1989) et qu'à présent, elle est ramenée entre 3 et 5 ans (Casey, 1998 ; Guillemard, 2003).

Parallèlement, la « dénormalisation » des trajectoires de vie (carrière fragmentée, etc.) amène les travailleurs à la « barrière des 40 ans » que nombre d'entreprises considèrent comme le seuil au-delà duquel les employés sont jugés trop vieux pour rester performants, et que leur formation ne peut plus offrir de retour sur investissement suffisant (Guillemard, 2003 ; Zanardelli & Leduc, 2006). Or, l'Europe considère (Commission européenne, 2002), elle, que l'employabilité repose bien sur cet axe ❸ et que les employés devraient accéder à la formation continue sans discrimination (Guillemard, 2003 ; Wannell & Ali, 2002).

** Le terme « Capacité » se réfère à l'Approche des Capacités développée par A. Sen, M. Nussbaum et d'autres, qui étudie les fonctionnements de l'individu (états et actions) et ses capacités (liberté et opportunités de réaliser ces fonctionnements) (Sen, 2000).

Paradoxalement, ces discriminations dans l'octroi de formations au plus de 40 ans, et les tendances à la dévalorisation des compétences sont à relativiser pour les profils très diplômés (Wannell & Ali, 2002) ; cela pourrait partiellement s'expliquer, d'abord, par la volonté des sociétés de protéger l'investissement lourd lié à l'embauche d'un profil fortement diplômé mais aussi par la nature même du cursus universitaire largement tourné vers la meta-connaissance et forgeant chez l'individu une culture d'auto apprentissage, où d'aptitude à développer de nouvelles compétences professionnelles en cours de carrière, rendant donc pérenne son utilité au sein de l'entreprise.

Envisageons-nous de sortir de la crise de l'employabilité ?

Il semble que ce soit le cas. La volonté ne manque pas et les moyens, politiques, financiers, structurels ou encore technologiques, sont mis en oeuvre. La formation tout au long de la vie devient une réalité et derrière les leaders du domaine (Suède, U.K. et Danemark), les initiatives fourmillent et les résultats arrivent doucement (Commission européenne, 2006).

Ainsi, par exemple, en Europe, Internet a un taux de pénétration très élevé dans les entreprises et dans les ménages (Commission européenne, 2001 ; Statec, 2006) où il progresse en outre comme outil pédagogique et de formation semi-structurée (Statec, 2006 ; Dryburgh, 2002).

Le testing devient également important car ces initiatives font plus que repenser les contenus des formations en visant aussi à les améliorer en termes de pédagogie, de coordination et d'adéquation.

PRECODEM, une initiative prometteuse

PRECODEM (« PRÉparation aux COmpétences de DEMain ») est un projet cofinancé par l'Union Européenne et mené par le Centre Henri Tudor en partenariat avec l'Agence pour l'Emploi du Luxembourg (ADEM).

L'objectif du projet est de concevoir, d'expérimenter et d'évaluer de nouvelles approches en matière de formation et de lutte contre le chômage (Szelest & Reiter, 2006). Dans cette optique, trois phases d'activité ont été définies :

- Elaboration de référentiels de compétences relatifs à deux métiers (niveau cadre et ingénieur) choisis par l'ADEM (sur base d'une veille « métier » effectuée par le Centre de Veille Technologique du Centre de Recherche Public Henri Tudor), intégrant les pratiques et technologies avancées du métier ;
- Conception et réalisation par des experts des domaines concernés, de 4 modules de formations innovantes relatives à ces pratiques et technologies avancées ;
- Evaluation – au moyen, notamment, d'un dispositif de type Testing Assisté par Ordinateur – de l'apport du dispositif (référentiels/formations/testing), d'une part, dans le capital compétences du demandeur d'emploi et, d'autre part, pour l'ADEM et les structures intervenant dans la prise en charge des demandeurs d'emploi.

Afin de travailler au plus près des besoins et des attentes des candidats demandeurs d'emploi, ceux-ci ont participé à des interviews et une analyse de leur curriculum vitæ a été effectuée (cf. « Figure 2 – Graphe des âges » et « Tableau 1 – Quelques statistiques relatives aux candidats PRECODEM », à la page suivante).

21-25	26-30	31-35	36-40	41-45	46-50	51-56	ans
3,3	10,0	13,3	40,0	13,3	16,7	3,3	%

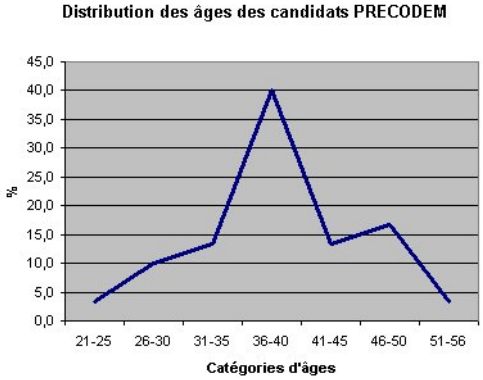


Figure 2 - Graphe des âges.

Tableau 1 – Quelques statistiques relatives aux candidats PRECODEM.

Candidat(e)	Résultat aux tests (en %)		Formation choisie		Homme ou Femme		Inactivité en mois	Nombre d'emplois	Age
1	31,37		AP		H		8	5	37
2		3,33		SI		F	15	6	40
3		47,43		SI	H		env. 24	5	37
4		30,77		SI	H		14	5	39
5		32,20		SI	H		9	4	34
6	34,57		AP		H		12	6	40
7		55,67		SI	H		5	10	44
8		25,47		SI	H		env. 12	8	37
9	17,67		AP		H		8	6	39
10		34,90		SI		F	10	14	36
11		28,23		SI	H		52	7	50
12	24,57		AP		H		12	8	45
13		54,63		SI	H		6	5	42
14		22,67		SI	H		2	1	21
15	14,27		AP		H		9	1	27
16		35,47		SI	H		14	6	33
17	32,47		AP			F	18	6	43
18		31,97		SI	H		36	11	40
19	10,77		AP			F	55	3	46
20	20,87		AP		H		3	2	46
21	12,23		AP		H		7	2	34
22		38,67		SI	H		env. 6	11	46
23		0,00		SI	H		26	9	39
24	31,83		AP		H		env. 24	2	39
25	34,57		AP		H		22	2	26
26		37,50		SI	H		env. 12	11	48
27	10,73		AP		H		72	2	37
28		30,40		SI	H		env. 12	0	26
29	51,50		AP		H		3	4	31
30		39,80		SI	H		12	5	56
Moyenne	25,18	32,30					18	5,6	39
Pourcentage	29,2		43,3	56,7	86,7	13,3			

Par ailleurs, les formateurs ont préalablement été formés à l'outil de TAO afin de l'exploiter au mieux et, ont pu, dans le cadre du projet FormIT, bénéficier de cours de perfectionnement à la formation (gestion de groupes, création de supports de cours, etc.) (vandenabeele & Giltay, 2005).

L'évaluation par TAO des candidats se passe en deux temps ; une fois avant les formations et, une deuxième fois après celles-ci. La comparaison des résultats de ces tests combinée aux informations des profils permettra notamment d'évaluer la progression du niveau de compétence des candidats et de déterminer la qualité des formations et leur degré d'adéquation vis-à-vis des objectifs fixés.

Les résultats définitifs du projet seront connus fin septembre 2006 et à cette fin, un rapport de synthèse et de recommandation sera destiné aux instances nationales compétentes pour, le cas échéant, généraliser le dispositif au niveau national.

Conclusion

Outre le besoin de persévérance au niveau politique (Commission européenne, 2006), la mise en place d'une gamme d'outils adéquats est indispensable au soutien de l'effort favorisant le maintien de l'employabilité. De tels outils permettront d'abord de définir les profils recherchés sur le marché de l'emploi et les compétences attendues pour lesdits profils grâce à de la veille économique, métier et technologique.

Dans une volonté d'instrumenter une telle politique active et réactive, la formation continue et structurée, de toutes formes, interne ou externe à l'entreprise, de promotion sociale ou autre, doit s'équiper d'outils idoines permettant d'évaluer la qualité des formations, le gain en compétences des apprenants et le niveau d'adéquation entre formations et apprenants (Latour et al., 2005).

Des outils doivent également permettre de mieux cerner les attentes des personnes désirant trouver un emploi ou en changer afin d'évaluer le déficit de compétences qu'elles doivent combler, de les aider et de les orienter, quels que soient leur formation initiale, leur sexe et leur âge car on peut toujours apprendre et garder de la valeur sur le marché du travail tout au long de sa vie.

Références bibliographiques

Allen Consulting Group. (2004). Development of a Strategy to Support the Universal Recognition and Recording of Employability Skills. A Skills Portfolio Approach. Employability skills. *Final report*. Camberra : Australian Government Department of Education, Science and Training.

Bolton, B. (1981). Assessing Employability of Handicapped Person : The Vocational Rehabilitation Perspective. *Journal of Applied Rehabilitation Counseling*, 12 (1).

Casey, B. (1998). Retraite anticipée ou tardive : incitations et désincitations. Document de travail, OCDE, AWP 3.3F, projet « Préserver la prospérité dans une société vieillissante ». Paris : OCDE.

Commission européenne. (2001). Direction générale des Affaires Economiques et Financières. Economie Européenne : l'économie de l'Union, bilan 2000. IP/06/618. Bruxelles : Communautés européennes.

Commission européenne. (2002). Evaluation de la mise en oeuvre des lignes directrices pour l'emploi 2002. Document d'appui pour le rapport conjoint sur l'emploi 2002. COM (2002) 621 final. Bruxelles :
Disponible à l'adresse suivante
http://ec.europa.eu/employment_social/employment_strategy/report_2002/supp/supp_fr.pdf

Commission européenne. (2006). Education et formation en Europe : de nouveaux efforts sont nécessaires pour atteindre les objectifs de Lisbonne. IP/06/618. Bruxelles : Communautés européennes.
Disponible à l'adresse suivante
http://ec.europa.eu/education/policies/2010/objectives_en.html

Dryburgh, H. (2002). L'acquisition de compétences en informatique. *Tendances sociales canadiennes*. Ottawa : Statistique Canada, Ministre de l'Industrie.

Gille, L. & Marti, R. (2000). Les technologies de l'information et des communications et l'emploi en France. Ministère de l'Economie, des Finances et de l'Industrie Direction Générale de l'Industrie, des Technologies de l'Information et des Postes. BIPE.
Disponible à l'adresse suivante
http://www.telecom.gouv.fr/fonds_documentaire/rapports/macro-eco.pdf

Guillemard, A.-M. (2003). *L'âge de l'emploi. Les sociétés à l'épreuve du vieillissement.* Paris : Armand Colin.

INTEREF. (2001). Résultats du Programme d'étude « Gestion des Ressources Humaines Demain », Glossaire. Espace d'échange des observatoires emploi-formation. Villeneuve d'Ascq.
Disponible à l'adresse suivante
http://www.interef.com/ateliers/grh_demain/fiches/employabilite.htm

Kearsley, G. (1989). Introducing new technology into the workplace: retraining issues and strategies. Background Paper No. 8b, Department of Labor, Washington, DC., Commission on Workforce Quality and Labor Market Efficiency. Washington, DC : the US Department of Labor, 457- 492.
Disponible à l'adresse suivante

http://eric.ed.gov/ERICWebPortal/Home.portal?_nfpb=true&_pageLabel=RecordDetails&ERICExtSearch_SearchValue_0=ED317675&ERICExtSearch_SearchType_0=eric_accno&objectId=0900000b80049143

Latour, T. et al. (2005). Covering Different Levels of Evaluation Needs by an Internet-Based Computer-Assisted Testing Framework for Collaborative Distributed Test Development and Delivery. Communication à EDMEDIA 2005, Canada.
Disponible à l'adresse suivante
http://www.tao.lu/html/index.php?option=com_content&task=view&id=24&Itemid=46

Ministère de l'Éducation nationale française. (2006). Prospective emploi-formation à l'horizon 2015. *Note d'information* n°06.03 février. Paris : Ministère de l'Éducation nationale.
Disponible à l'adresse suivante
http://www.education.gouv.fr/stateval

Sen, A. (2000). *Development as Freedom*. New York : Anchor Books.

Statec. (2006). L'équipement, la formation et les compétences informatiques des Luxembourgeois. *Statnews*. n°28, Luxembourg : Statec.

Szelest, L. & Reiter, S. (2006). PRECODEM - Préparation aux compétences de demain pour les demandeurs d'emploi - cadres et ingénieurs. *Fiche descriptive*. Luxembourg : Centre Henri Tudor.

Tuomi, I. (2002). The Lives and Death of Moore's Law. *First Monday*, volume 7, number 11.
Disponible à l'adresse suivante
http://www.firstmonday.org/issues/issue7_11/tuomi/

Vandenabeele, L. & Giltay, D. (2005). FormIT. *Fiche descriptive*. Luxembourg : Centre Henri Tudor.

Wannell, T. & Ali, J. (2002). Travailler intelligemment : le changement technologique influencé par les compétences. *Tendances sociales canadiennes*. Ottawa : Statistique Canada, Ministre de l'Industrie.

Zanardelli, M. & Leduc, K. (2006). Favoriser le vieillissement actif : les pratiques des entreprises en matière de formation continue pour les travailleurs âgés. *Population & Emploi, N° 15/2006*. Luxembourg : CEPS / Instead.

EFFETS DE L'INTRODUCTION D'UN DISPOSITIF INFORMATISE D'EEE DANS UNE UNIVERSITE FRANCAISE.

Nathalie Younes, UPMF et Université d'Auvergne, France

Résumé :
La communication présente une recherche de thèse achevée interrogeant les effets de l'introduction d'un dispositif informatisé d'évaluation de l'enseignement par les étudiants à visée formative dans un contexte universitaire français.

L'effet de seuil de l'EEE informatisée

La problématique interroge en quoi un tel dispositif qui représente un nœud de changement au sein de l'université (Bru, 2005), constitue un seuil critique de passage d'un système basé sur un enseignement individuel (Fave-Bonnet, 2004) – dans lequel les méthodes d'enseignement ainsi que les modalités d'évaluation des étudiants ne sont pas nécessairement explicitées – à un autre système supposant de formuler les objectifs et les méthodes pédagogiques, de rendre compte de ce qui se passe dans l'enseignement et de s'appuyer sur une démarche collective. Nous avons déterminé la notion d'effet de seuil non comme une limite clairement définie mais comme une suite de passages pouvant conduire à modifier profondément les caractéristiques du système d'enseignement. Nous considérons en quoi la participation des acteurs à la mise en place et au suivi de l'évaluation (Rossi and Freeman, 1985 ; Stake, 1986), les contradictions au sein de l'université quant à l'EEE entre contrôle et évaluation (Berger, 1990) et l'impact de l'informatisation du dispositif d'EEE (Teall and Franklin, 2001) sont des facteurs d'effet.

Méthodes et outils

L'étude empirique, située à l'Université d'Auvergne (France) s'est appuyée sur deux opérations successives : d'abord la conception et la mise en place d'un dispositif informatisé d'EEE à partir d'un ancrage théorique dans l'évaluation comme processus de régulation intersubjectif inhérent au développement et à l'évolution de l'action pédagogique (Stufflebeam et al. 1980). Au niveau méthodologique, nous avons retenu la référentialisation en tant que processus assorti de procédures critériées (Figari, 1994, 2001). La deuxième opération a été le repérage des effets liés à l'introduction de ce dispositif. Ces effets ont été étudiés par des méthodes qualitatives et quantitatives, à partir du positionnement de l'institution universitaire et de celui des acteurs vis à vis de l'EEE, à partir de la participation à l'évaluation, à partir des informations issues des EEE et à partir des effets perçus et escomptés par les acteurs.

Résultats : ambiguïté et paradoxes à tous les niveaux d'acteur

Les résultats de l'investigation des effets de l'introduction du dispositif d'EEE ont mis en évidence, de manière transversale, l'ambiguïté et les paradoxes des effets de

l'EEE qui constitue un seuil critique accompagnant la remise en cause du système d'enseignement traditionnel. Nous avons conclu que l'implantation d'un dispositif d'EEE à visée formative n'a pas suffi à établir le passage d'un système d'enseignement S1 (basé sur un enseignement individuel dans lequel les méthodes d'enseignement ainsi que les modalités d'évaluation des étudiants ne sont pas nécessairement explicitées) à un autre système d'enseignement S2 supposant au contraire d'expliciter les visées et méthodes pédagogiques, de rendre compte de ce qui se passe dans l'enseignement et de favoriser une démarche collective. L'analyse des résultats de notre étude nous a en particulier amenée à constater que l'inertie et la résistance à l'EEE ont bloqué un tel passage. Les différents types d'effets tels qu'ils ont été étudiés sont paradoxaux, reflétant l'ambiguïté de l'institution vis à vis de l'EEE qui oscille entre trois positions antagonistes : la considérer comme une formalité administrative, comme une ressource formative pour améliorer l'enseignement ou comme un outil de contrôle. Ce qui pourrait expliquer qu'un dispositif d'EEE même à visée formative ait suscité plus d'inertie et d'opposition que d'adhésion chez les enseignants qui redoutent son utilisation à des fins de contrôle. Par ailleurs, nous avons considéré que l'informatisation du dispositif et sa mise sur Internet ont accentué les effets de l'EEE.

Perspectives

Au terme cette recherche, deux perspectives peuvent être ouvertes :

1/ Travailler sur les effets de l'EEE, c'est prendre la mesure de l'EEE en tant que seuil critique de transformation des pratiques pédagogiques et du système d'enseignement. Alors que l'enseignant fonctionne sur un modèle pédagogique plutôt individuel et largement implicite, avec une culture de l'évaluation intuitive et personnelle, l'évaluation visant des effets dans la formation renvoie au contraire à un processus concerté et collectif qui nécessite la construction d'un autre modèle qui serait explicite et partagé. L'inertie et la résistance que suscite l'EEE mettent en évidence l'importance du seuil à franchir et l'ambiguïté des réactions que ce passage provoque. L'institution n'a peut être pas pris toute la mesure de l'effet global de cette évaluation que ce soit dans la prise en compte des effets paradoxaux qu'elle provoque ou que ce soit dans la mise en place des moyens pédagogiques qu'elle requiert.

2/ L'EEE met en évidence que l'enseignement supérieur « reste une source de questionnement » (Bru, 2004) que ce soit au niveau des modèles d'enseignement, des pratiques pédagogiques ou des carrières universitaires. Si le modèle économique professionnel gagne du terrain, le primat du modèle académique valorisant l'évaluation par les pairs et le savoir scientifique semble continuer à primer sur le modèle pédagogique. L'évaluation renvoie à la question délicate des bonnes pratiques pédagogiques, alors même que ce cadre de références n'est ni identifié ni débattu par l'ensemble des enseignants, voire considéré suivant des visées contradictoires. L'évaluation formative vise à susciter une dynamique pédagogique d'explicitation, de négociation et de partage. Ce vecteur pourrait être renforcé en s'inscrivant dans une réflexion globale incluant les finalités des formations ainsi que les modalités de l'enseignement et que les autres formes d'évaluation que l'EEE.

L'importance donnée à la pédagogie dans l'institution est un autre facteur déterminant. Alors qu'il y a une injonction institutionnelle d'évaluation de l'enseignement, ce dernier reste cependant peu valorisé par l'institution elle-même

puisqu'il a peu d'impact sur la carrière, qu'il n'y a pas de formation à la pédagogie non plus qu'un système d'accompagnement pédagogique. L'EEE est un vecteur stratégique pour amener les enseignants à expliciter et à débattre de leurs enseignements. Les nouvelles dynamiques universitaires en Europe rendent particulièrement nécessaire de considérer que l'enseignement ne peut pas reposer sur les seules intuitions ou intentions implicites que l'enseignant ne saurait élucider. Il exige au contraire qu'un savoir pédagogique universitaire se constitue rationnellement et collectivement à partir des données disciplinaires mais aussi à partir des données d'expérience qui peuvent être partagées. Mais pour que les enseignants s'y engagent, certaines confusions doivent être levées. Il importe notamment de clarifier l'EEE comme participant de l'amélioration de l'enseignement et de la distinguer de son utilisation comme contrôle ou comme gestion du personnel. Il s'agit également de ne pas inscrire ce type d'évaluation dans le cadre d'un quelconque référentiel conçu comme une « norme » de comparaison prescrite de l'extérieur (Figari, 2001), qui serait perçue comme particulièrement menaçante et inappropriée.

PRINCIPES ET OUTILS DE L'EVALUATION ACADEMIQUE DES COMPETENCES EN FORMATION INITIALE D'INFIRMIER-IERES

Brigitte Kampel, Salomé Anderfuhren Fiaux, Haute Ecole Cantonale Vaudoise de la Santé, HECVSanté, Lausanne

Mots-clés : Outils de formation et d'évaluation, pédagogie socioconstructiviste, cohérence curriculaire et évaluation

Résumé :

A l'occasion de changements majeurs dans le cadre de la formation initiale d'infirmier-ères en Suisse Romande, notre institution (HecvS) a complètement repensé sa philosophie pédagogique et son curriculum. Le dispositif d'évaluation a donc également subi un remaniement total, tant au niveau des principes que des outils utilisés. Nous développerons ces deux points dans notre présentation. Parmi les principes qui guident notre pratique actuelle, citons la centration constante sur les compétences professionnelles (processus et résultat), l'utilisation majoritaire de l'évaluation formative, l'évaluation sommative se déroulant à la fin de l'année. Un élément fort du dispositif, qui participe grandement à son efficacité, est la cohérence entre dispositif didactique et évaluatif et entre les deux volets de l'évaluation. En effet, les mêmes outils sont utilisés pour l'enseignement, la régulation et la certification. De plus, leur nombre a été limité à cinq, choisis pour leur potentiel par rapport à l'apprentissage, leur intégration à une approche réflexive et leur pertinence par rapport à la professionnalisation des étudiants en soins infirmiers. Ce sont le projet de soins, la carte conceptuelle, la base d'orientation, le récit et analyse d'expérience et l'analyse de traces issues du portfolio. Grâce à un accompagnement méthodologique et aux liens constants réalisés avec la pratique infirmière, les étudiants se sont appropriés ces outils et perçoivent le sens de leur utilisation pour la pratique professionnelle et pour la formation. Du côté des professeurs, ce changement a également été une rupture par rapport aux pratiques antérieures ; après quatre ans de

travail, la maîtrise des outils est en très bonne voie ; l'harmonisation de l'évaluation formative et la restitution d'évaluation descriptives précises, touchant aussi bien les processus que les résultats, restent à améliorer. Nous sommes particulièrement satisfaits d'avoir pu, grâce au nouveau dispositif, évaluer, en situation scolaire, le niveau de compétence de l'étudiant et non seulement ses connaissances disciplinaires.

Principes et outils de l'évaluation académique des compétences en formation initiale d'infirmier-ières

En Suisse Romande, la formation professionnelle d'infirmier/ère s'effectue depuis 2002 dans une Haute Ecole Spécialisée, et se situe au niveau tertiaire. Elle prépare l'étudiant à agir et réagir dans des situations humaines complexes, d'évolution difficilement prévisible ; et vise le développement de 9 compétences du référentiel métier, de niveau réflexif.

A cette occasion, la filière infirmier/ère de la HECVSanté (Hante Ecole Cantonale Vaudoise de la Santé) a complètement repensé sa philosophie pédagogique, organisant le programme de telle manière que l'apprentissage, le développement des compétences des étudiants soient le centre du dispositif[1]. Parallèlement à une réorganisation complète du curriculum, structuré autour de l'apprentissage par problèmes, le dispositif d'évaluation a également été complètement transformé.

Etant donné que la compétence est un savoir agir efficace en situation, est-il possible d'évaluer le développement des compétences professionnelles à l'école ou est-ce réservé à la formation pratique ? Nous avons clairement affirmé cette possibilité, mais comment aller au-delà de l'évaluation des connaissances et rendre visible, en situation scolaire, le processus de combinaison d'un ensemble de ressources pour un agir professionnel réflexif ?

Nous vous présentons ici comment nous avons résolu ces questions ainsi que les principes et outils qui sont à la base du dispositif d'évaluation.

---

[1] NB. Le dispositif de formation fait l'objet d'une autre présentation par les des mêmes auteures

Principes à la base de l'évaluation

- Le référentiel de compétences professionnelles est base de toute évaluation. Chaque compétence est déclinée en indicateurs de développement annuels,
- Le développement des compétences n'est pas linéaire, nécessite du temps. L'étudiant doit pouvoir se confronter à plusieurs situations d'une même famille,
- La régulation des apprentissages est constante au cours de l'année. Elle est formalisée par des évaluations formatives de travaux semestriels obligatoires et un bulletin semestriel descriptifs précis pour les compétences particulièrement développées jusque là ; il comporte un projet de développement pour la suite et intègre l'autoévaluation de l'étudiant,
- Les travaux réalisés montrent les actions réalisées, la combinaison des ressources utilisées (savoir, savoir-faire, savoir-être, savoir sur soi), l'autoévaluation et les possibilités de transfert,
- Les évaluations sont toujours critériées, les critères sont connus d'avance des étudiants,
- L'évaluation sommative se déroule dans la continuité de l'évaluation formative.

Outils d'évaluation

Ils sont au nombre de cinq, choisis pour leur potentiel par rapport à l'apprentissage, leur intégration à une approche réflexive, et leur pertinence par rapport à la professionnalisation en soins infirmiers. Ils sont utilisés à des fins formatives et sommatives.

Ce sont :

- *Le projet de soins* : outil de travail quotidien des soignantes, base de l'offre en soins infirmiers, formalisant la contribution infirmière à la prise en charge. Il consiste en une démarche de résolution de problèmes,
- *La carte conceptuelle* : représentation graphique hiérarchisée d'un ensemble de concepts et des liens significatifs établis entre eux, donnant une vue d'ensemble qui facilite la réflexion, la compréhension et la mémorisation,

- *La base d'orientation* : outil permettant, dans un contexte donné, d'anticiper, de réaliser et d'évaluer de façon critériée une activité, un produit, fait par soi-même ou par autrui. Son utilisation favorise la maîtrise des processus d'anticipation et de planification de l'action, l'appropriation des critères d'évaluation et l'autogestion des erreurs,
- *Le récit et l'analyse d'expérience* : ils permettent, après l'action, de revenir sur des événements particuliers pour leur donner sens. Son usage aide à comprendre comment le vécu, les pratiques et les ressources cognitives s'associent dans la compétence,

Chaque semestre comporte deux de ces travaux formatifs. L'épreuve sommative de synthèse réalisée à l'école en fin d'année se base sur une situation professionnelle complexe ; elle est constituée d'une combinaison de ces différents outils (sauf le récit d'expérience), basée sur les compétences professionnelles particulièrement travaillées à l'école durant l'année et mobilise une sélection des connaissances apprises. Le 5ème outil est

- *l'analyse de traces issues du portfolio* : dès le début de la formation, l'étudiant constitue un portfolio d'apprentissage. Il y insère donc des traces pertinentes de ses réalisations et de ses réflexions (processus et résultat), issues des temps de formation à l'école et en stage. A la fin du 1er semestre (visée formative) et en fin d'année (évaluation sommative), l'étudiant doit produire un travail réflexif approfondi du développement de ses compétences professionnelles. A partir d'une expérience emblématique documentée dans le portfolio, il analyse les activités réalisées, les intentions qui portent ses actions, les conditions et prescriptions réellement influentes, met en évidence les ressources (savoir, savoir-faire, savoir-être, savoir sur soi) mobilisées et leur combinaison pour la réalisation des activités, et identifie les apprentissages généralisables et les possibilités de transfert.

Conclusion : satisfactions, difficultés et perspectives

Ces modalités d'évaluation représentent une rupture importante par rapport au vécu antérieur des étudiants. Sans accompagnement méthodologique et sans possibilité

d'exercice formatif au cours de l'année, elles ne seraient pas éthiquement acceptables. Les deux épreuves sommatives présentent une bonne complémentarité : l'épreuve de synthèse sollicite des connaissances mémorisées, le travail rattaché au portfolio valorise le parcours individuel de chaque étudiant. Les compétences travaillées particulièrement dans l'année sont évaluées, tant du point de vue des performances (action efficace) que du processus de combinaison de ressources diverses sollicitées pour agir.

Nous avons donc, grâce à ce dispositif, vraiment évalué, en situation scolaire, le niveau de compétence de l'étudiant, et non seulement les connaissances disciplinaires.

Pour les professeurs, la maîtrise de ces nouveaux outils a nécessité un grand investissement, notamment de nombreuses intervisions. L'harmonisation de l'évaluation formative, et la restitution d'évaluations descriptives précises, touchant aussi bien les processus que les résultats, demeurent des difficultés.

Si l'élaboration et la correction des épreuves de synthèse se sont déroulées relativement sans heurt, le travail d'analyse rattaché au portfolio a requis des ajustements. Un groupe restreint de professeurs travaillant en intervision conduit maintenant les séminaires d'accompagnement, l'évaluation des travaux formatifs et les épreuves sommatives. Cela a amélioré la fidélité de l'évaluation, et diminué l'anxiété des étudiants, plus à même de se concentrer sur le contenu du travail réflexif que sur la compréhension de la consigne.

Il demeure néanmoins que chacune de ces épreuves représente un poids important en crédits ECTS ; ceci devra être modifié pour le programme débutant en automne 2006 car de nouvelles prescriptions (Accords de Bologne) liées à la mise sur pied d'un programme aboutissant au titre de Bachelor in Science in Nursing.

L'évaluation est le domaine qui a constitué la plus grande rupture pour les professeurs lors de la mise en place du programme HES. Bien que certains d'entre nous vivent difficilement le fait de déléguer l'évaluation sommative, la pratique de la régulation et la référence constante aux compétences se sont intégrées progressivement à notre pratique.

VAE EN LIGNE ET PRATIQUES REFLEXIVES : LES NOUVELLES PRATIQUES DE L'EVALUATION.

Padma Emilie Gouiller , Jacques Serizel, Université François Rabelais Tours

Résumé :
Nous entrons dans l'ère de la société cognitive et du numérique avec une transformation accélérée du critère d'employabilité. Ce nouveau contexte engendre de nouveaux repères, de nouveaux référentiels et modes de vie, et remet en cause les grands systèmes de légitimité. Le projet individuel se décline dans une société traversant une crise de la représentation de l'avenir brouillé entre les âges et les générations, entre temps et contre-temps. On parle dès lors d'orientation dépressive[1], où les acquis de l'expérience sont remis en cause.

Présentation

Dans le cadre d'études menées auprès de dispositifs de formation permanente universitaire[2], nous avons pu relever des éléments transférables pour envisager tout dispositif dans le domaine de l'orientation professionnelle par Internet[3]. C'est ainsi que notre projet d'accompagnement des candidats à la VAE[4] en ligne, consiste non plus à diriger dans une certaine voie, mais à accompagner des transitions et des ruptures dans des parcours moins fréquemment linéaires, dans l'évaluation des apprentissages expérientiels. Il s'agit dès lors d'un accompagnement pour une orientation individuelle et collective, entre présentiel et distantiel, qui navigue entre horizon utopique et réalités dans une certaine alternance sociocognitive des temps, des savoirs et des acteurs[5].

D'emblée une question fondamentale se pose à nous : Comment envisager une évaluation des compétences dans des dispositifs d'alternances formatives entre présence-distance?

Comment proposer une évaluation dans un dispositif évolutif et en négociation continue? En effet, étant donné que la VAE est un système de validation des acquis de l'expérience basée sur des preuves réelles et d'explicitations verbalisées, comment entrevoir une E-évaluation ?

---

[1] Danvers F., *Les dimensions culturelles de l'orientation*, in revue POUR, n°183, septembre 2005.

[2] Serizel J., *L'utilisation d'Internet en formation : une alternance entre le physique et le réel*, in Education Permanente n°53, p 71-79.

[3] Gouiller P.E, *L'orientation professionnelle par Internet, contribution pour un accompagnement en auto-co-hétéroformation*, mémoire DIRST, Octobre 20005, UFR Tours, France.

[4] Validation des Acquis de l'Expérience, 2002.

[5] Gouiller E.P, ibid.

Notre objectif étant de proposer un accompagnement VAE en ligne, ce dispositif s'est construit sur un principe de co-construction, s'installant progressivement sur les besoins de chaque acteur, dans la philosophie de la « médiaction[6] ». L'accompagnant faisant le repère, le relais, régulateur et le référent interlocutif des candidats travaillant sur leur trajet, celui-ci utilise la méthode discursive[7] des forums pour les lancer, les initier au vif du sujet de l'expérience acquise. A travers les discussions engagées en fonction des thématiques liées à la VAE, les candidats se concertent, se confrontent, lisent et relisent, produisent individuellement et collectivement leur *epoké*[8].

Entre présence et distance, nous amenons une idée d'évaluation par la re-connaissance collective des compétences au sein des échanges à travers les forums, plus exactement par la notion de rencontres productives mobilisées et ancrées par les interactions des acteurs. Ces échanges dans un entre-deux temps du récit de pratique sont d'autant plus facilités par l'alternance de différents mondes synchrones et asynchrones, et l'alternance des lieux de formations présence-distance. Les candidats ont le sentiment de participer directement à l'évènement, à l'information de leur trajet, comme une sorte d'emprise directe sur la réalité hyperéelle, dans une certaine « maternité temporelle »[9]. C'est une histoire de co-évaluation, ou d'évaluation co-constructive, car l'identification des acquis de l'expérience est *revisitée* par le regard de l'autre[10] sur les forums et autres outils de discussions en ligne. A travers les échanges écrits des pratiques de chacun, le groupe interagit en fonction des « échos » renvoyés sur les compétences. Le prescrit laisse place au vécu grâce à cette oraliture qui comble l'absence physique, mais permet quelque part de participer directement à l'activité[11], à l'évènement de l'autre et de soi pour modifier et donner forme au trajet professionnel. L'approche de l'écrit « brut » par le moyen du tchat au cours des entretiens de mise en situation permet à tout un chacun de s'exprimer entre le prescrit et le vécu, ni complètement l'un, ni complètement l'autre : une compétence inattendue, différente, une altérité[12]. C'est une co-évaluation au cœur du métier, mais aussi au cœur de l'action par des mises en situation de compétences discutées, décrites au moyen de la conversation, des outils discursifs et formant le récit d'expériences. Ces discussions par une certaine oraliture, un langage écrit, un langage

---

[6] Serizel J. *Accompagnement médiactif et autoformation. Contribution à l'étude des interactions entre adultes en formation DESS, par le biais des outils Internet*, mémoire de master, juin 2005.

[7] Henri.F & Lundgren Cayrol K, *l'apprentissage collaboratif à distance*, Presse de l'Université du Québec, 2001.

[8] Gouiller E.P., *L'orientation professionnelle par Internet*, article POUR, n°183, octobre 2003.

[9] Stiegler B., *La technique et le temps, La désorientation*, Tome 2, éd. Galilée, 1996.

[10] Klein A., *Les homepages, nouvelles écritures de soi, nouvelles lectures de l'autre*, in Spirale, Lille, 2001.

[11] Weissberg J-L. ; *Croyances, expérimentation, fiction : la crise des croyances dans les masses m édias*, in , in revue Sociologie et Sociétés, Les promesses du cyberespace, vol. XXXII, n°2, Les presses de l'université de Montréal, p 73-79, automne 2000.

[12] Lainé A., *VAE quand l'expérience se fait savoir*, édition Eres, 2005.

commun que l'on comprend à-demi mot, mais pourtant clair et concis, sont des sortes de métaphores de l'expérience acquise. Ces métaphores réécrivent la réalité de l'activité par les échanges discursifs, la négociation et les transactions au cours des interactions, entre rhétorique et poésie, entre intimité et extimité[13], entre langage public et langage intime. L'évaluation doit se comprendre dans une alternance du langage discursif partageable pour tous, tout en étant proche des expériences archaïques de l'individu. Le récit de pratiques se coconstruit entre monologue et dialogue interlocutif par l'interpellation médiatrice, entre dialogue et récit. Plus le dialogue est « vraisemblable », plus le récit prend sens et forme dans une sorte d'entendement par dialogie interlocutive. Dialogue pris entre monologue intérieur, et conversation-transaction à plusieurs. Le récit est une alchimie de la grammaire discursive[14] et des expressions archaïques, de la rhétorique et de la poésie, et pourrait être une « métaphore vive »[15].

Tout discours, tout récit de pratique posté sur un forum ou mis en exergue à travers un tchat, est adressé à un tiers ou à des tiers. Il est influencé par son contenu par le destinataire et la représentation que le narrateur s'en fait. Le candidat en VAE ne s'adresse pas seulement à l'accompagné, mais aussi aux autres participants. Ce ne sont pas pour autant des propos éloignés de la réalité de l'activité au profit du discours de façade. En fonction des « échos », des miroirs, des questions, des réponses, des attentes, il s'agit d'un regard revisité par le regard d'autrui, un nouveau regard sur sa pratique, sur son action. Ce nouveau regard sur son expérience acquise, le candidat repère les lacunes, les manques, et les éléments qu'il ne voyait pas au début de son parcours « récité » et « revisité ». Le sujet apprend, ré-apprend la réalité de sa propre activité, à travers la re-lecture que l'autre en fait. Cette re-lecture de l'expérience passé-présente et future, c'est une co-évaluation des compétences identifiées dans une activité. Les échanges en distantiel permettent d'ancrer encore plus la réalité des activités dans le temps et dans l'espace d'un récit écrit-réécrit à chaque interaction, comme une fiction, un virtuel hyperéel, ré-accessible par un nouveau mode du langage.

Un dispositif de VAE entre présence-distance est une nouveauté dans le sens où l'évaluation n'est pas prédéterminée, mais adaptée au public, alliant pratiques d'accompagnement traditionnel et nouveaux modes d'autoformation. Elle est construite selon un référentiel réalisé au travers des échanges du groupe, et non comme une donnée préalable imposée aux participants. L'évaluation a du sens car les candidats ne la subisse pas, mais y participe dans sa construction, dans l'action.

---

[13] Klein A., ibid
[14] Ferry J-M., *Les grammaires de l'intelligence*, édition Cerf, janvier 2004.
[15] Ricoeur P., *La métaphore vive*, édition du Seuil, Paris, 1975.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

REPRESENTATIONS ET TRADUCTIONS DANS LA VALIDATION DES ACQUIS DE L'EXPERIENCE EN FRANCE

Mireille Pouget, Université de Stirling, Ecosse UK

Mots clés : validation, acteurs, traduction

Résumé :

Cet article présente une analyse des processus impliqués dans la validation des acquis de l'expérience, (VAE) en France, où la loi de modernisation a crée un accès quasi universel à la validation, obligeant les universités en particulier à reconsidérer la nature des savoirs expérientiels et leurs méthodes d'évaluation de ces savoirs. Ce papier se sert des résultats émergeant d'une recherche doctorale conduite dans une université française à travers des observations d'interactions et des entretiens semi structurés avec tous les acteurs principaux. La recherche s'est tout d'abord appuyée sur l'approche des histoires de vie et des récits de formation. Cependant, l'analyse de discours des candidats, et les interactions observées entre les acteurs ont emmené la recherche à considérer une analyse basée sur la théorie de l'acteur –réseau (ou ANT) (Callon 1986, Latour 1987) dans le but de pouvoir considérer :

1. les candidats en tant qu'acteur-réseaux, avec leur propres capacités à « construire le monde » (Latour 1999 p.20), et avec leur réseaux relationnels qui 'interactent' avec les réseaux relationnels des autres acteurs principaux de la validation, en particulier l'accompagnatrice.
2. Comment ces savoirs sont representés, traduits et re-construits par-delà les espaces temporels et physiques, « à travers les réseaux de relations reliant les objets animés et inanimés combinés et déployés par devers des contextes complètement disjoints » (Nespor 1994 :161) pour que le candidat puisse obtenir une reconnaissance officielle de la part de l'université.

Ces considérations ont des implications d'une part, pour les individus-apprenants dont les propres réseaux de construction de savoirs sont complexes, et souvent culturellement différents des réseaux formels et normatifs propres aux institutions éducatives avec lesquelles ils négocient leurs savoirs, et, d'autre part, pour les nouveaux professionnels de la validation dans leur rôle de 'traducteurs' de ces savoirs expérientiels. Le papier conclura que ANT peut offrir une façon différente de penser la validation en donnant un compte-rendu de la complexité de sa dynamique et des ses résultats.

Introduction

L'objectif de cette communication est de d'explorer la possibilité d'appréhender les pratiques de la Validation des Acquis de l'Expérience avec la théorie de l'acteur-réseau (ou ANT. Cette discussion sera basée sur les résultats émergeant d'une analyse de terrain, conduite dans le cadre d'une recherche doctorale. L'analyse de terrain s'est

faite dans le service de formation continue d'une université française. Ce texte commencera par esquisser les étapes de réflexion qui ont marqué cette recherche; en effet, il avait été envisagé au départ de 'dévoiler' ce processus de validation, que l'on pourrait aussi nommer un processus de transformation de savoirs expérientiels, dont la nature allait également faire objet d'analyse ; un processus donc perçu comme complexe, voire paradoxal (Lauriol et al 2004) et somme toute mystérieux, particulièrement en ce qui concerne la pratique de l'accompagnement, cette 'nébuleuse' (Paul 2002). L'analyse de terrain fut amorcée dans une perspective d'histoires de vie. En effet, considérer le processus de validation comme un acte de formation semblait aller dans le sens donne par Lainé (2000), à savoir que les formations basées sur l'environnement professionnel se servait aussi de récits de vie, bien que de façon plus limitée. Mais il est devenu apparent, pendant le travail sur le terrain, que l'approche propre à celle des histoires de vie n'allait pas être suffisante pour rendre compte des interactions et leurs effets durant le processus de la validation ; que le processus de 'mise en ordre'' qui pouvait être observé était un processus complexe impliquant non seulement un acteur – le candidat – mais plusieurs autres acteurs, et que le processus de 'produire du sens' était collectif aussi bien qu'individuel : un processus émergeant de plusieurs phases de négociations. La recherche, qui s'est effectuée dans une université française, avec une complète coopération du centre de formation continue, des enseignants et des candidats, a consisté en une série d'entretiens semi structurés et d'observation d'entretiens avec candidats, accompagnatrices et personnel enseignant, et en interrogation de documents tels que le dossier de validation, les dossiers des cursus, les référentiels. Cette histoire, car finalement c'est aussi une histoire de transformation et de traduction, l'histoire de cette recherche relate donc un mouvement vers une compréhension différente des processus de la validation, avec la réalisation de l'existence d'un acteur inattendu mais influent, un objet, le dossier de validation. A partir de là, il a été postulé que l'histoire de la validation serait une histoire d'acteur-réseau, ou une histoire de traductions et de transformations.

La méthodologie

Cette 'étude' combine les méthodes qualitatives de collection de données, telles que des observations d'entretiens enregistrées, entretiens semi-structurés et enregistrés avec les informants clés, analyse de discours et analyse de documents. Ces méthodes constituent la panoplie habituelle de pratiques à la disposition du chercheur, dont la description de 'bricoleur' semble apte (Denzig and Lincoln 1994 p. 20.) La recherche était concernée principalement par les questions suivantes: quels sont les processus impliqués dans la validation et transformation des acquis de l'expérience en diplômes ? Que révèlent les pratiques quant au processus de représentations mis en place par les différents acteurs et que nous disent-elles sur la nature de ces savoirs (voir Pouget et Osborne 2004) ? Ce n'est que plus tard que la question suivante prendra en compte une nouvelle énonciation, portant sur les mouvements de transformation et les réseaux interrelationnels entrant en jeu pour une traduction effective des savoirs expérientiels à fin de reconnaissance officielle.

Douze candidats furent interviewés par le chercheur, ainsi que quatre accompagnateurs/rices, six enseignants, dont trois membres de jury, dont le président

et le représentant de la branche professionnelle ; sept entretiens entre une accompagnatrice et un candidat furent observés et enregistrés, ainsi qu'un jury (dont les deux candidats furent interviewés par la suite après leur validation totale). Les dossiers, dans leurs formes complètes et incomplètes, les cursus de formation et autres documents furent aussi mis à la disposition du chercheur. L'analyse des données est toujours un travail en cours, avec à l'appui l'usage du logiciel de recherche qualitative 'NVivo'.

Les acteur-réseaux et modèle de la traduction

Il serait problématique de dire que l'on puisse 'appliquer' le modèle de la traduction pour caractériser le processus itératif qui fait que l'objet - c'est-à-dire ici le dossier de validation - est considéré comme le résultat d'une négociation entre acteurs; un objet qui est transformé selon les traductions opérées, et selon le processus d'intéressement des acteurs impliqués. Problématique parce que l'ANT est une théorie, qui, pour reprendre le dialogue du Professeur de la London School of Economics avec son étudiant, (Latour 2003, 2005) n'est utile que si on ne l'applique pas à quelque chose, et qui essentiellement « porte sur la façon d'étudier les choses, ou mieux, sur la façon de ne pas les étudier » (Latour 2005 p. 142). Propos paradoxaux s'il en fut, qui reflètent justement la difficulté de résumer de quelque façon cette théorie ! L'ANT est avant tout une réponse aux théories du social qui distinguent entre les agents (ou la puissance d'agir[1]) et structure, agent étant toujours agents humains et structure étant la société et le contexte social, de quelque façon qu'ils soient analysés (Law 1999, Latour 1999), voire même expliqué à ces mêmes acteurs qui ont fourni la source de l'analyse. L'ANT, par contre, part du principe qu'il faut laisser aux acteurs un certain espace pour s'exprimer; que les acteurs sont ceux qui « font tout, même leurs propres cadres explicatifs, leurs propres théories, leurs propres contextes, leurs propres métaphysiques et même leurs propres ontologies … » (Latour 2005 p.147). L'observateur (dans notre cas ici, la doctorante) n'imposera pas un cadre d'analyse, mais « suivra les acteurs pour identifier comment ceux-la définissent et associent les éléments différents avec lesquels ils construisent et expliquent leur monde, que ce soit celui du social ou du naturel » (Callon 1986). Mais ANT est aussi une théorie qui considère humains et non-humains comme acteurs égaux, refusant la dichotomie du monde social et du monde naturel. John Law (2000 p.1) définit un objet comme un effet provenant d'un « éventail de relations », c'est-à-dire que l'objet est aussi l'effet d'un réseau ; une approche, dit-il, qui tire ses sources de la version post-structuraliste de la sémiotique ou chaque terme dépend pour sa signification de sa relation avec un autre terme, en particulier d'une relation de différence.

Avant de considérer le modèle que nous apporte Callon dans son article séminal ayant trait à des coquilles Saint-Jacques et à des marins pêcheurs dans la baie de St. Brieuc (Callon 1986), il importe de définir avant tout les acteurs principaux de la validation, et ceux dont cet article va parler principalement.

---

[1] Le terme anglais '*agency*' n'a pas de traduction facile en Français, à moins que l'on ne s'amuse aux néologismes peu élégants *agentisation* ou *agencé…*

Dans le cadre restreint de cette communication nous ne considérons que trois acteurs, le candidat, le dossier et l'accompagnatrice. Cependant il faut faire référence ici aux autres acteurs, tout aussi influents, tels que le jury et les membres du jury, les enseignants-chercheurs et spécialistes de leur discipline, directeurs de cursus de formation universitaire, le centre de formation continue et ses bureaux ouverts au public, les documents de cursus et les référentiels tels que le ROME[2], et les employeurs, les collègues, les agences de l'emploi. Il est important de se rappeler que les distinctions ne sont pas si distinctes: les frontières sont floues, les mouvements fluides entre l'institution et son environnement, intérieur et extérieur, par exemple entre les enseignants chercheurs et les professionnels de l'industrie avec qui ils travaillent en collaboration dans leurs projets de recherche ; ou le candidat à la validation qui par exemple est aussi un employer prenant des étudiants stagiaires en informatique.

Le candidat : sans doute au centre du processus nous observons le candidat de la validation, qui, selon Law (1992 p 4) en tant qu' acteur, est un réseau de relations d'éléments hétérogènes, ou « un effet produit par ce réseau », ce qui signifie que le travail du candidat (la construction du dossier), devient une représentation, sa propre re-construction de son monde, lui-même fait de réseaux relationnels où s'assemblent des éléments hétérogènes, animés et inanimés, par delà le temps et l'espace. Les candidats sont très capables 'd'expliquer', de décrire, d'analyser eux-mêmes, de reconstruire et de traduire leur monde professionnel et extra-professionnel, pour l'accompagnatrice durant leurs entretiens, sur les pages et dans les cases du dossier, puis devant le jury avec leur dossier comme témoin de tout ce travail ; les objets, ou entités qui composent ce monde sont multiples, omniprésents et relationnels : l'importance des machines, des ouvrages produits (par exemple les œuvres d'art – les ponts - accomplies par le candidat ingénieur civil), les équipes dont le rôle est considéré presque unanimement essentiel, les documents tels que réglementations ou protocoles qui jonchent la narration des candidats.

Le dossier : il serait possible de postuler que le dossier représente la tension que Law (1999 p. 5) qualifie d'oxymoron, dans le sens qu'il reflète deux intentions contradictoires :, lui-même un produit hétérogène (dans la mesure où il consiste en différents éléments : lettres, annexes, documents originaux, copies, articles de journaux, photos etc…) son rôle est de 'centrer', d'organiser, de 'mettre en ordre' une réalité disparate, hétérogène, touffue, 'dé-centrée', pour reprendre les termes de John Law, qu'est la vie professionnelle (et extra-professionnelle) des candidats. Mais c'est une mise en ordre qui est « contestable et souvent contestée » (Law 1992 p 5) tout du moins au début de la construction du dossier, parce qu'il y a résistance à cette mise en ordre, résistance à se conformer au cadre prescrit du dossier, avec ses colonnes, ses cases, ses en-têtes de pages. Law va jusqu'à dire que « l'analyse de la lutte sur cette mise en ordre est centrale à la théorie de l'acteur-réseau. Le but est d'explorer et de décrire le processus local de *mise en forme,* l'orchestration sociale, la mise en ordre et la résistance » (Law 1992 p 5).

[2] Répertoire Opérationnel des Métiers et des Emplois

Cela signifie que le candidat et l'accompagnatrice sont engagés dans une 'mise en ordre' des savoirs expérientiels, pour ne pas dire de l'expérience – des candidats. Les observations d'entretiens entre candidats et accompagnatrice, et les entretiens conduits avec ces acteurs séparément montrent comment le dossier est construit souvent avec une résistance initiale de la part des candidats a 'rentrer dans les normes' prescrites du dossier :

> *« j'ai eu du mal avec Madame Colette, de la canaliser …alors c'est vrai je suis pas dans l'attitude de….. j'aime pas froisser les gens…j'ai essayé d'imposer, vraiment de tenter …ces 5 thèmes, je sentais que c'était vraiment capital pour elle, donc c'est vrai je voulais pas la froisser… pour dire attention… il faut renter dans ce dossier normé… d'un autre côté je suis toujours partagée à me dire … il faut laisser la liberté aux gens de s'exprimer au travers de ce dossier normé, mais à leur façon aussi… donc c'est vrai je suis partagée par rapport à ça, laisser la liberté aux candidats d'écrire comme ils le souhaitent, dans la forme, ou, ….leur imposer ce dossier de 6 cases qui est quand même figé…. donc j'ai senti que Madame Colette… elle avait fait un beau travail autour de ces 5 thèmes, qu'elle avait vraiment repris … elle s'est appropriée le référentiel, donc on a trouvé l'alternative, de dire, on l'intègre qd même en sous-partie… y aura une partie de présentation sous-normée, et une partie fondamentale…. » Alice, accompagnatrice, entretien avec MP.*

L'accompagnatrice reconnaît la résistance de sa candidate, qui a construit, qui a travaillé ses propres thèmes et tient à les garder, même si ça ne se conforme pas au cadre du dossier. L'accompagnatrice aussi résiste à la candidate, car, si la candidate organise son monde à sa façon, l'accompagnatrice représente tout le réseau universitaire, administrative, politique et gouvernemental qu'est la validation des acquis de l'expérience. Ce processus de mise en ordre, similaire à l'action de rationalisation (Moser and Law 1999) est aussi un processus de traduction, lui- même suggérant un processus de transformation : l'expérience est transformée par la négociation entre l'accompagnatrice et la candidate.

> *« c'est surtout ça, refaire resurgir des choses qu'on avait oublié, comment restructurer un petit peu, bien qu'il y ait une trame… euh ! Reclasser, je sais pas si c'est sur tous les dossiers pareils …où on doit parler en fait de ses compétences entre le savoir faire, le savoir être…[…] donc j'avais quand même travaillé quand même un petit peu […].alors, on a triché avec Alice, parce que moi j'avais dis domaine de compétence et compétences… j'y arrivais pas…[…] et les savoir-faire, on a scindé à nouveau par rapport à la gestion du personnel, puisque moi, c'est propre à la gestion du personnel, à l'administration du personnel, propre au métier véritablement…[…] donc j'ai rescindé ensuite par rapport au domaine savoir faire dans le domaine des ressources humaines […] oui, ben moi j'avoue que, alors ce qui a été le plus difficile au final, c'est la lettre de motivation car j'arrivais pas à retranscrire… en fait ce que j'avais dans ma tête… » Thérèse.*

« Ca était plus long de …de chercher exactement tout ce qui…qui était en relation et d'écrire tout ce qu'il y avait quoi, de savoir c'était ça surtout, et de savoir par où, de quel côté commencer, par ou attaquer, et après s'orienter, en lisant le dossier on se demande même à la fin si on est vraiment…si on répond vraiment aux attentes…. on arrive …moi je suis arrivé maintenant à me demander si je répond vraiment aux attentes… me dire…c'est ça qu'y veulent ou pas, parce qu'alors là je sais plus, c'est c'que j'ai fait, maintenant est-ce que ça correspond vraiment au dossier, en regardant le diplôme, .on se dit ..ben merde ….est-ce que ça correspond vraiment au dossier ? [….] » Christian

« On met toutes les expériences qu'on a pu avoir sur le papier […] et on essaie de les coordonner et de les organiser [..] c'est bcp d'organisation parce qu'en fait heu…car au départ on a …heu…plein d'idées de taches que l'on peut effectuer mais en fait chaque tâche que l'on peut effectuer se décompose ensuite en plusieurs tâches …une tâche que l'on peut effectuer dans un domaine peut être la même dans un autre domaine donc en fait c'est beaucoup de regroupement… […]Que c'est en fonction d'un référentiel disons que pour une licence y'a plusieurs matières a faire et en fonction des matières on essaie un petit peu de recaser les expériences dans chaque matière […] Beaucoup de temps de réflexion, de mise en forme » Julien

L'accompagnatrice: il serait tentant de se reposer sur le discours humaniste Rogérien de développement personnel, qui nous a donné les pratiques de guidance, d'accompagnement, de mentoring, coaching répertoriées par Paul (2002) et évoqués par Pouget (2006) . Usher and Edwards' (2005), d'un autre cote, analysent ces pratiques en les rapprochant du terme utilisé par Foucault « pouvoir pastoral » qui est exercé dans ce qu'ils appellent les relations confessionnelles typique de la relation de guidance. Ils se servent de l'analyse de Foucault pour discuter du phénomène d'internalisation de la 'discipline' ou 'gouvernementalité'

« lorsque la discipline devient discipline de soi, occultant ainsi les formes de pouvoir exercées de l'extérieur » (p 400). Usher et Edwards identifient, parmi les pratiques de guidance telles que les pratiques de plans d'action, reconnaissance des acquis, évaluation basée sur portfolio un glissement de pouvoir disciplinaire à un pouvoir pastoral » (p.400). Dans cette perspective, il est clair que le rôle de l'accompagnatrice est effectivement un rôle confessionnel où justement elle devient le véhicule invisible d'un pouvoir situé ailleurs, dans les règles du jeu décidées par, non seulement le jury, les enseignants, l'université, mais tout un réseau d'entitités mouvantes ou, comme le dit le candidat Julien, « le système »…

« on accepte les règles du jeu, c'est comme ca, le système est d'une telle façon et puis en fait on est obligé de le faire ; on est obligé de passer par là […] disons qu'à partir du moment ou y'a des règles qui ont été mises en place …elles ne sont peut être pas toutes bonnes, c'est peut être pas à moi d'en juger non plus … mais à partir du moment où les règles ont été fixées par les personnes qui vont vous juger, ça parait difficile de pas les suivre …» Julien

Cependant, en ANT, le rôle de l'accompagnatrice n'est pas de considérer les candidats comme des individus isolés avec leur perspective psychologique unique, mais plutôt, son rôle entre dans une perspective ou « le point focal est sur le social, sur les interactions mutuelles de facteurs multiples » (Usher and Edwards 2005 p. 406). De plus, l'acte de guidance en ANT peut être considéré comme un processus produisant des savoirs, où les savoirs produits sont la parole, ou […] les savoirs sont un discours qui passe par les transactions des dialogues de l'entretien d'accompagnement.

Les candidats le comprennent parfaitement :

> *« le rôle de l'accompagnatrice c'est de mettre au clair les idées que l'on peut avoir, c'est vrai que quand je regarde le 1er dossier que j'ai fait, c'etait quand même très brouillon, car je n'arrivais pas à organiser mes idées…et le dossier de maintenant c'est vrai qu' y'a eu beaucoup de coup de pouce pour dire comment l'organiser….[…] La relation…c'est un travail où on recherche énormément de choses sur soi même et c'est pas toujours évident de les étaler en quelque sorte devant quelqu'un d'autre. » Julien*

Ici je retourne à l'idée que la construction du dossier de validation (les savoirs produits par cet acte de guidance) représente un travail - tous les candidats interviewés reprennent cette idée de 'travail' [3] - précisément parce que cela génère une résistance : les éléments hétérogènes dont parle Law iront ou n'iront pas facilement se laisser organiser dans l'ordre prescrit du dossier, comme c'est le cas pour Colette et les autres candidats interviewés. Les entretiens donc, entre l'accompagnatrice et le candidat deviennent un « réseau mobilisé d'acteurs » (Usher et Edwards 2005 p. 407), rassemblés à travers le temps et l'espace, et incluant la hiérarchie de l'université, les enseignants-specialistes de leur discipline, le jury, les cursus, les référentiels, les collègues, le lieu professionnel avec ses collègues, machines, documentations etc.…tous conjurés dans le processus de dialogue et processus d'écriture de l'accompagnement et de la validation.

Modèle de la traduction ou la Théorie de l'acteur-réseau (ANT)

Callon (1986) suggère dans son article que ses trois biologistes chercheurs sont passés par 'quatre moments' de traduction dans leurs efforts stratégiques de s'imposer et d'imposer leur définitions de la situation aux autres (les pêcheurs en particulier), s'engageant dans ce qu' ANT appelle la controverse : c'est par la controverse que s'élaborent les faits, et elle précède toujours l'émergence d'un énoncé scientifique et d'une innovation[4]; en l'étudiant, on peut donc percer le processus qui permet au fait de se construire. Les quatre moments de traduction suggérés par Callon sont : la problematisation, l'intéressement, l'enrôlement et la mobilisation.

---

[3] En Anglais non seulement travail est *''work'* mais aussi *'labour'*, c'est-à-dire également mettre au monde, *'être en travail'*.

[4] Lauriol et al (2004) identifient bien la VAE comme un processus d'innovation juridique et social.

Problématisation : Le terme renvoie à la façon dont les acteurs-réseaux (les enseignants/ directeurs de formation, les accompagnatrices et directrice du centre universitaire de formation continue qui tous connaissent et contrôlent les règles du jeu) définissent les identités permissibles et intérêts pour les personnes (Nespor 1994) telles que les candidats ; les candidats, quant à eux, ayant identifiés leur 'projet', c'est-à-dire la formation qu'ils considèrent compatible avec leur expérience, devront passer une première fois par ce que Callon appelle le 'point de passage obligé', PPO, c'est-à-dire le moment où l'enseignant/directeur du programme visé, établira, en discussion avec le candidat et avec documents a l'appui (CV, preuves de formation et qualifications antérieures) le début du processus en donnant son accord (se basant sur ce qu'il considère, avec le candidat, le caractère réaliste ou pas du projet). Bien que la controverse n'est pas aussi forte que dans le cas donné par Callon, il est stipulé ici que ce stade correspond à un premier PPO dans la mesure où chaque acteur a des intérêts qui puissent ne pas converger. Cela dépendra après tout du 'réalisme' du candidat quant à l'évaluation qu'il fait de son expérience et de ses savoirs, et de la disposition de l'enseignant quant à la VAE, les savoirs du candidats et autres facteurs. Nous stipulons que le candidat et les autres acteurs, en particulier l'accompagnatrice, le jury et le candidat, auront à passer par un deuxième PPO lors de l'élaboration du dossier.

Intéressement : Le succès de la validation dépendra des négociations entre tous les acteurs, et de la transformation du monde professionnel ou social du candidat en un objet-lien, objet négoce et objet négocié, son dossier, son alter ego, comme ce candidat, chercheur d'emploi à 53 ans, désespéré d'être ainsi « sur la touche depuis 2 ans », qui a son dossier avec lui dans tous ses déplacements de chercheurs d'emploi :

> *« [le dossier] Je l'ai dans ma mallette …c'est un apport potentiel … je suis en quête d'une entreprise qui veuille bien m'inviter à déballer mon potentiel. [il rit] » Jean-Marie*

Le processus d'intéressement peut être décrit grâce aux procédés à travers lesquels les candidats 'voyagent' : le rendez-vous avec l'enseignant-directeur de la formation, le rendez-vous initial avec l'accompagnatrice principale du centre de formation continue, où la structure législative et organisationnelle de la validation est présentée au candidat, les niveaux d'acquis discutés, une stratégie et des objectifs mis en place, et ou le dossier littéralement voit le jour avec les premières fiches identitaires du candidat. On octroie une accompagnatrice au candidat et le premier entretien, durant jusqu'à deux heures, est fixé.

A partir de là l'accompagnatrice prend le rôle de porte-parole, car son travail est de représenter, de définir, tenir en place, ou mobiliser tout un assortissement de pièces et morceaux juxtaposés.

> *« oui, pour moi c'est réellement une co-construction, c'est-à-dire, je fais équipe avec le candidat[…] Faut vraiment faire partie intégrante de sa démarche, comme ça fonctionne , pour continuer à le mobiliser… sa démarche, c'est pas une démarche facile, pour traduire et interpréter ce qu'on attend de lui, et lui faire connaître le monde universitaire, qui n'est pas simple pour les gens qui viennent du privé […] Donc On est vraiment traducteur à un moment donné, interprète[…]c'est là toute l'ambiguïté de notre démarche… c'est-à-dire on est une équipe… ça nous demande d'avoir*

un œil averti, de garants des preuves et de la cohérence du parcours…c'est là l'ambiguïté du poste » Accompagnatrice, entretien avec MP.

En définissant les limites et possibilités du dossier et de son 'formatage', elle établit ce processus d'interessement car, omniprésent dans son discours, est le jury, dont elle répète à chaque entretien qu'il est « souverain »; omniprésentes aussi sont les possibilités d'interprétations et de choix que le candidats peut faire à propos de ses propres savoirs et compétences, mais que l'accompagnatrice va en quelque sorte guider pour que le candidat éventuellement rencontre, métaphoriquement, puis concrètement, les « besoins » du jury. Pendant ce temps le dossier prend forme, et prend vie, littéralement il prend part aux discussion : il 'vit' sur le bureau de l'accompagnatrice, enter les deux interlocuteurs, il est manipules, tourne, feuilleté, dépecé, remis en place, relié, défait, emballé dans l'attaché-case, voire le sac en plastique. Tous les yeux sont sur lui.

L'intéressement de l'enseignant, qui sera le président du jury, est essentiel : le fait qu'il soit enseignant de la formation continue dans son département en fera un allié du centre de formation continue, bien que son alliance ne soit pas nécessairement pas un donné d'avance. « Etre intéressé est être inter-essé, être interposé […]. Intéresser d'autres acteurs, c'est construire des mécanismes qui peuvent être placés entre eux et toutes autres entités qui veulent définir leur identité autrement » (Callon 1986 p. 208). La compétition pour les enseignants vient d'autres alliances, d'autres discours sur la signification et le rôle de l'université et de ses diplômes, sur la problématique que pose l'évaluation de la VAE, ou elle peut émaner de collègues, de leurs étudiants de formation initiale, reflétant l'instabilité du nouvel environnement crée par la VAE.

Enrôlement : C'est à se stade que l'accompagnatrice et le candidat négocient des stratégies concernant le format et le contenu du dossier pour le rendre lisible pour le jury. Là, le problème est de présenter, traduire la masse d'informations déjà écrites produite par le candidat, qui travaille de longues heures chez lui a rassembler son 'monde' de souvenirs d'incidents et situations de travail ou expériences extra-professionnelles qui peuvent illuminer ses compétences et savoirs. L'élaboration du dossier passe par une description détaillée des acteurs invisibles de l'expérience des candidats. Des candidats offrent une résistance, le dossier aussi résiste : ses cases sont trop petites, les en-têtes pas au bon endroit…l'accompagnatrice doit rappeler les contraintes et demandes du jury futur constamment.

« le travail est là… c'est-à-dire de faire une liste… de repérer de façon exhaustive toutes les compétences que vous avez engrangées en liant avec les objectifs du diplôme, ne pas en oublier …pour que le jury ait une visibilité complète de votre évolution…[votre] parcours. C'est là le travail, ne pas en oublier…[…] je veux une visibilité de votre parcours…et que ce soit transcrit . » Accompagnatrice au candidat Fabien

Dire que le processus d'enrôlement soit un interessement réussi (Pol et al 2006) signifie que les acteurs, y compris les enseignants-décideurs, à travers leur porte-parole qu'est l'accompagnatrice, ont effectué un alignement à travers les négociations et transactions réelles et interposées.

Mobilisation : « Mobiliser, c'est, comme le mot l'indique, rendre mobiles les entités qui ne l'étaient pas auparavant. […] Tous les acteurs sont d'abord dispersés, puis rassemblés à un certain endroit et à un moment particulier. Cette mobilisation ou concentration a une réalité physique définie qui est matérialisée à travers une série de déplacements » (Callon 1986 p.217).

Dans le cas de la VAE, les entités de la vie (professionnelle par exemple) du candidat sont rassemblées en forme stable et composite dans le dossier, et devant le jury durant les vingt minutes de présentation finale du candidat. L'accompagnatrice, qui a mobilisé le candidat et le jury, interpelle ce dernier pour justement lui rappeler tout le réseau de complexité, législatif et pratique, à laquelle appartient la VAE, en leur énonçant leur devoir de jury, et en 'résumant' le candidat. L'accompagnatrice rassemble les acteurs invisibles de la VAE et se porte garant du travail du candidat. Comme elle le dit :

> *« le rôle d'accompagnateur c'est aussi un rôle d'évaluation et de recherche de preuve on est la aussi on est les garants de tout ce qui est écrit et de la formalisation du discours du candidat » Accompagnatrice : entretien avec MP.*

Les enseignants présents eux rallient leur discipline, la formation, le diplôme. Mais le jury est un lieu d'interessement réussi : en effet il a lieu parce que le président a déjà rencontre le candidat au début du processus, et les membres du jury ont déjà lu le dossier auparavant ; le jury est la pour confirmation, pour élucider des questions, pour entendre le candidat parler de son histoire.

Conclusion

Le but de cet article etait de poser un autre regard sur les processus de la validation à travers la théorie de l'acteur-réseau. Il n'est pas encore certain que cette exploration, puisque c'en est une, ait été entièrement réussie. Cependant, pour reprendre Callon (1986) « traduire c'est déplacer » et on peut sans doute être d'accord sur l'existence de nombreux déplacements: les savoirs expérientiels des candidats sont déplacés de leurs habitats naturels, l'accompagnatrice les traduit, traduit les candidats, traduit le dossier lui- même. Des savoirs sont produits à partir de ces mouvements, produits individuellement mais aussi et surtout collectivement. Latour lui-même (2003, 2005) insiste que c'est sur le labeur, le mouvement, le flux et les changements qu'il faut mettre l'accent. De cet enchevêtrement né des descriptions – et pour l'ANT la description du micro-social est primordial – c'est le rôle de l'accompagnatrice qui va sans doute attirer l'attention. En effet, c'est sur cette profession inédite (Mayen 2004) qu'advient un rôle de traductrice de savoirs et peut-être que l'analyse plus approfondie mettra les transactions entre l'accompagnatrice, les candidats et leurs dossiers en plus grand relief. Moser et Law (1999) organisent leur texte en parlant d'histoires *(stories)* où ils décrivent des passages racontés, entre le visible et l'invisible, le privé et le public. Ce sont ces passages à qu'il faut sans doute faire attention, à serrer de près les descriptions de leurs 'histoires' données par les candidats, qui se révèlent, comme Liv dans l'histoire de Moser et Law, des agents autonomes, compétents, et capables d'exercer cohérence, discrétion et de produire du

sens dans la grande confusion qu'est l'expérience. Le mystère de la VAE n'est peut-être pas à 'dévoiler' simplement, mais à suivre dans ses déroulements.

Références

Callon, M. (1986) Some elements of a sociology of translation: Domestication of the scallops and the fisherman in Law, J. (Ed) *Power, Action and Belief: A New Sociology of Knowledge?* Sociological Review Monograph, No 32 (University of Keele) London: Routledge and Kegan Paul

Lainé, A. (2000) L'histoire de vie, un processus de "métaformation. *Education permanente, No 142, 29-43*

Latour, B. (1986) The powers of association, in Law, J. (Ed.) *Power, Action and belief. A new sociology of Knowledge?* Routledge & Kegan Paul plc. London

Latour, B. (1987) *Science in Action.* Cambridge, MA: Harvard University Press

Latour, B. (1999) On recalling ANT in Law, J., Hassard, J. (Eds) *Actor Network theory and After*. Blackwell Pubs. Oxford

Latour, B. (2003) Comment finir une thèse de sociologie. Petit dialogue entre un étudiant et un professeur (quelque peu socratique)
http://www.ensmp.fr/~latour/articles/article/90-DIALOGUE%20ANT-FR.html
Accès le 80906

Latour, B. (2005) *Reassembling the Social. An Introduction to Actor-Network-Theory*. Oxford University Press, Oxford

Law, J. (1986) (Ed.) Editor's introduction : Power/Knowledge and the dissolution of the sociology of knowledge in Law, J. (Ed.) *Power, Action and belief. A new sociology of Knowledge?* Routledge & Kegan Paul plc. London

Law, J. (1992) Notes on the Theory of the Actor Network: Ordering, Strategy and Heterogeneity. Published by the Centre for Science Studies, Lancaster University, Lancaster LA14YN
http://www.comp.lancs.ac.uk/sociology/papers/Law-Notes-on-ANT.pdfn Accès le 30.08.06

Law, J. (1999) After ANT: complexity, naming and topology, in Law, J., Hassard, J. (Eds) *Actor Network theory and After*. Blackwell Pubs. Oxford

Law, J. (2000) Objects, Spaces and Others. Published by the Centre for Science Studies, Lancaster University, Lancaster LA14YN
http://www.comp.lancs.ac.uk/sociology/papers/Law-Objects-Spaces-others.pdfn
Accès le 30.08.06

Lauriol, J., Guérin, F., Zannad, H. (2004) *La validation des acquis de l'expérience en France: un objet frontière?*
At http://www.agrh2004-esg.uqam.ca/pdf/Tome3/Lauriol_Guerin_Zannad.pdf Accès le 08.09.06

Mayen, P. (2004) Caractériser l'accompagnement en VAE. *Une Contribution de didactique professionnelle. Education permanente, No 159 -2, 7-22*

Moser, I., Law, J. (1999) Good passages, bad passages in Law, J., Hassard, J. (Eds) *Actor Network theory and After*. Blackwell Pubs. Oxford

Nespor, J. (1994) *Knowledge in Motion. Space, time and Curriculum in Undergraduate Physics and management, 131*. London: the Falmer press

Paul, M. (2002) L'accompagnement : une nébuleuse. *Education permanente, 153 (4),* 43-49.

Pol, G., Legardeur, J., Minel, S., Merlo, C. (2004) Les phases informelles en amont des projets de conception. http://www.ergoia.estia.fr/documents/pol_legardeur_final.pdf Accès le 13.08.06

Pouget, M. (2006) Pedagogical and Social Aspects of APEL in Corradi, C., Valk, A., Van der Kamp, M. (Eds) Making LLL societies a reality: Recognition of formal, non-formal and in-formal learning in European universities. *University of Tartu http://www.eullearn.net/*

Pouget, M. & Osborne, M. (2004) Accreditation of Prior Experiential Learning: Knowledge and *Savoirs* in France – a Different Perspective? *Studies in Continuing Studies, 26 (1),* 45-65.

Rose, N. (1999) *Powers of Freedom: Reforming Political thought. 161*. Cambridge: Cambridge University Press

Usher, R., Edwards, R.G. (2005) Subjects, networks and positions: Thinking educational guidance differently. *British Journal of Guidance & Counselling, 33(3) 397-410.*

LA PLACE DES VARIABLES D'EVALUATION DANS LA PLANIFICATION ET LA GESTION DE PROCESSUS D'ENSEIGNEMENT-APPRENTISSAGE A L'ECOLE PRIMAIRE

Teresa Grange Sergi, Université de la Vallée d'Aoste (Italie)

Mots-clés : culture d'évaluation, variables d'évaluation, méthodes d'évaluation

Résumé:

La critique de la réforme scolaire introduite récemment en Italie, touchant aussi au système d'évaluation des élèves, a entraîné chez les enseignants l'émergence de représentations, de pratiques consolidées, d'orientations théoriques, de savoirs d'expérience qui constituent autant d'horizons de sens pour une réflexion autour de la *culture de l'évaluation* dans laquelle on développe les projets pédagogiques, on assume les décisions didactiques, on réalise les pratiques d'enseignement-apprentissage. Cette culture de l'évaluation sous-jacente à la décision et à l'action pédagogiques revêt un double intérêt, de recherche et de formation, pour la compréhension des pratiques d'évaluation dans leurs liens avec les pratiques d'enseignement-apprentissage et pour la construction pertinente de programmes de formation professionnelle des enseignants visant à la promotion de la qualité pédagogique à l'école.

Dans le cadre des recherches sur la culture de l'évaluation scolaire chez les enseignants (Becchi & Vertecchi, 1986; Pellerey, 1998 ; Vertecchi, 2003) et dans le contexte de questionnement éveillé par la réforme, nous avons conduit une étude ayant pour objet les variables d'évaluation que les enseignants déclarent de considérer dans leurs pratiques, la façon dont ils les opérationnalisent et la place qu'ils leur allouent dans les processus d'enseignement- apprentissage. Par l'interview de 25 enseignants d'école primaire, on a repéré, classé et analysé les variables d'évaluation considérées ou utilisées par les enseignants par rapport, notamment, aux fonctions de l'évaluation (Gosling, 1992 ; Allal, 1999), aux acteurs de l'évaluation (Perrenoud, 1998) , aux situations d'enseignement-apprentissage planifiées et mises en œuvre (Farr & Tone, 1998).

L'article discute l'écart conceptuel relevé entre didactique et évaluation et, au-delà de la situation italienne contingente, le rôle de la culture de l'évaluation dans la contextualisation de toute hypothèse d'innovation des pratiques d'évaluation à l'école.

Problématique et contexte

En Italie, dans un cadre législatif qui reconnaît aux écoles une pleine autonomie quant aux modalités et aux critères d'évaluation des élèves par rapport aux finalités et aux objectifs fixés par les programmes nationaux, une innovation des documents d'évaluation et de certification a été introduite par la récente réforme du système scolaire. Les ambigüités, les contraintes procédurales et les implications pédagogiques de ces nouvelles formes de certification ont suscité, au-delà des phénomènes de résistance et de marchandage qui accompagnent les phases initiales de toute innovation centrale (Huberman & Miles, 1984), un débat animé au niveau des acteurs et des chercheurs de l'éducation, qui a mis en cause et renvoyé de fait l'application des directives ministérielles en matière d'évaluation.

La critique de la réforme a également provoqué chez les enseignants l'émergence de systèmes de croyances, de pratiques affermies, d'orientations théoriques, de savoirs d'expérience qui représentent autant d'horizons de sens pour une réflexion autour de la *culture de l'évaluation* (Becchi & Vertecchi, 1986; Pellerey, 1998 ; Perrenoud, 1998 ; Vertecchi, 2003) dans laquelle on développe les projets pédagogiques, on assume les décisions didactiques, on réalise les pratiques d'enseignement-apprentissage. Cette culture de l'évaluation sous-jacente à la décision et à l'action pédagogiques revêt un double intérêt, de recherche et de formation, pour la compréhension des pratiques d'évaluation dans leurs liens avec les pratiques d'enseignement-apprentissage et pour la construction pertinente de programmes de formation professionnelle des enseignants visant la promotion de la qualité pédagogique à l'école.

Dans le construction conceptuelle dénommée *culture de l'évaluation* convergent les problématiques spécifiques à des composantes différentes de l'objet « évaluation », tout comme les questionnements liés au traitement de l'échec scolaire et au rôle de l'évaluation par rapport aux finalités de l'école. En Italie, les études de Vertecchi montrent, par exemple, qu'une culture de l'évaluation est encore à construire, que la présence capillaire de convictions persistantes relevant d'une tradition pédagogique développée pour une école à base sociale restreinte entrave une définition lucide et pertinente des compétences professionnelles nécessaires aux enseignants en matière d'évaluation, qu'une approche scientifique à l'évaluation fatigue à prendre le relais des références moralistes ou spontanéistes du métier. Il est, par ailleurs, frappant de relever que cette pauvreté de fond des pratiques et des conceptions de l'évaluation scolaire se manifeste dans un contexte imprégné d'évaluation, avec un débat permanent sur la nécessité de tout évaluer (du système scolaire dans son ensemble aux projets pédagogiques de classe, des compétences des enseignants aux acquis des élèves, du fonctionnement des organisations et des bureaucraties à la réussite d'une initiative de partenariat, et ainsi de suite), sur l'urgence de se donner des moyens appropriés, valides, efficaces et éthiquement acceptables. Il est possible qu'une *négociation de sens* - dont on a fait trop souvent l'impasse en privilégiant le raccourci de la prise directe sur l'intervention - par le dévoilement et la discussion des implicites, des présupposés et des enracinements des

principales résistances au changement puisse ouvrir à des nouvelles pistes : c'est justement par un travail des représentations qu'une approche réflexive au développement professionnel établit son environnement de pensée et d'action.

Dans ce climat de sensibilité accrue des enseignants aux problèmes de l'évaluation éveillé par le contexte de réforme, nous avons conduit une étude qui, tout en effleurant le spécifique des nouveaux instruments de certification, a élargi la perspective sur les représentations de l'évaluation scolaire qui sont à la base des attitudes et des comportements professionnels en acte chez les enseignants.

Questions de recherche

Dans le cadre des recherches sur la culture de l'évaluation scolaire chez les enseignants, nous avons focalisé des composantes particulières de cette culture, telles que les attitudes face à l'évaluation (Gosling, 1992; Vertecchi, 1993) et les supports conceptuels et instrumentaux de l'évaluation utilisés par les enseignants (Farr & Tone, 1998; Varisco, 2004). Nous avons essayé de cerner les variables d'évaluation que les enseignants déclarent de considérer dans leurs pratiques, la façon dont ils les opérationnalisent et la place qu'ils leur allouent dans les processus d'enseignement-apprentissage.

Méthodologie et procédures

On a interviewé 25 enseignants d'école primaire, le primaire étant le degré scolaire le plus directement affecté par la réforme de l'évaluation. Le choix des sujets a été fait sur la base d'une motivation personnelle à participer à une recherche sur l'évaluation scolaire ; d'une expérience d'enseignement d'au moins 10 ans de service en titulaires, assurant l'exercice dans la durée de plusieurs formes et modalités d'évaluation au niveau central et d'établissement; d'une répartition variée sur des écoles différentes afin de neutraliser la variable « culture d'établissement » et multiplier ainsi les apports et les perspectives. Le protocole d'entretien vis-à-vis a été construit en deux parties: une première section demi-structurée, à questions ouvertes, et une deuxième section structurée, composée de 70 item relevant sur une échelle Likert le degré d'accord ou de désaccord des enseignants face à des propositions concernant divers aspects de l'évaluation scolaire et le niveau d'intégration ou d'exclusion de certaines pratiques pédagogiques dans l'expérience quotidienne d'enseignement.

Les entretiens ont été enregistrés et successivement transcrits pour l'analyse de contenu et pour le codage.

A partir des critères, des stratégies, des instruments d'évaluation évoqués par les enseignants nous avons repéré, classé et analysé les variables d'évaluation considérées ou utilisées dans l'organisation et la gestion de processus d'enseignement-apprentissage par rapport, notamment, aux fonctions de l'évaluation, aux acteurs de l'évaluation, aux situations d'enseignement-apprentissage planifiées et mises en œuvre.

Résultats

A partir des textes des entretiens et, en particulier, des réponses au questionnaire, nous avons dégagé 6 catégories d'analyse : les attitudes face à l'évaluation, les fonctions de l'évaluation, les acteurs de l'évaluation, les pratiques d'évaluation, la place de l'évaluation dans les situations d'enseignement-apprentissage, les causes du succès et de l'échec scolaire. Après le classement nous avons exploré des corrélations entre les variables intervenant dans les différentes catégories ; les corrélations significatives sont mises en évidence dans la présentation des résultats de l'une des catégories correspondantes.

Attitudes face à l'évaluation

L'évaluation est unanimement considérée nécessaire mais difficile. Les difficultés évoquées concernent la validité des techniques et des instruments d'évaluation et elles varient en fonction des matières à évaluer. Par exemple, on considère qu'il est plus facile (moins difficile) évaluer une habilité de calcul que des connaissances en histoire. Tous les enseignants interviewés prennent leurs distances vis-à-vis des techniques et des instruments, mais avec des modalités différentes : de l'adaptation active et intentionnelle des matériaux et des stratégies à chaque situation de classe particulière, à l'affichage d'une confiance solide dans ses propres intuitions faisant de l'expérience personnelle le point de référence principal pour procéder à une démarche d'évaluation « de bon sens ». Cette distance est aussi associée à des sentiments d'impuissance que les enseignants expriment en termes d'irréductibles incertitudes : « il n'y a pas d'évaluation objective, la même copie corrigé par dix enseignants sera notée de dix façons différentes, il y a toujours des éléments qui échappent à l'évaluation, il y a trop de facteurs intervenant dans une prestation pour arriver à tout saisir … ». Une minorité se montre découragée et se confine volontairement dans ses quelques points fermes de repère, se déclarant sceptique envers le changement, tandis que les autres se disent intéressés à une amélioration tout en précisant qu'une bonne compétence en évaluation est impossible à atteindre.

L'évaluation est vue comme un moment important pour les élèves : dans certains cas elle est considérée un facteur essentiel de motivation externe et de valorisation des tâches scolaires, dans d'autres, elle aide l'élève à prendre conscience de ses apprentissages. Deux enseignants manifestent un certain malaise à assigner des mauvaises notes et préféreraient ne pas devoir certifier les résultats. La compilation du bulletin, toujours citée dans le courant des entretiens, est vécue de plusieurs manières : comme une obligation bureaucratique sans conséquences, comme une contrainte demandant un effort de synthèse inutile qui ne rend pas justice au travail accompli par l'élève et par l'enseignant, comme une occasion importante de communication école - famille.

L'évaluation est presque toujours entendue en tant qu'évaluation des résultats. Des références à l'évaluation de processus ou de décisions sont sporadiques et englobées dans l'évaluation des résultats, comme aspects particuliers ou explicatifs.

Fonctions de l'évaluation

L'évaluation sommative est considérée au cœur des fonctions d'évaluation. Les autres fonctions sont traitées comme accessoires et parfois optionnelles. Les

enseignants utilisent un langage spécifique : évaluation formative, évaluation initiale, évaluation des pré-requis, évaluation diagnostique, régulation des apprentissages, autoévaluation, mais ils ne renvoient pas leurs propos à des cadres conceptuels définis. Plusieurs d'entre eux attribuent une fonction régulatrice à l'évaluation sommative pratiquée fréquemment. Souvent l'évaluation formative est définie comme « celle qu'on ne note pas, celle qui ne compte pas, celle qui ne fait pas la moyenne » ou encore comme évaluation *in itinere* , l'évaluation sommative étant l'étape finale de bilan. Les enseignants ayant une attitude anxieuse face à l'évaluation évoquent l'évaluation formative comme idéale, à mesure d'élève, permettant à l'enseignant une plus grande liberté de différenciation. L'évaluation formative est aussi reliée à la légitimation (voire à la valorisation) de l'erreur comme partie du processus d'apprentissage. Les enseignants qui utilisent l'évaluation formative et qui soulignent les avantages d'une « zone franche » de l'évaluation ne font que très rarement le lien entre évaluation formative et évaluation sommative, ni en termes de facilitation des apprentissages ni en relation à des résultats atteints à un moment donné. On retrouve ce découpage à propos des pratiques pédagogiques quand il émerge un très faible réinvestissement des données d'évaluation dans la planification.

L'évaluation initiale présente des ambiguïtés à propos de l'étendue de ce que les enseignants considèrent comme initial. Par exemple, on affirme que avec l'évaluation initiale on se rend compte de ce que l'élève peut ou ne peut pas faire et on formule les objectifs en conséquence ; la pratique du « nivellement par le bas » est fréquemment citée comme réponse didactique destinée aux élèves en difficulté. Dans d'autres cas, et dans une autre posture pédagogique, l'enseignant utilise l'évaluation initiale pour « repérer la zone proximale de développement de l'élève par rapport à un objectif donné » et aménager des situations d'apprentissage progressives congruentes.

Acteurs de l'évaluation

L'évaluation est perçue comme une prérogative de l'enseignant, qui peut partager certains aspects des processus d'évaluation avec d'autres acteurs à des moments précis, mais qui détient toujours le dernier mot. En particulier, la plupart des enseignants interviewés ne pense pas que les finalités ou les critères d'évaluation doivent faire l'objet d'une négociation avec les élèves ; généralement ils considèrent souhaitable (mais pas nécessaire) de les communiquer à la classe. Certains d'entre eux affirment que l'évaluation est déjà très difficile pour les enseignants, qui sont des professionnels, et que c'est démagogique de penser que les élèves puissent s'évaluer eux-mêmes correctement. D'autres soulignent l'opportunité de l'autoévaluation, de préférence en contexte ludique, et lui reconnaissent une valeur formative. Ils n'explicitent pourtant pas la nécessité d'apprendre aux élèves à s'autoévaluer: ils la proposent comme une activité spontanée dont le réinvestissement n'est par ailleurs pas mentionné dans la description des pratiques. Les parents d'élèves sont normalement informés des résultats, avec des cadences très variables ; ils sont invités à participer et à se confronter avec les enseignants dans les cas les plus délicats et les plus graves, par exemple en cas de risque de décision de doublement. La concertation avec les collègues qui enseignent dans la même classe ou la même discipline est considérée importante et productive. Les enseignants préparent ensemble ou échangent des matériaux pour l'évaluation et se consultent pour les prises de décision

les plus ardues. Certains posent comme condition de la coopération une affinité personnelle et professionnelle marquée.

Pratiques d'évaluation

Les pratiques d'évaluation décrites par les enseignants étalent une grande variété de stratégies et d'instruments. En général, le lien avec les objectifs d'apprentissage est évoqué comme principe mais on ne trouve pas de correspondances dans la description des planifications.

Les interrogations orales et les épreuves écrites préparées par les enseignants au fur et à mesure sont les pratiques les plus fréquentes. L'observation continue des élèves est une source de données jugée précieuse mais, interrogés sur la façon dont ce matériel est utilisé, les enseignants formulent des réponses vagues portant sur « la formation d'une idée sur l'élève », sur l'intérêt suscité par telle ou telle activité, sur « ce qui fonctionne bien et ce qui ne marche pas ».

Le recours à un matériel structuré est assez fréquent quand les textes ou les manuels le proposent, tout comme la construction de la part des enseignants d'épreuves à réponse fermée ou à choix multiple. La question sur la manière dont ils contrôlent la validité des instruments reste sans réponse ou fait référence à une vérification *ex post* à partir des résultats des élèves.

Place de l'évaluation dans les situations d'enseignement-apprentissage

L'évaluation est considérée comme un moment en soi dans les situations d'enseignement-apprentissage, normalement positionnée à la fin d'un processus et bien distincte, par nature et statut, de tout le reste. Même quand les enseignants planifient, l'évaluation arrive après : aucun des enseignants interviewés ne prépare l'évaluation contextuellement à la programmation des activités. La boucle post-active du retour critique sur les choix didactiques n'est citée qu'en fin de parcours, généralement après une évaluation sommative.

La différenciation des stratégies et des procédures est plus souvent liée à une évaluation initiale plutôt qu'à une évaluation formative. L'évaluation formative sillonne parfois les étapes d'un cheminement linéaire d'apprentissage.

La place de l'évaluation dans la planification, dans la réalisation et dans la régulation des situations pédagogiques varie aussi en fonction des attitudes de fond de l'enseignant envers l'évaluation. Plus l'évaluation est vécue comme problématique, plus elle est réduite à sa fonction certificative.

Causes du succès et de l'échec

Le succès et l'échec est attribué par les enseignants principalement à l'engagement de l'élève dans ses études, à ses capacités, au soutien de sa famille. Les variables didactiques ne sont pas évoquées. Des variables affectives touchant à la création d'un climat sécurisant de la part de l'enseignant ou à l'encouragement à persévérer malgré les difficultés sont parfois intercalées dans un discours reconduisant pourtant à l'élève la responsabilité finale des résultats.

Discussion

Les résultats - qui se veulent révélateurs de criticités ou d'opportunités de valorisation de pratiques congruentes plutôt que descripteurs de tendances générales – montrent tout d'abord l'écart conceptuel qui sépare, dans les représentations des enseignants, la *didactique* (interprétée comme organisation des processus d'enseignement-apprentissage) des *pratiques d'évaluation* (considérées comme indépendantes des démarches de cours). On constate, par exemple, le contraste entre la richesse supposée des situations d'enseignement-apprentissage, qui font appel notamment à la différenciation pédagogique, à des démarches d'inspiration socioconstructiviste, à des modèles d'apprentissage coopératif, et la pauvreté des variables d'évaluation utilisées.

La marginalité de l'évaluation dans la cartographie des pratiques pédagogiques est, par ailleurs, un caractère significatif et récurrent dans un grand nombre de recherches pédagogiques, qui renvoie à des motivations s'insérant dans divers cadres de référence. Ainsi, sous l'angle de la relation éducative (Rey, 1999 ; Meyor, 2002), des enjeux de pouvoir interviennent, par exemple dans l'affirmation des prérogatives de l'enseignant par rapport aux critères, aux moyens et aux décisions, renforçant chez l'élève un sentiment d'arbitraire, source potentielle de conflit. Certains enseignants considèrent, au fait, ce pouvoir redoutable en tant que générateur de conflit: minimiser, éloigner voire supprimer l'évaluation peut créer l'illusion de la cessation des conflits, tandis que l'option de gérer et estomper ces troubles de relation par voie didactique n'est que très rarement envisagée. Du côté des compétences professionnelles de l'enseignant (Perrenoud, 1999), des marques d'insécurité vis-à-vis d'un domaine que l'on sent mouvant et controverse (Grange Sergi, 2001) vont de pair avec des attitudes de méfiance envers des instruments d'évaluation validés, jugés froids, distants, sans âme et donc inadéquats à leur fonction éducative (Vertecchi, 1993), ce qui débouche sur la banalisation de l'évaluation ou sur la surestimation des savoirs d'expérience sédimentés dans une pratique intuitive, non réfléchie, en dépit d'une gestion systématique compétente et formative de la progression des apprentissages. Du point de vue épistémologique, le discours sur l'évaluation tend à se distinguer du discours sur les actions d'enseignement-apprentissage là où les cadres conceptuels des enseignants pour leurs pratiques d'évaluation font état d'une certaine atomisation qui conduit chaque instituteur à intégrer implicitement des stratégies relevant de différents paradigmes, sans nécessairement les soumettre à un contrôle de pertinence, d'efficacité et de consistance par rapport aux processus didactiques qu'il promeut (Domenici & Moretti, 2006).

Cette mise en marge de l'évaluation se conjugue avec la manifestation de besoins de formation autour d'un thème difficile qui « pose problème » et dont on postule l'irréductible complexité; des sentiments d'impuissance peuvent d'ailleurs l'emporter sur la motivation à développer des compétences en matière d'évaluation et à remettre en question, si nécessaire de façon radicale, son propre fonctionnement professionnel dans la vie de classe. Les attitudes recueillies par les interviews révèlent une

considérable ambivalence face à une fonction pédagogique perçue comme importante, essentielle mais non parfaitement maîtrisée, en définitive insaisissable, et potentiellement conflictuelle.

Les fonctions d'évaluation, bien que parfois connues, pertinemment citées et idéalement évoquées comme nécessaires, ne sont pas toujours exploitées et souvent ne deviennent opérationnelles que dans le cas de l'évaluation sommative. L'accent sur les résultats, sur les bilans, rapprochent l'évaluation des jugements et s'accompagnent à une certaine déresponsabilisation de l'enseignant face aux succès ou aux échecs des élèves (Gosling, 1992). Il émerge une conception *statique* de l'évaluation comme constat, en relation avec une représentation de la profession enseignante où la co-responsabilité dans les processus d'enseignement-apprentissage est fortement déséquilibrée du côté de l'élève et où l'ancien clivage entre l'acte d'enseigner et l'acte d'apprendre reste latent, comme référence dans l'interprétation des situations et dans la prise de décisions.

Les réponses sur les causes du succès et de l'échec scolaire - qui renvoient prioritairement à des caractéristiques de l'élève, de préférence internes, généralement non modifiables ou modifiables à l'initiative de l'élève lui-même - confirment une certaine distance de l'enseignant en posture d'évaluateur, non seulement vis-à-vis de l'élève mais aussi face aux implications pédagogiques du processus d'évaluation. L'évaluation n'est pas perçue, dans ce cas, comme une fonction s'articulant au rôle de l'enseignant mais au contraire comme un rôle distinct qui peut arriver jusqu'à paraître antagoniste à celui de l'enseignant. Ces représentations séparatrices sont compatibles avec une idéologie de l'égalité des chances qui ne contemple qu'une prise en charge partielle de la part de l'enseignant du droit à l'apprentissage des élèves (Grange Sergi, 2006). Ainsi, l'évaluation réduite à sa fonction sommative, se configure-t-elle comme un prolongement linéaire des cours, figeant les choix didactiques dans l'alternance classique de processus de transmission et de contrôle. Dans cette perspective, on arrive également à composer la contradiction entre, d'une part, l'option pour une didactique qui, dans certains cas, n'est pas purement transmissive mais fait recours épisodiquement à des situations-problèmes, à des formats de leçon variés, à des stratégies tirées des pédagogies actives, constructives, non directives et, d'autre part, la pratique d'une évaluation réductrice de censeur. C'est en fait par le biais de l'élision que l'on dissimule le paradoxe: on décide, tout simplement, de ne rien évaluer jusqu'à la fin des processus et on adapte les objets d'évaluation aux moyens que l'on se donne. Cet implicite fallacieux, fondé sur la séparation des fonctions d'évaluation en relation aux processus d'apprentissage, permet de comprendre la coexistence de formes évoluées d'action pédagogique avec des formes pauvres d'évaluation. L'évaluation qui est repoussée *à la fin* (à la fin de la séquence, à la fin du trimestre, à la fin des cours de formation, souvent à la fin des manuels de didactique …) avec un statut *à part*, sanctionnant un résultat, recèle des pièges. Par exemple, l'illusion du pilotage des processus d'enseignement-apprentissage, par le recours à des méthodologies riches ou novatrices, en régulant l'activité sans pour autant réguler les apprentissages : les réponses aux questionnaire mettent en évidence que plusieurs enseignants, pendant la planification des cours, ne semblent pas chercher une cohérence entre les variables d'évaluation et les autres variables pédagogiques. Ensuite, des problèmes d'efficacité descendent des limites des actions

de remédiation *a posteriori* : le réinvestissement des données recueillies à l'occasion de pratiques d'évaluation sommative est faible et se traduit dans des démarches de compensation par la répétition de segments didactiques, souvent vécues comme des arrêts ou comme des « retours en arrière » un peu frustrants en soi, avec des bénéfices généralement modestes. Par ailleurs, les enseignants citent souvent l'observation des élèves comme pratique professionnelle, ils déclarent d'y faire recours « systématiquement » de façon spontanée ou, plus rarement, instrumentée, mais l'utilisation de ces données d'observation se borne à la formation d'impressions généralement globales ayant un lien assez précaire et aléatoire avec les décisions pédagogiques en situation.

Une autre conséquence, à long terme, de ce cloisonnement entre évaluation et apprentissage affecte notamment la motivation des élèves, car l'usage exclusif de l'évaluation sommative, avec son accent sur la performance et les résultats, est souvent associé à des phénomènes de décrochage ou d'engagement minimal et instrumental orienté à la reconnaissance externe des acquis plutôt qu'à la construction de nouveaux savoirs (Dweck, 1989). Les enseignants témoignent de l'authenticité de ce risque quand ils stigmatisent le comportement des élèves qui s'engagent dans une tâche à condition qu'elle soit notée, sans toutefois s'apercevoir qu'il s'agit d'une conduite apprise, à partir d'une expérience scolaire où la seule fonction d'évaluation exercée est justement l'évaluation sommative.

A l'opposé, une conception intégrant l'évaluation dans les processus d'enseignement-apprentissage émerge dans les discours sur l'évaluation formative et sur l'enchaînement de l'évaluation formative et de la différenciation pédagogique, mais elle présente des limites dans la nature de cette intégration. Dans l'échantillon considéré dans cette enquête, on trouve des évaluations formatives au pluriel, des exemples présentés par les enseignants eux-mêmes comme isolés, pulvérisés : des moments de travail de classe qu'on ne sait pas raccorder dans une matrice commune qui leur donne un sens au sein des démarches de construction des apprentissages. Il s'agit d'une évaluation formative un peu naïve jouant sur la parcellisation des étapes et sur les rémédiations précoces, s'inscrivant en général dans des modèles d'apprentissage mécanicistes. Il y a chez beaucoup d'enseignants la conviction que l'évaluation formative se distingue de l'évaluation sommative par le moment où elle intervient (*durant* le processus), sans retenir par contre la différence de fonction qui est à la base de cette différence temporelle. La différenciation, souvent évoquée en termes d'individualisation de l'enseignement en raison d'exigences particulières de l'élève (rythmes d'apprentissage, aptitudes reconnues, styles d'apprentissage, acquis préalables etc.), est fréquemment subordonnée à une évaluation diagnostique initiale et globale, par exemple à l'occasion de tests passés en début d'année portant sur un ensemble d'habiletés et de connaissances. Or, quand l'évaluation diagnostique est non spécifique, elle risque de devenir prédictive à long terme, et partant de transformer la différenciation dans une cristallisation des difficultés ou des acquis. Une évaluation diagnostique spécifique et contingente, au contraire, renforce le caractère provisoire des résultats d'évaluation et favorise chez les enseignants des attitudes évolutives face aux possibilités de progression des élèves dans leurs apprentissages. D'après les entretiens, on dégage un souci d'application des fonctions formative et diagnostique de l'évaluation qui se résout, dans la plupart des cas, dans une attitude plus pronostique que proactive : on comprend mieux la situation, mais le passage à une

planification conséquente favorisant le changement positif et débouchant sur l'atteinte des objectifs fixés ne va pas de soi.

Les fonctions d'orientation apparaissent, dans les cas examinés ici, seulement en relation avec l'évaluation sommative et sur des modèles déterministes: la projection dans l'avenir des résultats scolaires dans une discipline marque la stabilité supposée des évaluations des enseignants telle qu'elle est mise en évidence, entre autres, dans les études sur les effets Mathieu et Posthumus (Walberg & Tsaï, 1983) . La sollicitation du système métacognitif de l'élève et des perceptions de *self-efficacy* par des démarches d'évaluation formative et par les régulations qui en descendent stimulent, par contre, le développement de compétences d'orientation en termes de connaissance de soi, de sa propre façon d'apprendre, de ses stratégies ainsi que de son rapport aux savoirs, de façon contextualisée, situation par situation, par un effet démultiplicateur des ressources repérées chaque fois comme disponibles où en voie d'acquisition.

Quant aux acteurs de l'évaluation, l'autoévaluation des élèves est évoquée comme pratique impromptue et épisodique, sans contextualisation, ni finalisation à des processus d'apprentissage de niveau supérieur (comme, par exemple, métacognition ou métadécision). Parfois elle est proposée en situation d'évaluation sommative individuelle ou réciproque, sous forme ludique, apparemment sans avoir conscience des risques de confusion de rôles ni des conséquences au niveau de relations entre pairs (Allal, 1999).

La participation des parents d'élèves est confinée aux situations extrêmes de risque de doublement; dans les autres cas on a affaire à une communication unidirectionnelle de l'enseignant à la famille par la notation des travaux des élèves sur les cahiers, les carnets et les bulletins, et à l'occasion des entretiens trimestriels. Les contenus de ces communications sont souvent exprimés en termes de progression formative, mais on ne semble pas chercher de dialogue réel ou de partage des critères adoptés.

La cogestion de l'évaluation entre collègues présente, elle, un cadre plus souple et prospère de confrontation d'instruments et de co-responsabilité dans les décisions, inscrit dans une tradition de collégialité enseignante qui caractérise l'école primaire italienne depuis les années '80. Elle reproduit néanmoins les mécanismes de facilitation ou d'entrave liés à la condition d'interdépendance des enseignants faisant partie d'un conseil de classe ou d'un groupe de travail, une entente de fond au niveau personnel et de « la vision des choses à l'école » étant réputée indispensable pour mener à bien une concertation des stratégies et des instruments d'évaluation.

Le problème de promouvoir une culture de l'évaluation chez l'enseignant dans l'amélioration de la qualité pédagogique, et dans la contextualisation de toute hypothèse de changement évolutif des pratiques d'évaluation à l'école apparaît incontournable. Une évaluation à saisir dans toute sa complexité systémique, qui dynamise et articule de façon cohérente les interactions d'objets multiples (qu'est-ce qu'on évalue ?), de finalités multiples (pourquoi on évalue ?), de temps multiples (quand on évalue ?), de stratégies multiples (comment on évalue ?) de moyens multiples (par quels instruments on évalue ?) d'acteurs multiples (qui évalue ? qui est

évalué ?) de contextes multiples (où, dans quelles situations on évalue ?) est bien l'enjeu d'une prise en charge effective de cette fonction de la part des enseignants.

Il y a des aspects de l'évaluation qui, par des raisons d'expérience, de nécessité du quotidien ou de pression externe, ont pénétré plus que d'autres dans l'imaginaire et dans l'attention des enseignants ; les prendre en compte s'avère utile et nécessaire comme point d'ancrage des fils qui tissent la toile du réseau plus vaste d'une évaluation bien intégrée aux autres composantes des processus d'enseignement-apprentissage, d'où l'intérêt de mieux les explorer et les connaître. C'est donc d'abord une question d'intelligibilité, mais ensuite c'est une question d'investissement en formation avec la détermination d'inscrire toute proposition de changement - si limitée soit-elle, comme l'adoption d'un nouvel instrument de certification - dans une perspective réflexive, touchant à la construction progressive d'une culture de l'évaluation orientée à la réalisation du droit à l'apprentissage pour tous les élèves.

Références

Allal, L. (1999). Impliquer l'élève dans le processus d'évaluation: Promesses et pièges de l'autoévaluation. In C. Depover & B. Noel (Ed.), *L'évaluation des compétences et des processus cognitifs* , 35-56. Bruxelles: De Boeck.

Becchi, E. & Vertecchi, B. (1986). *Manuale critico della sperimentazione e della ricerca educativa*.Milano: Franco Angeli

Domenici, G. & Moretti, G. (2006). *Il Portfolio deell'allievo. Autonomia, equità e regolazione dei processi formativi*. Roma: Anicia

Dweck C.S. (1989). Motivation. In A. Lesgold & R.Glaser (Eds), Foundations for a psychology of education , 87- 137, NJ: Lawrence Erlbaum Associates

Farr, R. & Tone, B. (1998). *Portfolio and performance assessment: helping students evaluate their progress as readers and writers*. Orlando: Harcourt Brace College Publishers.

Gosling, P. (1992). *Qui est responsable de l'échec scolaire?* Paris: PUF.

Grange Sergi, T. (2001). *Verso un'identità professionale dell'insegnante, tra sapere e immaginario.* Aosta: Le Château.

Grange Sergi, T. (2006). Diagnosi didattica e modelli attribuzionali. Il successo scolastico nelle rappresentazioni degli insegnanti. In G. Domenici (Ed). *La ricerca didattica per la formazione degli insegnanti.* Roma: Monolite, in corso di stampa *(en cours d'édition)*

Huberman, M. & Miles, M.(1984). *Innovation up close. How school improvement works*. New York: Plenum Press

Meyor, C. (2002). *L'affectivité en éducation.* Bruxelles: De Boeck

Pellerey, M. (1998). *L'agire educativo. La pratica pedagogica tra modernità e postmodernità*. Roma: LAS.

Perrenoud, Ph. (1998). *L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages*. Bruxelles: De Boeck

Perrenoud, Ph. (1999). *Dix nouvelles compétences pour enseigner. Invitation au voyage*. Paris: ESF.

Rey, B. (1999). *Les relations dans la classe.* Paris: ESF.

Varisco, B. (2004). *Portfolio, valutare gli apprendimenti e le competenze.* Roma: Carocci

Vertecchi, B. (1993). *Decisione didattica e valutazione.* Firenze: La Nuova Italia.

Vertecchi, B. (2003). *Un'altra idea di scuola.* Roma: Anicia.

Walberg H.J. & Tsaï S.L. (1983). Mathew effects in education, *American Educational Research Journal*, 20 (3), 359-373.

EVALUER LES COMPETENCES ? LES RESSOURCES ET LES CONTRAINTES DU PORTFOLIO NUMERIQUE.

Baillat Gilles, Connan Pierre-Yves, IUFM Champagne Ardenne

Résumé :
La formation des enseignants est aujourd'hui entrée dans de nombreux pays dans une nouvelle phase marquée en particulier parce que l'on désigne souvent comme « l'approche par compétences ». Cette approche mobilise les instituts de formation (Universités, IUFM, Hautes Ecoles, écoles normales) et leurs acteurs en provoquant cependant toute une série de tensions liées aux mutation du paradigme de formation, mais aussi à la diffusion progressive des nouvelles technologies d'information et de communication. Parmi ces tensions, on note en particulier celles qui sont relatives aux besoins et aux difficultés de l'individualisation. Ces difficultés et besoins sont parfois (comme en France) renforcés par la diversification croissante des types de publics accueillis en formation aux métiers de l'enseignement. Les IUFM sont ainsi aujourd'hui susceptibles d'accueillir non seulement des étudiants passant directement de l'Université à l'institut de formation, mais aussi des étudiants récemment diplômés mais possédant déjà une première expérience professionnelle significative d'enseignement et, de plus en plus, des personnes en reconversion professionnelles complète. Une autre série de tensions est remarquable : celles relatives à l'évaluation des compétences professionnelles, dans des lieux de formation et avec des acteurs plus habitués à l'évaluation de connaissances académiques.

Présentation

Les nouvelles technologies se présentent parfois comme un ensemble de nouvelles ressources pour traiter ces difficultés et y apporter des réponses pertinentes, malgré les limites qui sont souvent soulignées. Parmi les évolutions récentes, la diffusion du portfolio numérique semble offrit des perspectives nouvelles, illustrées par exemple par la récente « **Première Conférence Internationale Francophone sur l'e-portfolio** tenue récemment à Quebec (avril 2006). On a pu notamment relevé lors de cette conférence, l'essor des initiatives en milieu scolaire dans certains pays (Karsenti et alii, 2006), mais aussi et surtout des initiatives développées en milieu universitaire (Cerisier,JF, Branciard, L., Casteignau, G., 2006) . Si la problématique du portfolio n'est pas une problématique vraiment nouvelle aujourd'hui il est clair que l'utilisation des nouveaux moyens technologiques font du portfolio numérique une opportunité originale, au regard de la question des compétences. La variété des conceptions (Belair 2005) du portfolio numérique impose cependant que l'on précise ici le point de vue retenu. Le port folio numérique peut en effet être appréhendé comme un « dossier d'apprentissage » permettant d'obtenir une vision globale des travaux réalisés par les étudiants ; il peut aussi être perçu comme « portfolio professionnel » en offrant une somme des travaux réalisés mais aussi une distanciation prise vis à vis de ces travaux ; il peut aussi, véritable « portefeuille de compétences », se présenter comme

un curriculum vitae avec l'exposé des preuves de compétences, le tout dans un but d'employabilité ou de reconnaissance qualifiante (VAE) ; il peut enfin se présenter avant tout comme un « processfolio » dans lequel le processus d'élaboration est aussi important que l'objet lui-même, ce qui fait de l'outil avant tout un portfolio de formation. C'est bien dan cette dernière perspective que nous nous situons, même si les trois autres conceptions peuvent aussi être considérées comme des dimensions qui ne peuvent jamais être complètement absentes.

La présente communication se propose de rendre compte d'une analyse des résultats d'une expérimentation menée dans un cadre qui vise à donner une réponse cohérente, en terme de formation et d'évaluation, aux nouveaux défis de la formation des enseignants français aux usages « TICE », défis qui sont actuellement résumés par la promotion d'une nouvelle certification de ces usages (le « Certificat Informatique et Internet » ou C2I)(Baillat,G., Trestini, M. 2006). L'expérimentation analysée, qui porte sur plus de 450 étudiants de l'IUFM de Champagne Ardenne, était inscrite dans un cadre relevant des logiques du portfolio numérique : le carnet de bord informatique (CBI). Ce carnet de bord informatique se présente comme un nouvel outil d'interlocution entre d'une part, des étudiants ayant à mener à bien leur formation à des usages identifiés par un référentiel officiel, et susceptibles de réaliser cette formation dans des cadres formels ou informels et d'autre part des enseignants et formateurs susceptibles de participer à la formation mais aussi de valider les éléments participant à la délivrance du certificat. Le corpus retenu dans l'analyse concerne des étudiants à l'enseignement pour le premier degré comme pour le second degré. Les premiers résultats illustrent clairement le contraste entre ces deux populations quant aux pratiques pouvant donner lieu à des validations d'acquis de l'expérience professionnelle, la différence actuelle des modalités de stage expliquant largement cette situation. L'utilisation tout au long de l'année 2005-2006 du CBI par ces étudiants a donnée lieu à un ensemble de flux de messages et d'échanges qui ont fait l'objet d'une analyse quantitative. Le contenu des échanges a fait lui l'objet d'une analyse du contenu des différentes interventions (étudiants et enseignants participant à l'expérimentation

Références.

Baillat, G. Vincent, J., Portfolio numérique et formation des enseignants, 1ère Conférence Internationale Francophone sur le e-portfolio. Pour une technologie au service de la valorisation de la personne, Quebec, 10-12 avril 2006.

Branciard, L. Barreau L., Un portfolio numérique pour le suivi, l'évaluation et la certification des professeurs de l'enseignement agricole public français, 1ère Conférence Internationale….

Casteignau G., le concept de portfolio virtuel dans la société du savoir, 1ère Conférence Internationale….

Cerisier, J-F., Utilisation d'un portfolio numérique dans l'enseignement universitaire pour l'évaluation de stages professionnels, 1ère Conférence Internationale….

Karsenti,T., Ledoux M., Goyer S., Edu-portfolio.org : le portfolio numérique malléable pour maximiser l'intégration des TIC en milieu scolaire

Paquay L.,(sous la direction de), L'évaluation des enseignants. Tensions et enjeux, L'Harmattan, 2004, 330p

Tardif, J., Intégrer les nouvelles technologies de l'information : quel cadre pédagogique ?, ESF éditeur, 2000

Trestini, M. La validation numérique des compétences du C2I enseignant dans le cadre d'une démarche de portfolio numérique, 1ère Conférence Internationale Francophone sur le e-portfolio. Pour une technologie au service de la valorisation de la personne, Quebec, 10-12 avril 2006.

LE PORTFOLIO AU SERVICE DES COMPÉTENCES TRANSVERSALES

Hervé Bernard, Université de Liège
Joëlle Vlassis, Université du Luxembourg

Mots clés : portfolio, situations problèmes, métacognition.

Résumé:
En Communauté française de Belgique, la réforme des socles de compétences a entraîné des implications pédagogiques considérables, et par delà les intentions, il convient de souligner qu'actuellement les enseignants manquent de moyens. En particulier, les compétences transversales posent problèmes. Celles-ci consistent en "démarches mentales, manières d'apprendre et attitudes relationnelles" (Socles de Compétences, 1999, p.4) qui transcendent les disciplines. Nous avons élaboré une recherche visant à concevoir, expérimenter et évaluer une séquence d'enseignement destinée à développer, chez les élèves du premier degré du secondaire des compétences transversales à travers quatre disciplines : le français, les sciences, l'histoire et la géographie. Pour chaque discipline, une séquence respectant la démarche par situations problèmes a été construite. Cette approche implique en effet le recours à des compétences transversales. En outre, des moments de réflexion métacognitive ont été prévus en vue d'amener les élèves à identifier puis à utiliser les stratégies efficaces afin de mettre en œuvre les compétences transversales. Enfin, ces options pédagogiques sont soutenues et intégrées par la réalisation d'un portfolio interdisciplinaire. Ce dernier outil, qui permet lui aussi de développer les compétences transversales, matérialise le caractère interdisciplinaire du présent dispositif.

I. Introduction

Dans une société où le savoir et les connaissances évoluent très rapidement, l'école ne peut prétendre être en mesure de transmettre tous les savoirs nécessaires à la vie future de l'élève. Ce dernier devra en effet continuer à s'auto-former afin d'actualiser ses connaissances en vue de faire face aux exigences de la société en matière de formation. En outre, la société demande également des individus capables d'adaptation face à des situations non routinières. Il est dès lors primordial d'amener nos élèves à s'approprier des compétences d'auto-formation et de résolution de problèmes. En Communauté française, le Décret Missions (Conseil de la Communauté française de Belgique, 1997) préconise le développement de telles compétences reprises généralement sous l'appellation de "compétences transversales". Celles-ci sont définies comme des "attitudes, démarches mentales et démarches méthodologiques communes aux différentes disciplines à acquérir et à mettre en œuvre au cours de l'élaboration des différents savoirs et savoirs-faire ; leur maîtrise vise à une autonomie croissante d'apprentissage des élèves" (Décret Missions, 1997). Ces compétences définies, encore faut-il préciser comment les développer et les évaluer concrètement. C'est l'objet de la présente recherche. Celle-ci ambitionne de

développer des compétences transversales au travers de quatre disciplines: le français, l'histoire, la géographie et les sciences. Trois options pédagogiques fondamentales sont à la base du présent dispositif : la démarche par situations problèmes, la métacognition et le portfolio. Dans cet article, nous commencerons par développer ces trois options pédagogiques ainsi que leur articulation au sein du dispositif, nous présenterons ensuite le dispositif construit sur base de ces options pédagogiques et, enfin, nous présenterons les premiers résultats récoltés au terme de la deuxième année de recherche. Des résultats plus précis seront disponibles au terme de la troisième année de recherche.

II. Présentation du cadre théorique

Trois options pédagogiques sont à la base du dispositif : la démarche par situations problèmes, le portfolio et la métacognition. Nous les développons ci-dessous.

A. La démarche de résolution de problèmes

La démarche par situations problèmes est actuellement préconisée en Communauté française de Belgique dans les programmes de français, histoire, géographie et sciences (sous un vocable parfois différent tels "énigmes, enquêtes, problèmes"). Il s'agit de mettre l'élève dans la peau d'un scientifique faisant face à un problème. La tâche de ce dernier est alors, de façon quelque peu simplifiée, d'émettre des hypothèses, de chercher la (ou les) solution(s) et de la (ou les) communiquer. Quatre caractéristiques nous paraissent fondamentales pour la création d'une véritable situation problème (voir à ce sujet De Vecchi, 2005 ; Fabre, 1999 ; Perrenoud, 1995).

- Tout d'abord, la situation doit avoir du sens pour l'élève et donc, soit l'interpeller dans ses représentations spontanées, soit être contextualisée afin de montrer comment ces savoirs sont mis en jeu dans la vie réelle.
- Ensuite, la situation doit prendre en compte les représentations des élèves et amener à leur modification.
- La situation problème correspond à une situation complexe laissant aux élèves la liberté de recourir à différentes stratégies en vue de sa résolution. L'enseignant se positionne en tant que personne ressource qui aide, conseille l'élève. En outre, la situation doit également mettre en jeu plusieurs variables afin que plusieurs hypothèses de résolutions soient possibles.
- Enfin, une situation problème débouche sur un savoir au sens large du terme. Celui-ci comprend deux dimensions : des savoirs disciplinaires prescrits par les programmes d'une part, et des compétences transversales d'autre part. Rappelons que l'on ne peut exercer des compétences à vide, sans contenu.

Dans chacune des quatre disciplines, nous avons construit une situation problème respectant ces principes. L'intérêt de recourir à de telles situations est qu'elles mettent en jeu les quatre compétences transversales dont nous visons le développement.

B. Le portfolio

Il existe plusieurs types de portfolios poursuivant divers objectifs (Goupil, 1998). Dans le cadre de notre recherche, nous avons choisi d'exploiter le portfolio du type "dossier progressif d'apprentissage" qui répond au mieux à nos objectifs. Selon Paulson, Paulson et Meyer (1991, in Goupil, 1998), ce type de portfolio :

- consiste en un assemblage finalisé des travaux de l'élève qui démontre ses efforts, ses progrès et ses acquisitions dans un ou plusieurs domaines,
- et implique la participation de l'élève à la sélection des contenus, à la définition des critères de sélection et d'appréciation des travaux, ainsi que des manifestations de réflexion personnelle.

Parmi les principes pédagogiques liés au portfolio, trois nous ont paru déterminants à retenir dans le cadre de notre projet : l'évaluation de processus (démarche en cours) autant que des produits, une évaluation formative basée sur des concertations multiples et corollairement, le développement de comportements métacognitifs.

B.1 L'évaluation des processus d'apprentissage

L'idée d'utiliser le portfolio dans le contexte scolaire est née afin de proposer une alternative à l'évaluation standardisée de compétences. Cette évaluation était focalisée sur des produits, sur des performances attestant de compétences en vue de leur certification. Or, Schillings et Vanhulle (2001) rappellent qu'une compétence ne peut que se construire dans le temps et à travers une succession de travaux adéquats. Celle-ci se construit par le biais de l'acquisition de réseaux de connaissances sur des contenus et des manières de faire, tout en intégrant un travail sur des attitudes.

Le dossier d'apprentissage qui garde la trace des progrès réalisés permet une évaluation des différentes facettes des compétences en train de se construire. Ainsi, selon Schillings et Vanhulle (2001), le portfolio sert une évaluation authentique dans la mesure où il recueille diverses indications à partir des apprentissages de l'élève à travers un ensemble d'activités variées. Au lieu de s'appuyer sur des tests ponctuels, ou sur des échantillons réduits de connaissances à un moment donné, cette évaluation consiste à recueillir des informations sur une période prolongée. Cette évaluation, profondément ancrée dans les apprentissages sera non plus exclusivement centrée sur les produits, résultats de l'apprentissage, mais également sur les processus c'est-à-dire les démarches, la réflexion et le raisonnement qui accompagnent cet apprentissage.

Schillings et Vanhulle (2001) rappellent que le but d'une telle forme d'évaluation "enchâssée" dans les apprentissages comme le dit Tardif, est de favoriser l'implication et la motivation des élèves. Celle-ci est formative. Elle apprend à l'élève, via la co-évaluation et l'auto-évaluation (voir point suivant), où sont ses points forts, ses lacunes, et comment y remédier. L'élève peut ainsi se définir des projets de progression dans ses apprentissages.

B.2 Une évaluation formative basée sur une concertation multiple

Un des points forts du portfolio consiste à baser l'évaluation et les apprentissages sur des interactions sociales. En effet, Allal, Cochet-Kaeser, Bonaiti-Dugerdil, et Wegmuller (1998) rappelle que trois formes d'engagement de l'élève dans le processus d'évaluation sont mises en place : l'auto-évaluation, la co-évaluation élève-élève(s), la co-évaluation maître-élève. L'ensemble de ces exploitations du portfolio contribue à la fonction formative de l'évaluation en tant moyen de régulation.

À ce sujet, ce sont les échanges avec l'adulte et/ou ses pairs qui permettent à l'apprenant d'accéder à des opérations mentales élaborées selon un processus que Vygotsky (1985) appelle "internalisation". Pour Vygotsky (1985), ce processus est à la base des apprentissages : le sujet "transforme" son activité sociale et collective en activité individuelle. Par exemple, lorsque, avec l'aide de l'adulte, l'élève apprend une démarche, il va intérioriser le processus qui a été mis en place et tenter de le reproduire mentalement et individuellement lorsqu'il sera confronté à un problème qui nécessitera une démarche similaire. Au fil des interactions avec l'adulte, l'élève affine son efficacité. Dans le cas du portfolio, c'est au travers des moments de co-évaluation que celui-ci perfectionnera son mode d'autoévaluation. Mais qu'entend-on précisément par co-évaluation et auto-évaluation? Quels sont leurs intérêts? Et en quoi l'auto-évaluation peut-elle découler de la co-évaluation?

La co-évaluation

La co-évaluation consiste en "un processus suivant lequel l'enseignant et l'élève confrontent leur point de vue sur une production de l'élève, ou suivant lequel, les élèves entre eux donnent leur point de vue sur les productions des uns et des autres" (Dore, Michaud, Muchararugi, 2002, p.17).

Pourquoi parler de *co-évaluation maître-élève* et pas d'évaluation tout court? Parce que dans la perspective du portfolio, il s'agit bien d'une évaluation conjointe où l'enseignant évalue le travail ou les progrès d'un élève, mais où l'élève donne également son avis sur ses apprentissages. Ce mode d'évaluation s'effectue le plus souvent au cours d'entretiens (voir point suivant). Cette évaluation n'exclut en aucune manière des moments d'évaluation plus traditionnelle du travail de l'élève par le maître.

L'avantage de la *co-évaluation élève-élève,* quand elle est bien menée, est de permettre aux élèves, de se référer aux idées des autres ou d'échanger avec les pairs sur les stratégies utilisées pour résoudre un problème. "La co-évaluation peut être une occasion de parler de ses réussites et d'aider des camarades à corriger leurs erreurs. Le fait d'expliquer à l'autre demande de bien structurer et articuler ses propres connaissances. L'élève apprend aussi à développer son esprit critique et à l'utiliser de façon constructive" (Dore *et al.*, 2002, p.17).

L'auto-évaluation

Selon Dore, Michaud, Mukararugi (2002), l'auto-évaluation est un moyen privilégié d'appropriation du savoir par l'élève. "Par l'auto-évaluation, l'élève acquiert un mode de pensée, la capacité d'exercer un contrôle sur ses actions et de prendre des

décisions quant à ses apprentissages" (Dore *et al.*, 2002, p. 12). Ces acquisitions seront d'autant plus intériorisées qu'elles auront été accompagnées par des co-évaluations avec le maître et/ou des pairs.

"Il s'agit d'un apprentissage inestimable", car, comme le souligne Greene (1997), "une fois que l'élève quittera les bancs de l'école, l'enseignant ne sera plus là pour lui dire si son travail est de qualité ou non" (p. 16).

Les pratiques de co-évaluation et d'auto-évaluation amènent l'élève à participer activement à l'évaluation de ses apprentissages. Dore, Michaud, Mukuragagi (2002) soulignent que cet engagement de l'élève agit positivement sur la motivation. Etant responsable de ses apprentissages, l'élève se sent moins à la merci des notes reçues, il n'est plus simplement objet d'évaluation. En outre, toujours selon ces auteurs, le portfolio agit sur l'estime de soi. L'élève se rend compte de son droit à l'erreur. Celle-ci fait partie de l'apprentissage, l'élève comprend qu'il est en cheminement, voit ses progrès et se fixe de nouveaux défis. L'évaluation n'est plus un mystère pour lui.

B.3 Pratiques visant à développer chez l'élève le sentiment que le portfolio lui appartient

Les différents ouvrages traitant du portfolio insistent sur le fait que l'élève doit avoir le sentiment que le portfolio lui appartient véritablement. Dans cette perspective, l'élève doit participer activement à la constitution de son portfolio et à la sélection des documents. À ce sujet, Schillings et Vanhulle (2001) soulignent que le portfolio doit, pour son auteur, consister en une collection de travaux rassemblés dans la logique d'une intention précise, et sur lesquels les élèves sont amenés à réfléchir. Dans notre recherche, le portfolio devra donc contenir tout ce qui, selon l'enseignant et l'élève, semble propre à favoriser le développement des compétences transversales retenues et précisées dans les fiches-processus.

Farr et Tone (1998) insistent sur ce sentiment d'appartenance dans la mesure où celui-ci contribue à inculquer aux élèves l'habitude de s'auto-évaluer. En effet, selon ces auteurs, "les élèves qui se sentent tenus de respecter une procédure préétablie régissant la sélection des documents à insérer dans le portfolio n'ont en général pas l'impression qu'il leur appartient véritablement. De plus, cette procédure dispense l'élève d'une bonne partie du travail de réflexion et d'évaluation que le portfolio requiert [...] À l'opposé lorsque les élèves choisissent les documents à conserver, qu'ils les classent et les reclassent, en déterminant pour ce faire des catégories souvent très générales, ils s'initient à l'auto-évaluation.

Dans notre dispositif, trois démarches ont été développées afin de développer chez les élèves cette conviction: l'ajout de documents personnels, le tri, et la table des matières.

L'ajout de documents personnels

Schillings et Vanhulle (2001) insistent sur l'idée de ne pas standardiser le portfolio. En effet, il vaut mieux éviter une sélection trop dirigée du portfolio avec une série de

rubriques bien spécifiées à l'avance. Le risque d'une telle option serait de restreindre trop hâtivement les apports des élèves sur ce qu'ils jugent, eux, significatifs pour apprendre. Idéalement, selon Farr et Tone (1998), l'élève devrait être encouragé à mettre dans son portfolio à peu près tout ce qu'il veut et l'enseignant devrait y ajouter ce qu'il juge utile à son évaluation et susceptible d'amener l'élève à s'auto-évaluer. Il convient de ne pas imposer des critères de sélection trop restrictifs. Cette stratégie contribue à l'efficacité du portfolio. En effet, celui-ci doit conserver ce caractère ouvert afin que l'élève ait le sentiment qu'il lui appartient, sinon sa capacité de favoriser l'auto-évaluation serait compromise.

Concrètement, le portfolio qui est exploité dans nos séquences est subdivisé en fonction des compétences poursuivies. Dans ces catégories, certains documents sont imposés (fiche-processus, ...). Mais à l'intérieur de celles-ci, les élèves sont libres d'organiser leurs documents et d'en ajouter d'autres. Afin d'augmenter l'aspect métacognitif de cette démarche, il est demandé aux élèves de compléter une (très) courte fiche pour justifier leur choix. Afin que cette démarche ne constitue un obstacle à l'ajout de document, cette fiche prend la forme d'un QCM.

Le tri et la table des matières

Dans la même perspective, nous pensons qu'il est important de prévoir au cours du dispositif des moments de tri des documents. En effet, après avoir inséré ceux qui sont imposés et ajouté ceux qui lui sont personnels, l'élève réorganisera son portfolio en fonction de ses propres critères. Ceux-ci peuvent être discutés avec des pairs ou avec l'enseignant. Pour Forget (2001), le tri, pierre angulaire du processus, conduit l'élève à réfléchir à ce que chacun des documents présents dans sa farde représente pour lui. Ce moment donne rétrospectivement tout son sens aux activités organisées puisqu'il invite l'enfant à construire une signification personnelle à chacun des documents. Lorsque les élèves ont trié leur portfolio, ils peuvent, s'ils le souhaitent, créer de nouvelles catégories. Farr et Tone (1998) proposent de les encourager à créer une table des matières. Certains élèves pensent alors à numéroter les pages de leur portfolio afin que leur table soit plus complète (titres des documents retenus et pagination). Enfin, ces auteurs suggèrent que l'enseignant invite ses élèves à classer ou reclasser le contenu de leur portfolio avant les entretiens maître-élève. C'est un moyen, parmi d'autres, d'éviter que la table des matières ne cristallise définitivement l'organisation de leur portfolio. Ce moyen leur permettra par la suite d'analyser et de classer d'une autre manière leurs documents, et de donner ainsi à leur collection un sens nouveau.

C. La métacognition

Le travail sur les aspects métacognitifs est au cœur du dispositif dans les situations problèmes et est renforcé par le portfolio. La métacognition, littéralement "connaissance sur la connaissance ", est généralement définie comme une activité mentale permettant à un individu de prendre du recul, de réfléchir sur ses apprentissages et d’en assumer ainsi le contrôle (Poissant, Poëlhuder & Falardeau, 1994 ; Tardiff, 1997). Il s’agit donc d’un processus de contrôle et de régulation conscient des actions réalisées par un individu. La métacognition agit sur trois facteurs qui nous paraissent essentiels :

- La métacognition augmente la motivation.

  En amenant l'élève à prendre conscience de ses stratégies de résolution et en l'amenant à identifier les stratégies les plus efficaces, l'élève en vient à percevoir que le résultat de ses actions est sous son contrôle et il développe par ailleurs une image plus positive de ses compétences (Viau, 1994 ; Presseau, 2004).
- La métacognition est liée à la réussite scolaire.

  Différents auteurs (Tardiff, 1997 ; Ghaye et Rahier, 1999) insistent sur le fait que les élèves qui réussissent bien à l'école mettent en œuvre des actions mentales de nature métacognitive.
- La métacognition favorise le transfert des apprentissages.

  Le travail sur les aspects cognitifs de la métacognition vise à doter l'élève de stratégies efficaces et économiques et à former ce dernier à identifier les stratégies à mettre en œuvre dans une situation particulière. Il s'agit ainsi d'autonomiser l'élève face aux tâches auxquelles il est confronté.

Le travail sur les aspects métacognitifs vise donc à doter les élèves de stratégies efficaces pour l'exercice des compétences transversales. Ainsi, au cours du dispositif, différents moments de réflexion et d'échanges de nature métacognitive sont prévus. Nous les présenterons plus précisément dans la suite de cet article.

Les trois principes pédagogiques présentés ci-dessus sont, on le voit, en interrelation. Le portfolio, qui poursuit des objectifs motivationnels et de développement des compétences transversales s'appuie largement sur des moments de réflexion de nature métacognitifs (diverses évaluations, construction des fiches processus). Si le portfolio réalise le lien entre les différentes disciplines, l'exercice de compétences transversales similaires dans ces disciplines est assuré par une démarche identique dans les quatre branches : la démarche par situations problèmes. Nous allons dans la partie suivante préciser davantage en quoi consistent ces séquences et comment elles s'articulent.

III. Méthodologie

Le dispositif que nous présentons au point suivant a été conçu au cours d'une recherche de terrain prévue pour une durée de trois ans. Cette recherche est basée sur la collaboration entre chercheurs et enseignants pour la création des séquences et des outils utilisés dans le cadre d'un portfolio. Quatre disciplines sont concernées par le dispositif. Il s'agit du français, des sciences, de la géographie et de l'histoire. Les analyses et les résultats présentés dans cet article reflètent les travaux menés au cours des deux premières années. Une première expérimentation a été réalisée dans trois classes, une classe de première secondaire et deux classes de deuxième secondaire. Outre la construction des séquences, le travail des chercheurs a consisté en des observations participatives au sein des classes. En vue d'évaluer l'impact de cette première expérimentation, les enseignants ont été soumis, au terme de celle-ci, à un questionnaire de satisfaction. Les élèves quant à eux ont été soumis à deux types de questionnaires : un premier de satisfaction et un second d'auto-évaluation sur les compétences. Une évaluation plus précise sera réalisée lors de la troisième année de recherche sur les aspects motivationnels et le développement des compétences transversales travaillées.

IV. Présentation du dispositif

Dans cette partie, nous présentons les modalités concrètes de l'implantation du dispositif dans les classes, en particulier du portfolio.

A. Le portfolio et la métacognition

Le portfolio, au service du développement des compétences transversales, poursuit l'objectif de développer les capacités métacognitives. Selon Dore, Michaud et Mukarugagi (2002), les pratiques de co-évaluation et d'auto-évaluation vont aider l'élève à progressivement prendre conscience de sa façon d'apprendre et peut agir sur ses apprentissages. Il peut davantage dire quelles sont ses forces et ses faiblesses ou encore demander de l'aide sur un point précis. Cette prise de conscience et ce contrôle sur ses propres apprentissages sont des aspects majeurs de la métacognition

Outre les entretiens entre l'élève et l'enseignant et entre les élèves qui sont de nature à développer la réflexion métacognitive (Schillings & Vanhulle, 2001; Paris & Ayrès, 2000), la métacognition est développée au cours de la construction de *fiches-processus*. Celles-ci consistent en fiches précisant la démarche à adopter afin de développer une compétence donnée. Dans le cas de notre projet où il a été décidé de travailler sur quatre compétences (présentées au point B), quatre fiches-processus ont donc été réalisées. Chaque fiche est présentée de la même manière (excepté celle de l'orthographe). Un exemple de fiche est présenté à la page suivante (voir figure 1). La compétence est tout d'abord définie sur la base des discussions avec les élèves. Cela correspond à l'intitulé "De quoi s'agit-il?" de la figure 1. Ensuite, les différents critères à respecter pour une compétence sont énoncés sous l'intitulé "fiche-outil" (voir figure 1) afin de respecter une terminologie que les élèves utilisent déjà dans certaines des classes impliquées dans le projet.

L'originalité de notre démarche consiste à impliquer activement les élèves dans la construction des fiches-processus. En effet, pour chaque compétence, les élèves testeront dans un premier temps leurs démarches spontanées, puis une mise en commun donnera aux élèves l'occasion de partager leurs stratégies et de discuter de leur efficacité. La fiche-processus sera alors élaborée puis réinvestie dans les autres matières. Concrètement, les fiches-processus seront réalisées dans le cadre du cours de français dans la mesure où celui-ci dispose de cinq périodes hebdomadaires. À cette fin, la séquence de français démarrera quelques semaines avant les autres.

Il convient de souligner que ces fiches, une fois élaborées, ne sont pas figées une fois pour toute. Il se peut que certaines adaptations soient nécessaires en fonction des différentes matières, même s'il s'agit d'une même compétence ou bien qu'au fil des expériences certaines catégories doivent être ajoutées ou affinées.

L'objectif des fiches-processus est triple. Il consiste d'une part à aider les élèves à prendre conscience des stratégies à mettre en oeuvre pour développer les compétences, mais aussi à leur permettre de travailler de manière autonome à l'aide de leur fiche. Enfin, ces fiches-processus serviront de base pour les fiches d'auto-évaluation.

JE CHERCHE DE L'INFORMATION dans un ouvrage de référence

DE QUOI S'AGIT-IL ?

Je cherche de l'information = Je cherche ce qui va me permettre de répondre à une question ou de vérifier une hypothèse.

FICHE - OUTIL

Les étapes de la recherche	Ce que je dois faire
1. Je vérifie si le document traite bien de mon sujet.	- Je lis la 1re ou la 4e de couverture.
2. J'identifie dans l'ouvrage les chapitres qui traitent de mon sujet.	- Je consulte la table des matières. - Je consulte l'index thématique.
3. Dans le chapitre, je cherche les passages qui traitent précisément de mon sujet.	- Je lis les titres. - J'observe les iconographies. - Je lis les sous-titres. - Je lis les éléments mis en évidence (mots en gras, mots soulignés, passages encadrés, ...)
4. Je lis le paragraphe que j'ai sélectionné.	

Figure 1 - Un exemple de fiche-processus pour la compétence "Je cherche de l'information".

B. Déroulement général des séquences

B.1 Répartition des thèmes

L'ensemble des séquences a été testé dans deux classes de 2e année du secondaire et une classe de 1ère année du secondaire. Le thème de chaque séquence est présenté dans le tableau ci-dessous.

Tableau 1 - Répartition des thèmes.

Classe	Discipline	Thème/question de la situation problème
1ere année	français histoire	Pourquoi le tremblement de terre au Pakistan a-t-il été si meurtrier? Quelle est l'image des Gaulois véhiculée dans la BD?
2e année	français histoire sciences géographie	Pourquoi le tremblement de terre au Pakistan a-t-il été si meurtrier? Comment expliquer l'étendue du monde musulman aujourd'hui constatée sur la carte? Comment la sportive produit-elle l'énergie nécessaire à sa course? Quelles sont les conséquences de l'évolution des moyens de transports?

Le choix des thèmes est bien sûr dicté par les directives des programmes. Pour les sciences, l'histoire et la géographie, les questions des situations problèmes sont issues des programmes concernés. En 1re année, pour des raisons indépendantes de notre volonté, seules deux (plutôt que les quatre prévues initialement) disciplines ont été concernées par le dispositif. En français, le thème intègre plusieurs objectifs du programme concernant l'écriture et la lecture du texte informatif.

B. 2. Les compétences retenues

Nous avons choisi de développer quatre compétences transversales. Les deux premières sont directement liées aux situations problèmes, les deux autres à des activités de communication. Les compétences retenues sont reprises des directives officielles de la Communauté française de Belgique (Socles de Compétences, 1999) :

*1. Émettre des hypothèses sur la base d'indices (**C_1**)*[1]

Plus précisément, cette compétence consiste à faire "identifier des indices et dégager des pistes de recherche propres à la situation" (Socles de Compétences, 1999, p. 35).

Nous exploitons en deux temps la compétence "Emettre des hypothèses" :

- sur la base de leurs connaissances antérieures
- sur la base d'indices

Ce deuxième aspect de cette compétence, généralement moins présent dans les classes, est exploité à partir d'un dossier d'indices remis aux élèves. Les indices consistent généralement en photos, en courts extraits de presse ou de Journal Parlé. Ces indices permettent également aux élèves d'élargir leurs hypothèses spontanées à d'autres pistes auxquelles ils n'auraient pas pensé.

Cette compétence donne lieu à une fiche-processus (***FPC_1***), et une auto-évaluation (***AEC_1***).

[1] Les abréviations entre parenthèses correspondent aux abréviations utilisées dans le schéma de la figure 2.

2. Récolter des informations (C_2)

Cette compétence est recommandée par les Socles de Compétences (1999) tant en éveil scientifique, qu'en éveil géographique et historique.

Idéalement, cette compétence implique d'inviter les élèves à chercher eux-mêmes, avec l'aide de l'enseignant, les documents nécessaires, mais des contraintes de temps et de praticabilité ont rendu cette démarche impossible dans le contexte de la recherche. Des dossiers ont donc été constitués par l'équipe de recherche pour la plupart des séquences. Ceux-ci rassemblent un ensemble d'informations autour des thèmes visés et présentent une table des matières. Les dossiers sont identiques pour tous les élèves. La tâche de ceux-ci consiste donc d'une part, à identifier dans le document le (ou les) chapitre(s) correspondant à la piste qu'ils doivent investiguer et d'autre part, à repérer ensuite le (ou les) passage(s) fournissant les informations utiles. Cette compétence donne lieu à une fiche-processus (***FPC$_2$***), et une auto-évaluation (***AEC$_2$***).

3. Communiquer des résultats (C_3)

Dans les Socles de Compétences (1999), cette compétence est recommandée tant en français, qu'en éveil scientifique, qu'en en éveil géographique et historique. Celle-ci sera subdivisée en deux sous-compétences :

- *Orienter son récit en fonction de la situation de communication (**C_{3a}**)*

  Nous avons envisagé trois moments importants de communication dans le développement de la démarche scientifique, quelle que soit la matière :

  1. *Le rapport d'hypothèses* où l'élève, individuellement, rédige sur la base d'indices, trois ou quatre hypothèses à l'intention de l'enseignant.
  2. *Le rapport de recherche* où l'élève, individuellement également, développe les informations qu'il a recueillies pour répondre à la piste qu'il a dû investiguer.
  3. *La présentation des résultats* : cette dernière phase fait l'objet d'un travail collectif. Elle consiste en un travail de synthèse qui prend la forme soit d'un exposé, soit d'une synthèse écrite sous la forme de tableau ou de schéma commenté.

- *Assurer la présentation orale d'un exposé*, en particulier "utiliser et repérer les procédés propres à assurer la clarté du message" et "utiliser et identifier les moyens non verbaux" (p. 20) *(C_{3b})*

  Parmi les différents modes de communication cités dans la compétence C_{3a}, la préparation d'un exposé fait l'objet d'un travail approfondi dans le cadre de notre recherche, dans la mesure où ce type de préparation fait rarement l'objet d'un travail de préparation dans les classes.

La deuxième sous-compétence fait partie de la première. Ces compétences sont donc rassemblées dans une même fiche-processus (***FPC$_3$***), intitulée "orienter son récit en fonction de la situation" qui est développée au fur et à mesure des situations. Elles font également l'objet d'une auto-évaluation (***AEC$_3$***).

4. Orthographier les productions personnelles (C_4)

Cette compétence est inscrite dans les Socles de Compétences de français (p. 17). Elle est reprise dans une fiche-processus (***FPC$_4$***) qui est propre à chaque élève, et qui est complétée par l'enseignant et l'élève, au fur et à mesure des productions personnelles formelles et informelles, en fonction des erreurs les plus communément commises par chacun. Cette fiche constitue ainsi pour l'élève un outil personnel qu'il pourra utiliser chaque fois qu'il sera amené à produire un écrit.

B. 3 Organisation générale du dispositif

Chacune des séquences est développée selon trois grandes étapes (voir tableau 2). Chacune de ces étapes se clôture par une production écrite qui rend compte des travaux propres à chacune des phases. La dernière phase, la communication des résultats, prend deux formes et est préliminaire à la synthèse écrite que le professeur distribuera à ces élèves au terme des travaux :

L'exposé en géographie et en français : Les élèves sont invités à communiquer aux autres, oralement, le résultat de leur recherche. Cet exposé utilise les ressources communicatives les plus adéquates en fonction du thème (photo, texte et schéma pour le thème en français, graphiques et textes pour la géographie).

La synthèse en histoire et en sciences : Celle-ci est réalisée en groupe et reprend l'ensemble des recherches effectuées pour investiguer toutes les hypothèses. Elle prend la forme soit d'un tableau de synthèse (en histoire), soit d'un schéma (en sciences). Ces synthèses sont ensuite échangées entre les groupes. Chaque groupe effectue alors ses commentaires sur la synthèse proposée par les autres. Une mise en commun est alors l'occasion pour l'enseignant de reprendre avec les élèves les points importants pour établir la synthèse générale.

Tableau 2 – Les séquences se divisent en trois grandes étapes.

	Phases	**Production écrite**
Étape 1	Émettre des hypothèses	Rapport d'hypothèse
Étape 2	Rechercher de l'information	Rapport de recherche
Étape 3	Communiquer les résultats	- Exposé - Tableau ou schéma de synthèse

La figure 2 ci-dessous présente l'organisation générale du dispositif. Elle montre le décalage entre la séquence en français et les autres séquences. En effet, rappelons que c'est dans le cadre du cours de français que les fiches-processus présentées dans la première partie sont construites avec les élèves. Ce cours permet un premier exercice de ces compétences et débouche sur une réflexion de nature métacognitive à travers la réalisation de ces fiches. Ces fiches construites en français sont réutilisées dans le cadre des autres cours qui constituent également des occasions d'évaluer les élèves

sur ces compétences. Cette option didactique a entraîné un décalage de deux semaines de la séquence de français par rapport aux autres séquences.

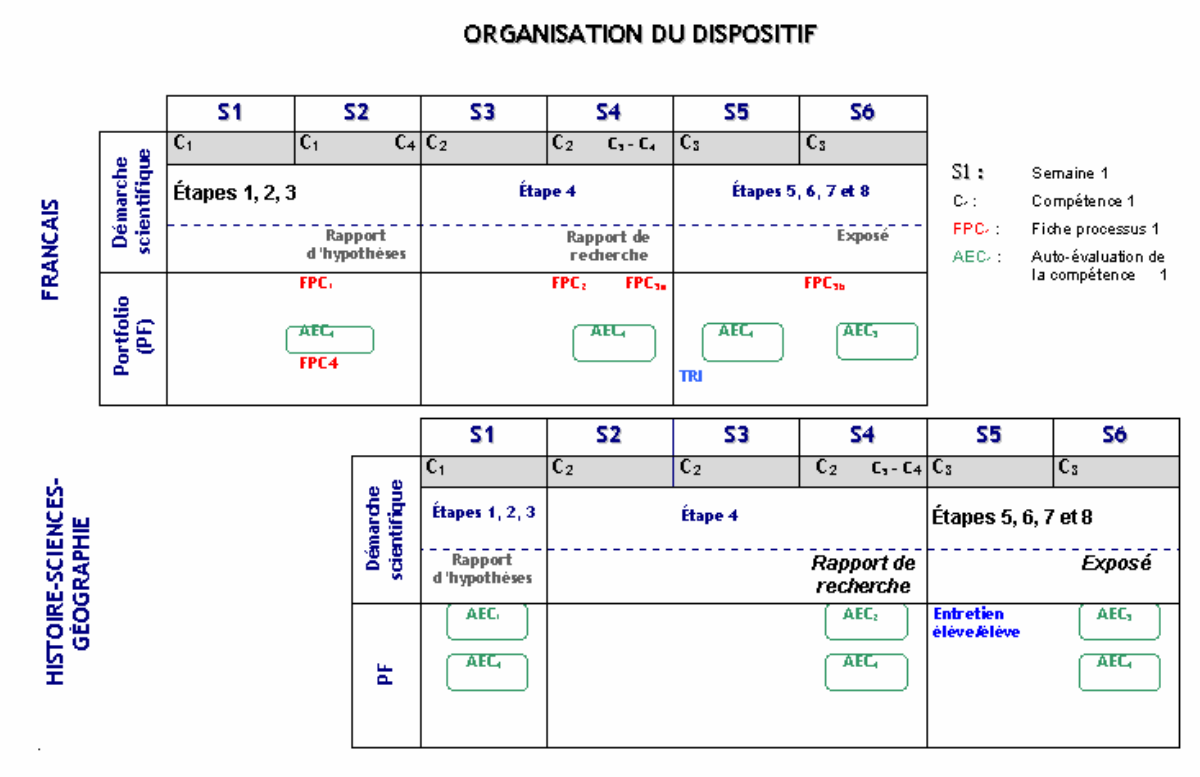


Figure 2 – Organisation du dispositif.

IV. Résultats et perspectives

Bien qu'il soit prévu d'évaluer l'impact du portfolio au cours de la 3e année de recherche au niveau de la motivation et du développement des compétences travaillées, nous avons soumis, fin de la deuxième année de recherche, les élèves ayant participé au projet ainsi que leurs enseignants à des questionnaires afin de récolter des indices sur l'effet de cette première expérimentation. Pour rappel, trois classes ont participé la recherche. au cours de cette deuxième année.

A. L'analyse des questionnaires d'auto-évaluation des élèves

De manière générale, les élèves des trois classes s'auto-évaluent positivement au niveau des quatre compétences travaillées. Entre les classes, des différences apparaissent : la classe de 2A est celle où les élèves auto-évaluent plus favorablement leurs compétences tandis que celle de 1C les auto-évaluent le moins favorablement et la classe de 2B propose une auto-évaluation intermédiaire. Ces différences peuvent sans doute s'expliquer par plusieurs facteurs :

- En 2A, le professeur de français est de suite entré dans la logique portfolio et a proposé aux élèves différentes activités de renforcement des compétences exercées (par exemple, il a proposé une séquence sur la rédaction et la formulation d'une hypothèse). Par ailleurs, c'est la seule classe où le décalage prévu initialement entre le français et les autres disciplines a été respecté et où les fiches outils ont été construites dans les temps prévus. Cet élément semble, en comparaison des résultats des autres classes, particulièrement déterminant dans le dispositif.
- En 2B, le professeur de français a pris du retard dans les activités portfolio et par manque de concertation entre les enseignants, cette information n'a pas été communiquée à temps aux autres membres de l'équipe. Aussi, les fiches-processus ont-elles été construites avec retard en français et n'ont pas toujours pu être exploitées dans d'autres disciplines.
- En 1C, sont apparus des problèmes organisationnels (absence répétée d'un enseignant pour raison de voyage et excursion scolaire) qui ont entraîné des difficultés de concertation entre les deux professeurs de l'équipe et interrompu momentanément à plusieurs reprises le travail portfolio avec les élèves. Par ailleurs, la faible auto-évaluation (par rapport aux deux autres classes) est peut-être "réaliste" dans la mesure où ces élèves-ci sont en première année et n'ont pas encore autant exercé ces compétences transversales que les élèves de deuxième année. Enfin, pour cette

classe, seules deux disciplines étaient concernées par l'expérimentation. Il était dès lors prévisible que les progrès seraient moins marqués chez ces élèves.

B. L'analyse des questionnaires de satisfaction des élèves

De façon globale, l'avis des élèves sur le dispositif et assez positif. L'intérêt du portfolio a bien été perçu par les élèves qui le considèrent surtout comme une "aide pour progresser" et comme "une façon intéressante de travailler". La majorité considère que cet outil leur a permis de réaliser des progrès au niveau des quatre compétences et déclare vouloir continuer à travailler selon cette méthodologie. L'intérêt des élèves est manifeste.

C. L'analyse des questionnaires de satisfaction des enseignants

Les avis des enseignants sont positifs. Ils considèrent en effet que le dispositif a permis d'atteindre les objectifs fixés au départ (développement de compétences) et que les activités proposées aux élèves ont suscité leur intérêt ("Très enthousiastes"; "Ils ont vraiment apprécié"). Par ailleurs, ils soulignent que l'aspect transversal est non seulement intéressant pour les élèves mais également pour les professeurs ("Il est grand temps de trouver une méthode d'enseignement qui relie tous les cours", "l'idée est excellente"). Certains déclarent avoir éprouvé des difficultés mais les éléments pointés semblent facilement surmontables (difficulté de déplacer les portfolios d'un cours à l'autre; nécessité de préparer (aussi) l'examen du district).

Ces différents résultats sont très encourageants. Il semble en effet que le dispositif construit permette le développement des compétences transversales et que les activités proposées génèrent l'intérêt (motivation) des élèves. En outre, les enseignants et les élèves se disent satisfait du dispositif. Cependant, des améliorations à apporter au présent dispositif sont apparues au cours de l'expérimentation. Tout d'abord, il semble que les progrès des élèves soient plus importants lorsque plusieurs disciplines participent au projet. Les progrès constatés dans le cas de la participation de deux disciplines sont en effet nettement plus faibles. D'autre part, il semble important de s'assurer que le travail des élèves n'est pas interrompu (par des vacances scolaires, par exemple) au risque de faire perdre le sens des activités aux yeux des élèves. Une troisième condition à l'efficacité du dispositif est le respect du timing au sein de chaque discipline et à cette fin de prévoir des réunions de concertation entre les enseignants. Ainsi, les outils (fiches-processus par exemple) construits dans le cadre du cours de français devant être réinvestis dans les autres disciplines, il est évident qu'un retard dans la construction de ces fiches met en péril l'ensemble du dispositif. Au terme de cette expérimentation, il est apparu plus prudent de

commencer plus tôt la séquence en français (soit quatre semaines avant les autres séquences au lieu des deux semaines prévues initialement). Enfin, les enseignants ont semble-t-il éprouvé, au début de l'expérimentation, des difficultés à identifier clairement ce qu'ils devaient réaliser avec leurs élèves et quand ils devaient le réaliser. A cette fin, nous avons aujourd'hui construit un descriptif précis pour chaque séquence. Celui-ci aura l'occasion d'être testé auprès d'enseignants n'ayant pas participé aux travaux des deux premières années de recherche.

Références

Allal, L., Cochet-Kaeser, F., Bonaiti-Dugerdil, S. & Wegmuller, E. (1998). Le portfolio dans la dynamique de l'entretien tripartite. *Mesure et Evaluation en Education, 20*, 5-32.

Conseil de la Communauté française de Belgique (1997). *Décret du 24 juillet 1997 définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre.* (Moniteur belge : 23 septembre 1997).

De Vecchi, G. (2005). *Une banque de situations-problèmes tous niveaux : tome 2*. Paris : Hachette Education.

Dore, L., Michaud, N., & Mukarugagi, L. (2002). *Le portfolio, évaluer pour apprendre*. Montréal : Chenelière.

Fabre, M. (1999). *Situations –problèmes et savoirs scolaires*. Paris : PUF.

Farr, R. & Tone, B. (1998). *Le portfolio au service de l'apprentissage et de l'évaluation*. Montréal : Chenelière.

Forget, A. (2001). Pour en finir avec les apprentissages en porte-à-faux : le portfolio? *Caractères*, 4, 18-30.

Ghaye, B. & Rahier, N. (1999). *Comment apprendre à apprendre?* Averbode.

Goupil, G. (1998). *Portfolios et dossiers d'apprentissage*. Montréal : Chenelière.

Green, B. (1997). *Nouveaux paradigmes pour la création d'écoles de qualité*. Montréal : Chenelière.

Ministère de la Communauté française de Belgique (1999). *Socles de compétences : Enseignement fondamental et premier degré de l'Enseignement secondaire*. Bruxelles : Auteur.

Paris, SG., & Ayres, LR. (2000). *Réfléchir et devenir. Apprendre en autonomie. Des outils pour l'enseignant et l'apprenant*. Traduction par M. Aussanaire-Garcia, de l'ouvrage américain: Becoming Reflective Students and Teachers. With Portfolios and Authentic Assessment (1994). Paris/Bruxelles: De Boeck Université s.a.

Perrenoud, P. (1995), Des savoirs aux compétences : les incidences sur le métier d'enseignant et sur le métier d'élève. *Pédagogie collégiale*, *9*(2), 6-10.

Poissant, H., Poëlhuber, & B., Falardeau, M. (1994). Résolution de problèmes, autorégulation et apprentissage. *Revue canadienne de l'éducation*, *19*(1), 30-44.

Presseau, A. (2004). *Intégrer l'enseignement stratégique dans sa classe*, Montréal : Chenelière.

Schillings, A. & Vanhulle, S. (2001). Petit parcours dans la littérature francophone sur le portfolio scolaire ou le portfolio en cinq questions. *Caractères*, 4, 6-13.

Tardif, J. (1997). *Pour un enseignement stratégique*. Montréal : Les Éditions Logiques.
Viau, R. (1994). *La motivation en contexte scolaire*. Bruxelles : De Boeck.
Vygotsky, L. (1985). *Thought and language*. Cambridge, MA : Harvard University Press.

UN PORTFOLIO CERTIFICATIF, CLE DE VOÛTE D'UN PROGRAMME DE FORMATION PEDAGOGIQUE OBLIGATOIRE DES ENSEIGNANTS DE L'ENSEIGNEMENT SUPÉRIEUR. VERS UNE ALLIANCE ENTRE UNE FONCTION FORMATRICE INTÉGRATIVE ET UNE FONCTION CERTIFICATIVE ?

Léopold Paquay, Université catholique de Louvain
Catherine Van Nieuwenhoven, HEG-ISPG

Résumé :
Cette communication constitue pour l'essentiel une analyse de pratique. En 2002, en Communauté française de Belgique, à la suite du décret imposant une formation pédagogique obligatoire des enseignants de hautes écoles, le programme de formation « CAPAES -*Certificat d'aptitude pédagogique approprié à l'enseignement supérieur* » développé à Louvain-la-Neuve a été charpenté autour d'un dispositif de portfolio. La communication consiste d'une part à expliciter les intentions des concepteurs de ce portfolio certificatif et des outils développés, et d'autre part, sur la base d'une évaluation fine de ce dispositif par les acteurs, de dégager quelques conditions transférables quant à la problématique de la tension entre évaluation formatrice à visée intégrative et évaluation certificative.

Le programme « Capaes »

Ce programme, dont les axes sont définis par décret (CFB, 2002), a été développé en collaboration entre l'Université catholique de Louvain et le Centre d'enseignement supérieur de promotion sociale du Brabant wallon (CPFB), Ce programme[1] fonctionne depuis 3 ans ; il s'adresse actuellement à une centaine de candidats, tous enseignants en exercice en haute école.

Les activités principales sont :
(1) un cours théorique de 60 heures axé sur les approches psycho-socio-pédagogiques de l'enseignement supérieur, précédé d'un cours de base pour les candidats non porteurs d'un titre pédagogique,
(2) des séminaires d'intégration (de 10 h. ou de 30 h.) et
(3) un accompagnement de la pratique en haute école.

Les logiques formatives sont dominantes : il n'y a pas d'examens portant sur les cours théoriques. La validation finale se réalise à travers un portfolio. Les enjeux certificatifs de ce portfolio sont importants : en effet, c'est sur la base de ce portfolio qu'au terme du programme, une attestation de réussite est accordée. Pour obtenir le

[1] Les axes du programme sont décrits sur le site
< http://serveur.adef.ucl.ac.be/Capaes/default.htm>

CAPAES, le candidat doit déposer cette attestation de réussite ainsi qu'un « dossier professionnel » à la Commission de la CFB qui attribue ou non le Certificat CAPAES.

En pratique, le portfolio constitue une préparation directe du « dossier professionnel ». Il s'agit d'un *travail de synthèse* dans lequel le candidat décrit et analyse son parcours professionnel et le parcours de formation mais également où il analyse ses pratiques pédagogiques en haute école à la lumière des grilles de lecture (concepts, modèles et théories) développées dans le programme. Ce sont surtout les séminaires d'intégration qui constituent des lieux privilégiés d'analyse des parcours et des pratiques avec un accompagnement collectif et individuel des titulaires.

Problématique

La littérature sur le portfolio est large. Le portfolio constitue un des outils reconnus pour favoriser la réflexivité sur la dynamique d'un parcours et permettre l'intégration d'acquis multiples ; il constitue aussi la base d'une évaluation authentique (Bélair, 2002 ; Lévesque & Boisvert, 2001 ; Mottier Lopez, 2005 ; Segers, Dochy & Cascallar, 2003 ;…).

Notre dispositif du portfolio sera analysé principalement en référence à deux questions aux enjeux pratiques et théoriques :

A/A quelles conditions le portfolio constitue-t-il un outil d'analyse réflexive et d'intégration d'acquis divers ? Comment et à quelles conditions le programme contribue-t-il à l'analyse de pratiques professionnelles et à l'écriture du portfolio ? Plus largement : la réalisation obligatoire d'un portfolio favorise-t-elle l'intégration des acquis ? (Desjardin, 2002 ; Vanhulle & Mottier Lopez, 2004)

B/La visée certificative dominante ne perturbe-t-elle pas la visée formative ?
Dans ce programme "Capaes", la volonté a été de concilier le caractère formatif du portfolio et d'utiliser celui-ci comme une base d'une évaluation certificative authentique. Pour des raisons de fiabilité de l'évaluation certificative du portfolio, des critères et indicateurs précis ont été définis au risque d'en perdre partiellement le caractère formatif, dans une dynamique personnelle.
Dans quelle mesure les enjeux certificatifs s'allient-ils à la dynamique formative ? Remplit-on les conditions favorables pour que ces pratiques d'évaluation contribuent au développement professionnel ? (Cf. Paquay, 2002 et 2004). Cette deuxième question sera traitée en profondeur dans la communication.

Méthodologie

Les deux questions ci-dessus seront documentées – en référence à la littérature- sur la base de trois ensembles d'informations :

- des questionnaires d'évaluation du programme appliqués à l'ensemble des candidats en 2004, en 2005 et en 2006 (entre 50 et 100 questionnaires reçus chaque année)

- des entretiens de groupe des enseignants intervenants dans le programme et des co-titulaires du séminaire d'intégration ;
- une analyse de quelques portfolios et rapports d'évaluation.

Traitements : dans les questionnaires écrits, nous procéderons à une comparaison des pourcentages relatifs aux items à choix de réponse et à une analyse catégorielle de contenu des commentaires écrits ; pour le reste, nous nous basons sur une analyse thématique de contenu des entretiens de groupe et des portfolios sélectionnés.

Résultats

L'analyse est en cours. Sur la base des données actuellement recueillies, quelques constats ressortent. Mettons ici en relief deux constats en lien avec la problématique exprimée :

- Dans le cadre des cours, des dispositifs de pédagogie active avec des activités fréquentes d'analyse de cas, de moments réflexifs sont reconnus par les étudiants comme des opportunités pour préparer leur portfolio. Par ailleurs, les séminaires sont des lieux privilégiés, par les interactions structurées entre professionnels, de susciter une dynamique de réflexion sur la pratique et d'accompagner les premiers pas d'écriture réflexive. Dans la communication, nous ferons état des conditions favorables à un tel « transfert » des acquis de formation dans la production intégratrice qu'est le portfolio.
- Les consignes précises et détaillées relatives au portfolio, ainsi que la grille critériée d'évaluation, semblent constituer davantage un support qu'un frein à une analyse réflexive personnelle. Nous préciserons les conditions qui déterminent cet effet positif déclaré sur le développement professionnel.

Perspectives

Nous clôturerons cette étude de cas par la comparaison des résultats avec ceux d'autres travaux récents (Struyven, Gielen & Dochy, 2003 ; Frenay, Wouters, Bourgeois & Galand, 2005 ; Mottier Lopez & Wegmuller, 2005 ; Van Nieuwenhoven, Ledur & Labeeu, 2005).

Références bibliographiques

Bélair, L. (2002). L'apport du portfolio dans l'évaluation des compétences, *Revue les sciences de l'éducation en question*, 1, 1, 17-37

CFB (2002). *Décret définissant le Certificat d'aptitude pédagogique approprié à l'Enseignement supérieur (CAPAES) en Hautes écoles et ses conditions d'obtention*.. Bruxelles : CFWB www.agers.cfwb.be/prof/espaces/sup/capaes.asp

Desjardin, R. (2002*). Le portfolio de développement professionnel continu*, Montréal : Chenelière/McGraw-Hill, 108 p.

Frenay, M., Wouters, P., Bourgeois, E. & Galand, B. (2005). *Teaching portfolios as a mean of evaluating the impact of two certification in university teaching and learning programs for higher education teachers*. Paper presented at the 2005 American Educational Research Conference, Montréal, 11-15/04/05.

Lévesque, M. & Boisvert, E. (2001). *Portfolio et formation à l'enseignement.* Montréal : éditions Logiques, 135 p.

Mottier Lopez, L. & Wegmuller, E. (2005). Le portfolio en formation initiale d'enseignants primaires : étude d'entretiens de co-évaluation dans une perspective d'évaluation située et authentique. *Communication au colloque de l'Adee-Europe*, Reims, 24-26 oct (Actes sur CD-Rom).

Paquay, L. (Dir.) (2004). *L'évaluation des enseignants. Tensions et enjeux*. Paris : L'Harmattan, 330 p.

Paquay L., Defêche, N. & Dufays, J-L. (2002). Comment concilier évaluation formative et évaluation certificative ? Quels apports de l'apprenant ? In L. Paquay, G. Carlier, L. Collès & A-M. Huynen (Eds), *L'évaluation des compétences chez l'apprenant. Pratiques, méthodes et fondements* (85-95), Louvain-la-Neuve, Presses universitaires de Louvain.

Segers, M., Dochy F. & Cascallar E. (2003). *Optimising new modes of assessment. In search of qualities and standards.* Doordrecht (Ndl) : Kluwer Academic Publishers.

Struyven, K., Gielen, S., & Dochy, F. (2003). Students' perceptions on new modes of assessment and their influence on student learning : the portfolio case. *European Journal of School Psychology, 1* (2), 199-226.

Université catholique de Louvain (2005). *Programme du Capaes UCL-CPFB.* Site du Capaes : < http://serveur.adef.ucl.ac.be/Capaes/default.htm>

Vanhulle, S. & Mottier Lopez L. (2005) Le portfolio, ou de l'évaluation comme levier de développement professionnel. *Communication au colloque de l'Adee-Europe*, Reims, 24-26 oct (Actes sur CD-Rom)

Van Nieuwenhoven, C., Ledur, D. & Labeeu. (2005) *Le portfolio en formation des enseignants, outil efficace de développement et d'évaluation des*

compétences professionnelles. Communication au 18e Congrès de l'ADMEE-Europe, Reims, IUFM Champagne-Ardenne, septembre 2005).

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

EQUITE ET EFFICACITE: QUELQUES PROBLEMES DE MESURES DANS UN CONTEXTE D'ANALYSE LOCALISE DE PISA

Jean Moreau, URSP Lausanne
Christian Nidegger, SRED Genève

Mots-clés : équité, efficacité, compétences.

Résumé :

L'enquête PISA est désormais devenue incontournable dans les débats sur l'éducation en Suisse et en Suisse romande. Cette enquête a été l'occasion non seulement de comparer la Suisse (en tant qu'entité) mais aussi les différents systèmes scolaires qui la composent; l'éducation étant du ressort des différents cantons. Souvent les résultats présentés s'appuient prioritairement sur des comparaisons de moyenne des performances dans le domaine le plus étudié de l'enquête, par exemple les mathématiques en 2003. Ces performances sont mises en relation avec différentes caractéristiques des élèves ou informations sur leur environnement familial ou scolaire. On trouve également des analyses mettant en évidence la dispersion des résultats. Toutes ces analyses essayent de comprendre le fonctionnement des systèmes éducatifs notamment en termes d'efficacité et d'équité des systèmes ou sous-systèmes scolaires.

Les premiers résultats présentés par PISA s'appuient sur des instruments sophistiqués qui permettent de mettre en évidence du point de vue statistique de faibles variations. Cependant le recours massif à des comparaisons basées sur les moyennes peut occulter des variations non négligeables. Toutefois, la prise en compte de l'écart-type ne suffit pas toujours à caractériser suffisamment ces variations. En effet, cette mesure de la dispersion ne prend pas en compte la structure des populations scolaires. De plus, la plupart des analyses des performances des élèves s'appuient sur les résultats dans le domaine prioritaire de l'enquête ou bien comparent les différents domaines entre eux. La constitution de profils d'élèves qui prennent en compte simultanément les différents domaines peut donner un éclairage autre de ces performances.

Place de Pisa dans un contexte local

L'enquête PISA est désormais devenue incontournable dans les débats sur l'éducation en Suisse et en Suisse romande. Cette enquête a été l'occasion non seulement de comparer la Suisse (en tant qu'entité) mais aussi les différents systèmes scolaires qui la composent; l'éducation étant du ressort des différents cantons. Souvent les résultats présentés s'appuient prioritairement sur des comparaisons de moyenne des performances dans le domaine le plus étudié de l'enquête, par exemple les mathématiques en 2003. Ces performances sont mises en relation avec différentes caractéristiques des élèves ou informations sur leur environnement familial ou scolaire. On trouve également des analyses mettant en évidence la dispersion des résultats. Toutes ces analyses essayent de comprendre le fonctionnement des systèmes

éducatifs notamment en termes d'efficacité et d'équité des systèmes ou sous-systèmes scolaires.

Caractéristiques des populations et organisation scolaire

La Suisse romande offre sur un espace restreint une assez grande diversité scolaire et démographique. Soulignons d'abord que les caractéristiques des populations (langue parlée à la maison, origine de l'élève, niveau socio-économique, etc.) des cantons romands sont relativement différentes comme le montre le tableau ci-dessous. De plus, pour le secondaire I (élèves de 12-15 ans) chaque canton est libre de son organisation et de sa structure scolaire. Sur sept cantons, on trouve ainsi 4 cantons avec un système à trois filières où les élèves sont regroupés dans des classes différentes en fonction de leur niveau, les conditions d'accès à ces filières étant spécifiques à chaque canton. Deux cantons ont une structure mixte, avec filières pour certains établissements ou groupes d'élèves et parallèlement une organisation en classes hétérogènes avec des niveaux dans certaines disciplines. Enfin, un canton a un système de classes hétérogènes avec niveaux dans certaines disciplines.

Tableau 1 - caractéristiques des populations et organisation scolaire

	systèmes à filières				systèmes mixtes		système hétérogène
	FR	NE	BE	VD	VS	GE	JU
NSE faible en %	30.5	26.9	34.9	24.5	28.1	*23.8*	**31.7**
NSE élevé en %	30.5	26.9	*19.5*	24.4	22	**30.7**	19.8
Non-natif en %	17.1	22.4	18.9	24.2	18.5	**44.3**	*10.2*
allophone en %	11.5	12.8	16.5	13.7	11.2	**21.7**	*7.6*

En gras, valeur la plus élevée - *en italique,* valeur la plus faible

NB: dans les tableaux et figures qui suivent les cantons sont désignés par leur abréviation usuelle: FR=Fribourg, NE=Neuchâtel, BE=Berne, VD=Vaud, VS=Valais, GE=Genève, JU=Jura

On observe que Genève, canton urbain par excellence, a la plus grande proportion d'élèves de niveau socioéconomique élevé, de non-natifs et d'allophones, alors que comparativement, le Jura affiche les proportions les plus faibles pour trois de ces caractéristiques.

Quelques interrogations d'ordre méthodologique

Les résultats présentés par PISA s'appuient sur des instruments sophistiqués qui permettent de mettre en évidence du point de vue statistique de faibles variations. Cependant, le recours massif à des comparaisons basées sur les moyennes peut occulter des variations non négligeables. Par ailleurs, la prise en compte de l'écart-type ne suffit pas toujours à caractériser suffisamment ces variations. En effet, cette mesure de la dispersion ne prend pas en compte la structure des populations scolaires. De plus, la plupart des analyses des performances des élèves s'appuient sur les résultats dans le domaine prioritaire de l'enquête ou bien comparent les différents domaines entre eux. La constitution de profils d'élèves qui prennent en compte

simultanément les différents domaines peut donner un éclairage autre de ces performances. On cherchera à mettre en évidence ces différentes questions en s'appuyant sur les analyses complémentaires de PISA 2000 et 2003 réalisées en Suisse romande. Nous illustrerons à travers quelques exemples les problématiques soulevées ci-dessus.

Quelques exemples liés à la question de l'équité

L'efficacité et équité des systèmes scolaires sont souvent décrites à l'aide de différentes statistiques globales portant sur les performances des élèves. La performance moyenne d'une population scolaire donnée rendrait compte de l'efficacité du système et la dispersion des performances (mesurée par la variance ou l'écart-type) de son équité. Les résultats des enquêtes internationales (PISA) comparent en général les résultats moyens et les écart-types pour juger de leurs performances respectives. Remarquons que ces comparaisons ne prennent pas en compte les structures des différentes populations scolaires. En effet, comme le montre le tableau 1 les populations scolaires comparées ne sont souvent pas homogènes et se distinguent par rapport à l'importance de certaines catégories scolaires particulièrement vulnérables sur le plan scolaire (élèves d'origine étrangère, élèves ne parlant pas le français à la maison, élèves en retard scolaire, élèves de milieu défavorisé, etc…). La seule moyenne ne peut donc pas, en raison des disparités des populations, refléter de façon adéquate l'efficacité du système et permettre des comparaisons valides. Il est utile dans ce cas, de comparer les systèmes scolaires pour ces différentes catégories d'élèves à risque, pour évaluer leur degré d'efficacité.
De même la variance ne saurait constituer à elle seule une mesure pertinente de l'équité. En effet, l'hétérogénéité de la population scolaire induit également par nature une variabilité des performances qui ne peut être attribuée à la seule absence d'équité. Lorsqu'il existe des sous-populations scolaires importantes d'élèves défavorisés, les compétences des élèves sont naturellement plus dispersées. L'analyse des résultats de l'enquête PISA montre que l'origine de la famille, la langue parlée à la maison, le niveau socioéconomique de la famille sont parmi les facteurs explicatifs les plus importants. Ces facteurs expliquent donc en partie l'importance de la variance des résultats. La comparaison des variances ou des écart-types des résultats (tableau 2) est donc problématique pour des populations scolaires qui seraient structurellement très dissemblables. En outre, il est difficile de distinguer certains cantons (les différences sont souvent non statistiquement significatives).

Tableau 2 – écart-type en mathématiques

	Ecart-type en mathématiques	Erreur standard
FR	78.05	2.13
VS	78.63	1.47
JU	74.35	2.59
NE	78.79	1.29
BE	84.63	2.57
VD	84.78	1.58
GE	85.33	1.62

Analyses en terme de risques

De même que celle de l'efficacité, l'analyse de l'équité des systèmes doit prendre en compte la structure de la population scolaire. Pour cela, on peut apprécier les différences d'équité entre les cantons en comparant certaines sous-populations scolaires en terme de risque.
On analyse ici les différences d'équité en terme de risque pour certaines catégories d'élèves (Moreau, 2004). Dans le cas de la comparaison des résultats des élèves dans l'enquête PISA, nous considèrerons le risque pour un élève d'avoir des difficultés (avoir un niveau inférieur ou égal au niveau 2) dans un domaine particulier (par exemple les mathématiques). Nous comparerons successivement les risques relatifs de certaines catégories d'élèves pouvant être considérées comme vulnérables dans le contexte scolaire et associées aux variables fondamentales structurant la population scolaire : les garçons comparés aux filles, les élèves en retard (16 ans) comparés aux élèves ayant l'âge normal (15 ans) en 9e année, les élèves parlant le français à la maison et les autres. Le facteur socioéconomique n'est pas une variable catégorielle et sera examiné à part. Le calcul de cet indice dépend bien sûr du seuil choisi pour définir le risque. Nous avons choisi ici le niveau 2 en mathématiques correspondant à un niveau de maîtrise souvent considéré comme insuffisant pour les élèves de 9e année.

On rappelle que le risque relatif pour une catégorie par rapport à une autre de présenter une certaine caractéristique est défini par le rapport des taux de chacune des catégories considérées. Par exemple dans le canton de Vaud, le risque est égal à 1,7 pour les non francophones en mathématiques. Donc il y a proportionnellement 1,7 fois plus de non francophones que de francophones dans les niveaux inférieurs ou égaux à deux en mathématiques. Le risque 1 correspondant à l'équité parfaite.

Le tableau suivant présente pour différentes catégories d'élèves les risques relatifs d'être en difficulté en mathématiques (enquête 2003). Dans le tableau, les cantons sont classés en fonction de leur moyenne en mathématiques, de la plus élevée à la plus faible. Ces comparaisons permettent de se rendre compte si une catégorie d'élèves est particulièrement prétéritée par rapport à une autre dans une population scolaire.

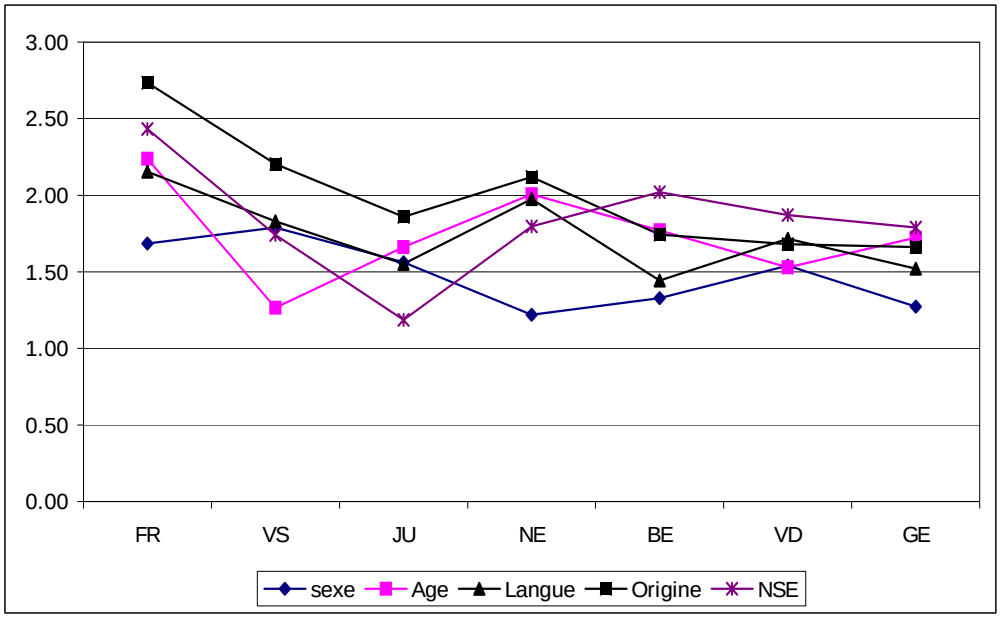


Figure 1 - Comparaison des risques relatifs pour certaines catégories d'élèves en mathématiques (2003)

Les résultats présentés corroborent les résultats de l'enquête 2000 où les cantons les plus efficaces, par exemple Fribourg, n'apparaissent pas toujours comme les plus équitables. Fribourg et Valais apparaissent comme peu équitables pour les élèves non natifs. Dans le canton de Fribourg, les élèves qui ne sont pas nés en Suisse ont par rapport à leurs camarades un risque près de trois fois plus grand d'être en difficulté en mathématiques. On notera un risque relativement faible pour les élèves allophones genevois (dont la proportion dans la population scolaire est élevée) qui pourrait s'expliquer par des mesures prises pour assurer l'intégration de ces élèves alors que le risque élevé observé à Neuchâtel (où la proportion des allophones est assez importante) pourrait être le signe d'une politique différente en la matière. Cette mise en évidence des disparités de performance pour certaines populations scolaires donne une image quelque peu différente de l'équité des systèmes éducatifs romands et nous semblent mieux à même de cerner le caractère inéquitable d'une situation scolaire que la variabilité globale des résultats.

Le cas du niveau socioéconomique

L'indice du statut socioéconomique n'étant pas une variable dichotomique ou catégorielle, on peut considérer différentes approches pour analyser cet aspect. L'influence du niveau socioéconomique est souvent appréhendée par le calcul du gradient, c'est-à-dire la droite de régression expliquant la performance (les mathématiques dans l'exemple ci-dessous) en fonction du niveau socioéconomique de la famille. La pente de cette droite ou coefficient de régression nous renseigne sur l'influence du niveau socioéconomique sur les performances. La progression de 1

point de niveau socioéconomique a un impact moyen de 18 à 29 points suivant les cantons.
Outre la difficulté d'interpréter ce résultat, on fait ici l'hypothèse que le modèle linéaire est le modèle le plus adapté pour caractériser la liaison entre ces deux variables. il est également difficile au vu des résultats de différencier les différents cantons (sauf le Jura), d'autant plus que les valeurs sont entachées d'une certaine imprécision (erreur standard égal à 3 environ).

Tableau 3 - moyenne en mathématiques et coefficient de régression(2003)

	Moyenne en mathématiques	**Coefficient de régression**
FR	553	22
VS	549	27
JU	540	18
NE	528	28
BE	526	29
VD	524	29
GE	508	29

On peut également considérer le risque relatif pour certaines catégories d'élèves en fonction du niveau socioéconomique de la famille. Les quartiles du niveau socioéconomiques fournissent quatre catégories d'élèves : les deux catégories extrêmes concernant les élèves dont les parents ont un niveau socioéconomique bas et ceux de niveau élevé et les deux catégories (2ème et 3ème quartile) formant les 50% médians.

Tableau 4 - risque relatif en fonction du niveau socioéconomique (2003)

	Risques relatifs des catégories socioéconomiques extrêmes (1e quartile et dernier quartile)	Risques relatifs des catégories socioéconomiques définies par le premier et le 2ème quartile	Risques relatifs des catégories socioéconomiques médianes (2ème quartile et 3ème quartile)	Risques relatifs des catégories socioéconomiques définies par le le 3ème et le 4ème quartile
FR	3.45	1.95	1.18	1.53
VS	2.74	1.28	1.31	1.64
JU	1.58	0.99	1.28	1.26
NE	2.73	1.28	1.42	1.50
BE	2.46	1.73	1.23	1.17
VD	2.80	1.49	1.11	1.70
GE	2.76	1.27	1.19	1.83

Le risque relatif, calculé pour les catégories socioéconomiques extrêmes, force un peu le trait. Cependant, il donne une idée des écarts maximaux pouvant exister entre des élèves de milieux socioéconomiques différents. On constate des différences importantes entre les cantons. L'analyse des risques pour ces élèves présente le canton de Fribourg comme plus inégalitaire (risque égal à 3.4 environ) et le Jura plus

équitable (risque égal à 1.58). La comparaison des catégories médianes montre des écarts plus faibles entre ces catégories et des différences moindre entre cantons. La comparaison des risques relatifs entre les catégories socioéconomiques successives révèlent une réalité complexe. Dans certains cantons, les deux catégories extrêmes singularisent (Fribourg, Vaud), dans d'autres, la catégorie. la plus basse est particulièrement prétérité (Berne) ou la catégorie la plus haute favorisée (Genève). Toutefois, il faut noter que dans le cas d'une variable continue comme le niveau socioéconomique, les valeurs du risque relatif dépendent des seuils choisis pour définir les catégories considérées.

Ainsi, par rapport aux analyses qui se centrent sur la moyenne, la variance et l'écart-type, on obtient une appréciation de l'équité quelque peu différente de celle qui a souvent été mise en avant dans la présentation des premiers résultats de PISA. Notamment, il s'avère que le canton de Fribourg, dont la performance moyenne est élevée et la dispersion relativement faible, est également caractérisé par un risque plus grand que dans la plupart des autres cantons, pour certains groupes d'élèves (allophones, niveaux socioéconomique faibles, etc.) de ne pas atteindre un niveau de performance satisfaisant.

Profils de performances et efficacité des systèmes

Un autre aspect mérite d'être étudié pour affiner l'analyse des résultats, il s'agit de prendre en compte les différents domaines testés par PISA à travers la création de profils de performances.

Comme on l'a vu ci-dessus, en Suisse romande, les résultats de PISA, aussi bien en 2000 qu'en 2003, ont montré des différences de moyennes entre les cantons mais surtout une forte variation des résultats à l'intérieur de chaque canton notamment en fonction des filières et/ou des niveaux des élèves. Par ailleurs, les résultats montrent que pour chaque domaine testé par PISA, des recouvrements importants entre les filières et/ou niveaux des élèves à l'intérieur des cantons quel que soit le domaine pris en compte. Par contre les différences entre les domaines sont relativement peu marquées. Par exemple l'ordre des cantons en fonction des moyennes est globalement le même en mathématiques et en lecture. De même, lorsque l'on compare les écarts entre les élèves les plus performants et les élèves qui obtiennent les résultats les plus faibles, mesurés ici par la différence de points entre le 5e et le 95e centile, montre que le Jura est le canton ayant l'écart le plus petit dans les deux domaines alors qu'il est le plus élevé à Genève, (voir tableau 5).

Tableau 5 - moyennes cantonales et dispersion, écart entre le 5e et le 95e centile (2003)

	moyenne math	moyenne lecture	dispersion math	dispersion lecture	total dispersion	différence dispersion*
FR	**553**	**519**	254	247	501	-23
VS	549	517	260	240	500	-22
JU	540	502	*242*	*235*	*478*	*0*
NE	528	495	259	259	518	-40
BE	526	491	**280**	268	549	-71
VD	524	498	**280**	**274**	**554**	**-76**
GE	*508*	*484*	**280**	273	553	-75

En gras valeur plus élevée, *en italique* valeur la plus faible
*nombre de points par rapport à la dispersion la plus faible (JU)

Dans le but d'observer s'il existait une hétérogénéité des performances des élèves en fonction des domaines testés, nous avons établi des profils de performances basés sur les performances en mathématiques et en lecture. Les élèves ont été répartis en trois groupes pour chaque domaine à partir des niveaux de compétences PISA. Le premier groupe rassemble les élèves ayant les niveaux 1 ou 2 en mathématiques ou en lecture. Le troisième groupe comprend les élèves ayant les niveaux 4 ou 5 en lecture ou 5 ou 6 en mathématiques. Les autres élèves sont attribués au deuxième groupe (niveau 3 en lecture ou niveau 3 ou 4 en mathématiques). Il n'existe pas de niveaux de compétences PISA en sciences, nous n'avons donc pas pu prendre en considération ce domaine dans notre analyse. La comparaison des profils ainsi obtenus donne le graphique ci-dessous.

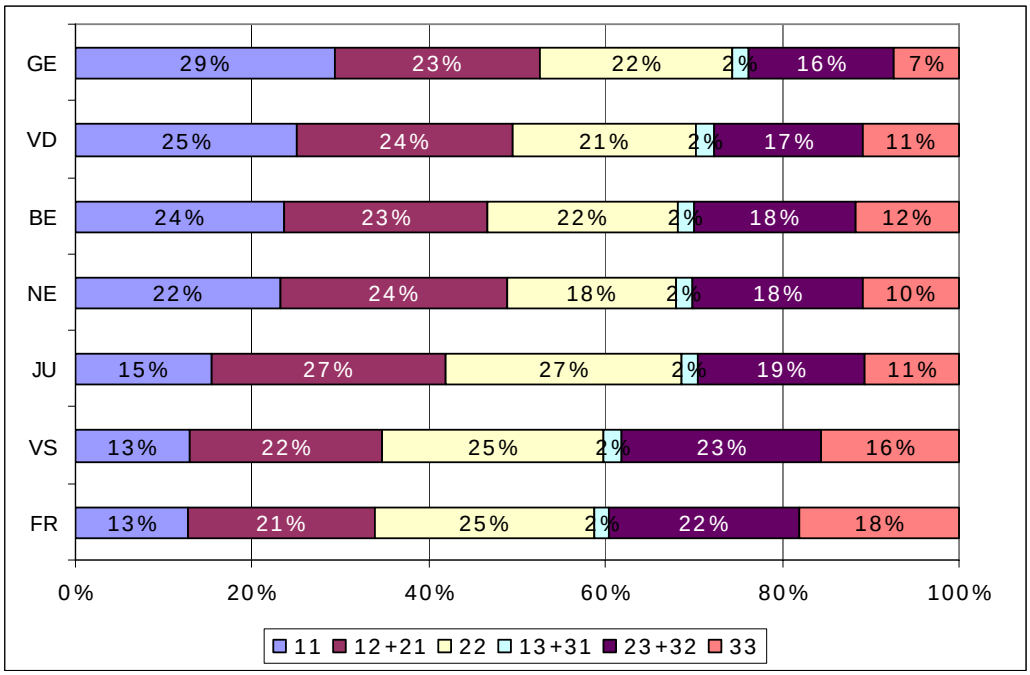


Figure 2 - Profils des élèves en fonction de leurs résultats en mathématiques et en lecture (2003)

La figure 2 montre les profils des élèves en mathématiques et en lecture. Les chiffres indiquent les différents profils. Le profil 11 correspond aux élèves qui se situent au niveau 1 (le plus faible) pour les deux domaines. Le profil 33 désigne les élèves qui ont un niveau 3 dans les deux domaines. Les autres profils correspondent aux autres combinaisons possibles de niveaux.

Les profils vraiment hétérogènes, niveau 1 dans un domaine et niveau 3 dans un autre sont rares (de l'ordre de 2%). On constate assez logiquement la fréquence des élèves les plus faibles suit le même ordre que l'ordre des résultats moyens selon les cantons, cependant les différences sont importantes entre les cantons. 13% des élèves fribourgeois ont un profil d'élève faible alors que c'est le cas de 30% des élèves

genevois. Les différences sont également importantes entre les cantons extrêmes en ce qui concerne les élèves ayant un profil d'élève performants, profil 33 (7% des genevois et 18% des fribourgeois). Toutefois, les différences sont faibles entre les 4 cantons situé au milieu (Vaud, Berne, Neuchâtel, Jura). La proportion d'élèves ayant des profils homogènes faible (profil 11) ou fort (profil 33) est légèrement plus grande dans les cantons ayant les moyennes les plus faibles. On notera que le Jura et Neuchâtel se distinguent par un pourcentage relativement plus élevé d'élèves faibles dans ce dernier canton alors que le Jura a un pourcentage plus faible d'élèves performants par rapport au Valais et à Fribourg. Par ailleurs, c'est le Jura qui a la part la plus faible de profils homogènes. Il est à noter que ce canton a une moyenne plutôt élevée dans les deux domaines, il est le seul canton avec une structure scolaire hétérogène pour tous ses élèves. On pourrait se demander si à cette structure scolaire hétérogène correspond des profils de performances plus hétérogènes des élèves. De plus, il est caractérisé par une plus faible proportion d'élèves allophones et de non-natifs.

Dans un deuxième temps, nous avons observé ce qui se passe à l'intérieur des cantons du point de vue des profils en fonction des filières cantonales. Rappelons que le choix d'une organisation en filière et la répartition des élèves dans ces filières est du ressort des cantons. A titre d'exemple nous prendrons en compte deux cantons qui ont une structure à trois filières; Vaud qui obtient des résultats moyens à l'enquête et Fribourg qui par contre a la meilleure performance moyenne dans l'ensemble des domaines testés par PISA. On obtient des recouvrements différenciés entre les filières scolaires si l'on prend en compte chaque domaine testé ou si l'on considère simultanément les mathématiques et la lecture.

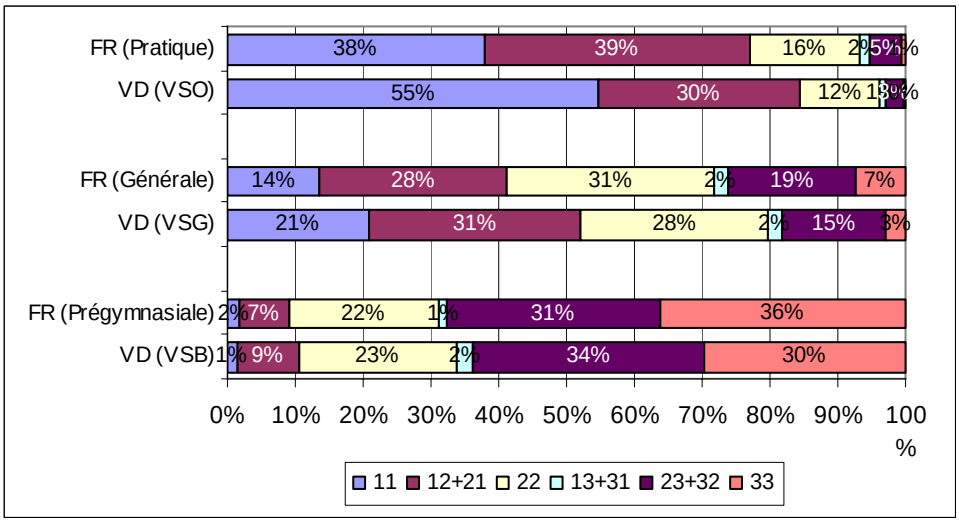


Figure 3 - Vaud et Fribourg: Profils des élèves en mathématiques et en lecture selon les trois filière cantonales (2003)

Les profils ont été établis selon les même règles que pour la figure 2. VD (VSO) désigne la filière la plus faible du canton de Vaud et FR (Pratique), la filière correspondante de Fribourg, VD (VSG) la filière intermédiaire de Vaud et FR (Générale) celle de Fribourg. Les filières regroupant les élèves les plus performants sont VD (VSB) pour Vaud et FR (Prégymnasiale) pour Fribourg. On constate que les superpositions observées entre profils selon les filières cantonales sont moins importantes que lorsque l'on compare la dispersion des résultats domaine par domaine. Pour le canton de Vaud, on notera en particulier que quasiment aucun élève de la filière vaudoise la plus faible (VD VSO) a un profil d'élève performant (profil 33), par contre plus de 50% des élèves ont un profil faible (profil 11). A l'opposé les élèves de la filière regroupant les élèves les plus performants (VD VSB) sont plus de 30% a avoir un profil fort (profil 33). Ainsi le recouvrement entre les filières est nettement moins important en prenant en compte les profils des élèves.

Pour le canton de Fribourg, on remarque également que le recours aux profils permet de distinguer les élèves des différentes filières. Cependant, on constate quelques différences par rapport aux profils obtenus pour les filières vaudoises. Ainsi, les élèves faibles (profil 11) de la filière la plus faible (FR pratique) sont moins nombreux (38%) que les élèves faibles vaudois (filière VSO) 55%. Il est à noter que les élèves de la filière fribourgeoise la plus faible (FR pratique) représentent un peu plus de 10% des élèves de 9e alors que les élèves vaudois de la filière correspondante (VD VSO) représentent plus de 30% des élèves vaudois. Par contre, la part des élèves performants (profil 33) est plus proche entre les filières regroupant les meilleurs élèves (VD VSB et FR Pratique), à savoir 36% pour Fribourg et 30% pour Vaud.

Discussion

Un des enjeux des enquêtes PISA est de fournir des informations fiables permettant de comparer les systèmes éducatifs entre eux sur de nombreuses dimensions. Les résultats sont destinés à un large cercle d'acteurs de l'éducation et nécessitent une communication accessible à un grand nombre de personnes. Ceci implique un travail de synthèse et de réduction de l'information qui parfois peut amener à ce que les résultats ne puissent pas prendre en compte des spécificités locales. Il est dès lors nécessaire de recourir à des analyses complémentaires ou d'utiliser des instruments différents afin de construire une information plus pertinente par rapport à la situation.

La lecture des premiers résultats romands mettait en évidence, par exemple, que Fribourg canton ayant des bons résultats avec une dispersion parmi les plus faibles et semblait donc être un système scolaire relativement équitable. Des analyses plus fines montrent que cela n'est toujours vrai. Les comparaisons des risques relatifs de certaines catégories d'élèves mettent en évidence des disparités entre les systèmes que les simples mesures de dispersion ne révèlent pas. Une autre statistique retenue pour la mesure de l'équité est le poids du niveau socioéconomique dans l'explication des performances. Les analyses menées sur les données romandes sur cet aspect montrent que le risque ne se distribue pas de façon linéaire et de la même façon dans les cantons romands. La construction de profils d'élèves permet de mieux cerner

l'hétérogénéité des performances des élèves en fonction des domaines testés. En particulier, les profils mettent en évidence la sélectivité des filières cantonales.

Les exploitations locales de PISA permettent une meilleure exploitation des données recueillies et une valorisation supplémentaire ayant une pertinence au plus près des acteurs locaux de l'éducation. Mais en contrepartie, elles montrent les limites d'interprétation de certains résultats d'un dispositif conçu pour une mesure à large échelle.

Ainsi, ces quelques exemples montrent que les résultats de PISA qui sont le plus souvent présentés et qui font l'objet des débats sur l'éducation sont des photos prises à une certaine distance et qu'elles ne permettent pas toujours une réelle prise en compte des situations à un niveau plus local, ceci est particulièrement vrai des systèmes éducatifs aux structures complexes et enchevêtrées comme ceux de la Suisse.

Références

Moreau, J. (2004). *Compétence et facteurs de réussite au terme de la scolarité. Analyse des données vaudoises de PISA 2000.* Lausanne: URSP

Nidegger, Ch. (coord.) (2005). *PISA 2003: Compétences des jeunes romands: Résultats de la seconde enquête PISA auprès des élèves de 9e année.* Neuchâtel : Institut de recherche et de documentation pédagogique.

OCDE (2004). *Apprendre aujourd'hui, réussir demain - premiers résultats de PISA 2003*. Paris : OCDE.

OFS/CDIP (2005) *PISA 2003: Compétences pour l'avenir – Deuxième rapport national*, Zahner, C. (éd.), Série monitorage de l'éducation en Suisse, Neuchâtel: OFS/CDIP.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

BASES DE DONNÉES INTERNATIONALES SUR LA QUALITÉ DE L'ÉDUCATION (1964/2005)

Nadir ALTINOK, IREDU-CNRS (Institut de Recherche sur l'Éducation)

Mots-clés : Qualité de l'éducation, Enquêtes Internationales, Évaluation.

Résumé :
Dans ce travail de recherche, nous utilisons une méthode qui nous permet d'obtenir des indicateurs qualitatifs du capital humain (IQCH), pour 105 pays, et pour chaque domaine de compétence (mathématiques, sciences, lecture) mesuré à travers l'ensemble des enquêtes internationales d'acquis des élèves. Cette méthode repose sur la possibilité d'un re-calibrage des résultats d'une enquête par rapport à une autre en ancrant les résultats sur les pays qui ont participé à au moins deux enquêtes différentes. Elle nous permet de construire des indicateurs de qualité du capital humain comparables pour un plus grand nombre de pays comparé à ce qui est actuellement disponible dans la littérature. Nous obtenons deux bases de données, l'une sous forme de coupe instantanée pour 106 pays, et l'autre sous forme de panel non cylindré de 1964 à 2005.
.

1. Introduction

Dans le champ de l'économie, nombreuses sont les études à tenter de mesurer l'impact de l'éducation sur la croissance économique. Les études les plus influentes tendent à conclure à une remise en cause d'un impact positif de l'éducation (Pritchett, 2001) Or, la mesure de l'éducation est pour le moins problématique : très souvent, les recherches prennent comme indicateur le taux brut de scolarisation ou encore le nombre moyen d'années scolaires. D'autres types d'études tentent d'estimer la fonction de production éducative, c'est-à-dire de relier des *inputs* éducatifs et financiers ("intrants") avec des *outputs* éducatifs *(*"sortants"*)*. Une fois de plus, les *outputs* éducatifs utilisés dans ces études sont trop limités : ils concernent la plupart du temps des indicateurs strictement quantitatifs de l'éducation. Est-il légitime de supposer qu'une année d'éducation dans un pays *i* est similaire à une année d'éducation dans un pays *j* ? Le rendement de l'éducation est-il similaire dans l'ensemble des pays ? Peut-on considérer l'éducation comme un bien homogène, à l'image du capital physique ? Dans cette contribution, nous partons de l'hypothèse de différences dans la qualité des systèmes éducatifs. Dans ce cas précis, nous utilisons les enquêtes internationales sur les acquis des élèves afin de mesurer cette qualité de l'éducation. Avec la construction de ces enquêtes internationales, nous obtenons des indicateurs qualitatifs du capital humain, pouvant donner une mesure alternative de l'éducation. Ainsi, cette nouvelle base de données peut apporter de nouveaux résultats dans l'estimation de la

contribution de l'éducation à la croissance ou encore l'estimation de la fonction de production éducative.

Au final, nous construisons deux bases de données sur la qualité de l'éducation, mesurée à partir des enquêtes internationales sur les acquis des élèves. La première base de données est en coupe transversale et fournit des indicateurs qualitatifs du capital humain (IQCH) pour 105 pays et déclinés pour trois domaines de compétence (mathématiques, sciences, lecture). La deuxième base de données est en panel et fournit des IQCH allant de 1964 à 2005 pour 105 pays et déclinés également pour trois domaines de compétences. Ces deux bases de données constituent une avancée dans la mesure où elles regroupent davantage de pays que les précédentes études.

2. Méthodologie générale

Nous avons utilisé 8 groupes d'enquêtes internationales d'analyse des compétences des élèves. Notre étude procède à un regroupement de ces enquêtes sur des échelles communes. La méthode utilisée se base sur la prise en compte des pays qui participent à plusieurs enquêtes simultanément et par le biais de leurs résultats, nous procédons à un re-calibrage des enquêtes les unes avec les autres. A l'issue de ces procédures, nous obtenons des indicateurs qualitatifs du capital humain (IQCH) pour trois domaines de compétences (mathématiques, sciences, lecture). Cette méthodologie diffère quelque peu selon que l'on travaille en coupe instantanée ou en panel.

Les indicateurs qualitatifs du capital humain (IQCH) peuvent être considérées comme une alternative aux variables strictement quantitatives de l'éducation, comme les taux de scolarisation. Les études de Hanushek et Kimko (2000) et Lee et Barro (2001) ont déjà entrepris ce genre de prise en compte. Notre étude consiste à poursuivre l'analyse en intégrant de nouveaux pays et de nouvelles données. En effet, les précédentes études n'ont concerné que la période 1964-1990.

Il s'agit à proprement parler de *quantifier* sur une échelle de 0 à 100 la *qualité* de l'éducation, plus précisément les taux de réussites d'échantillons représentatifs d'élèves de divers pays à des enquêtes internationales sur les acquis des élèves. Nous prenons en compte 8 différentes enquêtes internationales sur les acquis des élèves.

Nous obtenons deux type de bases de données. La première est en coupe instantanée et relate la qualité de l'éducation pour 105 pays. Cette base de données sépare la qualité de l'éducation en trois domaines de compétence (mathématiques, sciences, lecture). Elle reflète les disparités en terme de qualité de l'éducation pour un grand nombre de pays, et ce pour l'année la plus récente (entre 1998 et 2005). Afin d'obtenir cette base de données, nous avons recherché les pays qui avaient participé à au moins deux enquêtes différentes, afin d'obtenir une comparabilité entre les enquêtes. Nous avons toujours choisi les enquêtes de l'IEA comme enquête de référence, étant donné que ces enquêtes recouvrent le plus grand nombre de pays, que le niveau économique de ceux-ci est le plus hétérogène et que ces enquêtes sont réputées pour leur qualité d'échantillonnage. Nous avons mis à une échelle de 0-100 toutes les enquêtes, en supposant que le pays qui avait obtenu le plus grand score

aurait un résultat de 100. Ensuite, nous ajustons les enquêtes secondaires par rapport aux enquêtes de référence, en multipliant les premières enquêtes par un indice de passage calculé à partir des pays qui ont participé aux deux enquêtes. Ainsi, nous obtenons des indicateurs qualitatifs du capital humain calculés sur une même échelle et donc comparables.

La seconde base de données est en forme de panel et s'étend entre 1964 et 2005. Nous procédons à une compilation de l'ensemble des enquêtes relatives à la mesure des acquis des élèves au niveau primaire et secondaire. La méthodologie suivie reprend partiellement celle effectuée en coupe instantanée. Nous avons deux groupes d'enquêtes : celles où les Etats-Unis ont participé (série A) et qui permettent un ancrage avec une enquête spécifique (NAEP) et celles où les Etats-Unis n'ont pas participé (série B) et où nous utilisons principalement la méthode en coupe instantanée. Pour la première série d'enquêtes A, nous utilisons un ancrage sur une enquête américaine NAEP (National Assessment of Educational Progress) comme cela a été fait dans Hanushek et Kimko (2000). Le NAEP a été le principal instrument de mesure des acquis des élèves américains depuis 1969. L'IAEP est l'équivalent international du NAEP. Ainsi, la procédure d'évaluation est basée sur les résultats américains. A des périodes différentes depuis 1970, les élèves des Etats-Unis âgés de 9, 13 et 17 ans ont été questionnés sur leur acquis en sciences et en mathématiques. Ces tests peuvent donner une mesure de référence absolue pour le niveau des acquis des Etats-Unis. Afin de collecter à la fois les données des enquêtes IEA et IAEP, Hanushek et Kimko ont utilisé les résultats des Etats-Unis comme doublons. Ils ont ainsi modifié la moyenne des enquêtes de l'IEA afin de l'égaliser avec celles de l'IEAP qui leur étaient les plus proches (en terme d'âge, d'année et de domaine de compétence). A la différence de Hanushek et Kimko, afin d'obtenir des indicateurs comparables avec ceux obtenus du groupe B, nous n'avons pas repondéré les scores par les erreurs de mesure. Pour le deuxième groupe d'enquêtes B – celles où les Etats-Unis n'ont pas participé – nous avons utilisé la méthodologie présentée en coupe instantanée, à la différence que nous ne procédons pas à un ajustement à une échelle de 0 à 100 afin d'avoir des enquêtes comparables avec la série d'enquêtes A.

Au final, nous obtenons 56 séries d'enquêtes pour tous les groupes d'âge (9, 10, 13, 14, 15 et dernière année du secondaire). Afin d'avoir des données comparables dans le temps et des variables éducatives correspondantes, nous n'avons pas pris en compte les séries d'enquêtes pour les élèves dans leur dernière année au secondaire et les pré-tests pour l'enquête PASEC, ce qui réduit le nombre d'enquêtes à 42 séries. Dans une dernière étape, étant donné que des séries concernent la même année et le même niveau d'études (primaire ou secondaire), le regroupement de celles-ci conduit au final à 26 séries d'enquêtes.

L'ensemble des enquêtes utilisées et les informations principales relatives à celles-ci sont résumées dans les tableaux 1 et 2. Le premier tableau relate les enquêtes mobilisées afin d'obtenir des IQCH pour l'année la plus récente. Le tableau 2 montre les enquêtes utilisées pour obtenir la base de données des IQCH entre 1964 et 2005.

Tableau 1. Présentation des enquêtes internationales sur les acquis des élèves utilisées en coupe instantanée

Enquête Numéro	*Organisme-Enquête*	*Année(s) de référence*	*Nombre de pays*	*Domaine testé*	*Niveau économique des pays enquêtés*
1	IEA-TIMSS	1995, 1999, 2003	64	Mathématiques, sciences	Pays développés, en développement
2	IEA-PIRLS	2001	35	Lecture	Pays développés, en développement
3	OCDE-PISA	2003	41	Mathématiques, sciences, lecture	Pays développés, en développement
4	UNESCO-LABORATORIO	1997	13	Mathématiques, lecture	Pays d'Amérique Latine
5	CONFEMEN-PASEC	1995-2005	10	Mathématiques, lecture	Pays d'Afrique francophone
6	UNESCO-SACMEQ	1999, 2002	14	Mathématiques, lecture	Pays d'Afrique anglophone
7	UNESCO-MLA	1992-1997	11	Mathématiques, sciences, lecture	Pays d'Afrique

Abréviations: IEA (International Association of the Evaluation of Educational Achievement), TIMSS (Third International Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study), OECD (Organization for Economic Co-operation and Development), PISA (Programme for International Student Assessment), UNESCO (United Nations Educational, Scientific and Cultural Organization), CONFEMEN (Conference of Francophone Education Ministers), PASEC (Programme on Analysis of Education Systems), SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality), MLA (Monitoring Learning Achievement).

Tableau 2. Présentation des enquêtes internationales sur les acquis des élèves utilisées en panel

Enquête Numéro	*Organisme-Enquête*	*Année(s) de référence*	*Domaine testé*	*Nombres d'enquêtes utilisées*
1	USA-IAEP	1988, 1990	Mathématiques, sciences	6
2	IEA-TIMSS	1964, 1970, 1982, 1984, 1995, 1999, 2003	Mathématiques, sciences	16
3	IEA-PIRLS	1970, 1990, 2001	Lecture	4
4	OCDE-PISA	2000, 2003	Mathématiques, sciences, lecture	6
5	UNESCO-LABORATORIO	1997	Mathématiques, lecture	2
6	CONFEMEN-PASEC	1995-2005	Mathématiques, lecture	2
7	UNESCO-SACMEQ	1999, 2002	Mathématiques, lecture	3
8	UNESCO-MLA	1992-1997	Mathématiques, sciences, lecture	3

Abréviations: USA (Etats-Unis), IAEP (International Assessment of Educational Progress), pour les autres abréviations, voir le tableau 1.

3. Principaux résultats

Cette section a pour objectif de dresser les indicateurs statistiques usuels concernant les indicateurs qualitatifs du capital humain (IQCH) et de procéder à des comparaisons statistiques avec des variables éducatives. Nous utilisons dans cette section les éléments de la base de données en coupe transversale : les IQCH concernent alors les indices de qualité les plus récents dont nous disposons. Les indicateurs qualitatifs du capital humain obtenus ont été regroupés dans un même indice, en effectuant la moyenne arithmétique des résultats dans les trois domaines (lorsqu'ils sont disponibles) : nous obtenons alors les IQCH-Général. Le nombre des IQCH-Général est légèrement plus élevé que ceux des IQCH spécifiques du fait que les résultats pour chaque domaine de compétence ne comprennent pas le même échantillon de pays. Par ailleurs, il est logique de trouver une valeur maximum différente de 100 car si un pays a obtenu 100 pour un domaine de compétence, il peut tout à fait avoir obtenu un score différent dans un autre domaine (exemple du Singapour).

En premier lieu, nous dressons les indices d'analyse usuels pour chacune d'entre elles (tableau 3). On constate que les résultats en mathématiques sont les plus nombreux (104 pays), et c'est pour la mesure des IQCH en sciences que nous avons le moins de pays (79). L'analyse des moyennes montre que la moyenne la moins élevée est pour les mathématiques (69 % de réussite), tandis que pour les sciences on observe

une moyenne assez forte (81 % de réussite), la lecture se situant dans une position intermédiaire (72 % de réussite).

Tableau 3. Analyse statistique générale des IQCH

Variable	*Observations*	*Moyenne*	*Ecart-type*	*Minimum*	*Maximum*
IQCH-Général	105	72,41	17,32	23,20	97,74
IQCH-Mathématiques	104	69,80	16,58	22,69	100
IQCH-Sciences	79	80,71	13,03	42,21	100
IQCH-Lecture	89	72,24	20,39	23,71	100

Dans un deuxième temps, la corrélation entre les différentes mesures des compétences ont été calculées (tableau 4). Les résultats laissent supposer à une corrélation positive entre les IQCH en mathématiques, sciences et lecture : ainsi, plus un pays a un niveau de compétences élevé en mathématiques, plus ses niveaux de compétence en lecture et en sciences seront élevés (les coefficients de détermination sont de 0,90 pour sciences et 0,95 pour lecture). Remarquons que le modèle reliant la lecture aux mathématiques est plus explicatif que celui qui relie ces dernières aux sciences, ce qui peut en partie s'expliquer par la différence du nombre d'observations dans chacun des tests (88 observations en mathématiques contre 79 pour sciences).

Tableau 4. Matrice de corrélation des indicateurs de qualité du capital humain

	IQCH	*IQCH-M*	*IQCH-S*	*IQCH-L*
IQCH	1			
IQCH-M	0,98	1		
IQCH-S	0,95	0,90	1	
IQCH-L	0,98	0,95	0,85	1

Note : Le nombre d'observations varie entre 64 et 104. IQCH-M = indicateur qualitatif du capital humain en mathématiques ; IQCH-S : indicateur qualitatif du capital humain en sciences ; IQCH-L : indicateur qualitatif du capital humain en lecture.

Le tableau 5 présente la variation de l'IQCH selon les grandes régions du monde. Nous avons repris le découpage effectué par la Banque mondiale. On constate de fortes disparités entre ces régions : globalement, les régions les plus dotées en IQCH sont les pays d'Europe et ceux de l'Asie de l'Est et du Pacifique. Dans une position intermédiaire, les pays d'Amérique Latine, du Moyen Orient et d'Afrique du Nord ont des résultats avoisinant 70 % de réussite aux tests internationaux. L'Afrique subsaharienne obtient les moins bons résultats avec une moyenne de 50 % aux IQCH. Le pays de ce continent qui a obtenu le maximum aux tests (i.e. Madagascar) réalise ainsi environ le même score que le pays ayant obtenu le score le plus bas dans les pays de l'Asie de l'est et Pacifique (i.e. la Chine).

Tableau 5. Variation de l'IQCH selon les régions du monde

	Pays	***Moyenne***	***Ecart-type***	***Minimum***	***Maximum***
Afrique subsaharienne (ASS)	27	49,13	08,76	23,20	65,20
Asie de l'est et Pacifique (AEP)	13	85,50	12,68	63,27	97,74
Europe de l'est et Asie centrale (EAC)	16	84,25	05,33	75,01	91,64
Amériques et Caraïbes (AMC)	14	70,80	11,24	59,03	91,43
Moyen Orient et Afrique du Nord (MOAN)	12	68,39	06,85	56,15	84,43
Asie du Sud	0				
Europe continentale (EUC)	23	87,16	03,94	78,83	94,66
Total	**105**	**72,65**	**17,40**	**23,20**	**98,34**

Le tableau 6 indique la variation de l'IQCH selon le niveau économique des pays. Nous avons repris le découpage selon la Banque mondiale. On constate une forte disparité selon ce critère : plus le pays est doté d'un revenu élevé, plus son score le sera également. Ainsi, toute chose étant égale par ailleurs, les pays à revenus faibles (PRF) vont des scores 35 points de pourcentage inférieurs à ceux des pays à revenus intermédiaires (PRI) et inférieurs d'environ 64 points de pourcentage par rapport aux pays à revenus élevés (PRE). Les disparités de qualité du capital humain suivent ainsi les disparités au niveau économique. On remarque également que les disparités à l'intérieur de ces groupes de pays sont plus fortes pour des pays à revenu faible ou intermédiaire que pour les pays à revenu élevé. Par là même, les disparités de capital humain augmenteraient tout d'abord avec le niveau économique (passant de 10,96 à 13,31) pour ensuite connaître une homogénéisation assez forte lorsque le niveau économique des pays est élevé (passant de 13,31 à 6,70).

Tableau 6. Variation de l'IQCH selon le niveau économique des pays

	Nombre de Pays	***Moyenne***	***Ecart-type***	***Minimum***	***Maximum***
Pays à revenus faibles (PRF)	22	50,03	11,43	23,20	80,56
Pays à revenus intermédiaires (PRI)	49	72,11	13,26	42,68	97,74
Pays à revenus élevés (PRE)	34	87,31	06,60	64,88	96,71
Total	**103**	**72,71**	**16,90**	**23,2**	**97,74**

Dans le tableau 7, on s'intéresse au cas des pays les plus performants et à ceux les moins performants selon les grandes régions. En Afrique subsaharienne, on ne trouve pas de distinction selon la langue officielle quant à la place des pays les plus performants (Madagascar, Cameroun, Île Maurice). La Mauritanie et la Namibie se révèlent être les pays les moins performants dans ce continent. Concernant l'Asie du sud et de l'est, on retrouve dans les pays ayant un fort IQCH les "tigres asiatiques" (Taiwan, Singapour, Corée du Sud, Hong Kong), ce qui pourrait notamment expliquer leur forte croissance. Les résultats élevés de certains pays de l'Europe de l'Est (Estonie, République Tchèque) peuvent notamment expliquer leur niveau élevé dans les technologies et leur rapport au savoir, leur permettant un rattrapage rapide des pays de l'Europe continentale. Enfin, les résultats concernant l'Europe continentale rejoignent ceux de PISA 2003 en observant une présence des pays nordiques (Finlande, Pays-Bas) dans les pays les plus performants.

Tableau 7. Analyse de la performance en terme de qualité du capital humain selon les régions géographiques du monde

Régions	***ASS***	***AEP***	***EAC***	***AMC***	***MOAN***	***EUC***
Pays Les Plus Performants						
	Madagascar	Taiwan	Estonie	Canada	Israël	Finlande
	Cameroun	Singapour	Rép. Tchèque	Cuba	Iran	Liechtenstein
	Île Maurice	Corée du Sud	Hongrie	Etats-Unis	Bahreïn	Pays-Bas
	Ouganda	Hong Kong	Lituanie	Uruguay	Egypte	Belgique (Flandre)
Pays Les Moins Performants						
	Mauritanie	Chine	Macédoine	Rép. Dominic.	Rép. d'Oman	Chypre
	Bénin	Philippines	Turquie	Honduras	Arabie Saoudite	Belgique (Franco.)
	Sénégal	Indonésie	Serbie	Venezuela	Maroc	Portugal
	Namibie	Thaïlande	Arménie	Bolivie	Koweït	Grèce

Note : ASS = Afrique subsaharienne ; AEP = Asie de l'Est et du Pacifique ; EAC = Europe de l'Est et Asie Centrale ; AMC = Amériques et Caraïbes ; MOAN = Moyen Orient et Afrique du Nord ; EUC = Europe continentale.

4. Conclusion

A travers toutes les explications de la croissance économique des pays, une d'entre elles qui est le plus souvent acceptée est le niveau de capital humain. Cette évidence a toutefois rencontré de nombreuses incohérences dans la littérature existante. Les approches macro-économiques les plus robustes apportent ainsi des contradictions dans la relation éducation-croissance. Pritchett (2001) montre en effet que très souvent l'impact de l'éducation sur la croissance est négatif et de manière significative. Pour autant, la plupart des études ont ignoré le caractère qualitatif du capital humain, en ne s'appuyant que sur des indicateurs purement quantitatifs.

Or, la prise en compte des enquêtes nationales ou internationales sur les acquis des élèves peut permettre de combler ce manque de mesure qualitative. Hanushek et Kimko (2000) et Lee et Barro (2001) ont notamment pris ce chemin qualitatif, mais sans exploiter l'ensemble des enquêtes internationales et pour un échantillon très réduit de pays. Dans ce travail de recherche, nous avons utilisé une méthode qui nous a permis d'avoir des indicateurs qualitatifs du capital humain (IQCH) pour environ 105 pays, et pour chaque domaine de compétence (mathématique, sciences, lecture).

Deux bases de données sont obtenues : l'une en coupe transversale qui relate la qualité de l'éducation pour 105 pays et pour l'année la plus récente ; et l'autre en panel allant de 1965 à 2005.

Il resterait à présent à estimer à partir de ces bases de données, l'impact de l'éducation sur la croissance ou encore la fonction de production éducative. Ces nouveaux indicateurs permettent-ils d'améliorer notre compréhension de l'éducation ? Notamment sous sa dimension qualitative ?

Références

Hanushek, E.A. & Kimko, D.D. (2000). Schooling, Labor-Force Quality, and the Growth of Nations, *American Economic Review*, 90(5), 1184-1208.

Lee, J.W. & Barro, R.J (2001). Schooling Quality in a Cross Section of Countries, *Economica*, 38(272), 465-488.

Pritchett, L. (2001). Where has all the education gone ?, *World Bank Economic Review*, 15, 367-391.

ROLE DU FACTEUR LINGUISTIQUE DANS LES PRATIQUES PEDAGOGIQUES RAPPORTÉES AU PROGRAMME INTERNATIONNAL DE RECHERCHE EN LECTURE SCOLAIRE (PIRLS)

Marielle Simon, Catherine Turcotte, Université d'Ottawa

Mots-clés : évaluation de la lecture, analyses secondaires, évaluations comparatives, pratiques pédagogiques en lecture

Résumé :

En 2001, des élèves de la quatrième année dans trente-cinq pays ont participé au *Programme international de recherche en lecture scolaire* (PIRLS). La Suède, la Bulgarie, les Pays-bas et l'Angleterre se situaient au premier rang du palmarès des résultats. Ces pays obtenaient un score moyen significativement supérieur à la France et, au Canada, l'Ontario (majoritairement anglophone) performait significativement mieux que le Québec (majoritairement francophone). De même, ailleurs au Canada, les élèves francophones obtenaient un rendement significativement inférieur à celui des élèves de langue anglaise. Afin de tenter d'expliquer cet écart, nous avons comparé les données recueillies par le questionnaire contextuel à l'intention des enseignants lors du PIRLS 2001. L'étude se donnait comme objectif de comparer les pratiques d'enseignement et d'évaluation de trois populations d'enseignants : ceux du Québec français, ceux de l'Ontario français et ceux de l'Ontario anglais. Cette communication présente les résultats des analyses secondaires d'une trentaine de questions choisies sur une base théorique. Les résultats des analyses de correspondances multiples révèlent des différences marquées dans les pratiques pédagogiques des enseignants, les plus grandes s'observant entre les pratiques des enseignants anglophones et francophones. Pour la majorité des variables, les enseignants des deux populations francophones montrent des pratiques semblables, sauf en ce qui a trait au temps consacré à l'enseignement de la lecture de façon hebdomadaire. La discussion porte sur les implications de l'étude aux plans de la recherche, de la pratique et de la conception des enquêtes en lecture.

Contexte

En 2001, des élèves de la quatrième année de trente-cinq pays participaient au *Programme international de recherche en lecture scolaire* (PIRLS). La Suède se situait au premier rang du palmarès des résultats. Chez les quelques populations francophones présentes à ce programme, les élèves obtenaient un rendement inférieur à celles des pays anglophones. Par exemple, la France se situait au 18e rang comparativement à l'Angleterre (3e rang), aux Etats-Unis (9e rang), à la Nouvelle Zélande (13e rang) et à l'Écosse (14e rang). Au Canada, les élèves du Québec, majoritairement francophones, affichaient également un rendement inférieur à celui des élèves de la province de l'Ontario, une province majoritairement anglophone (Bussière, Cartwright, Crocker, Ma, Oderkirk, & Shang, 2001; Ministère de l'éducation du Québec, 2003). De plus, en Ontario, les élèves des écoles de langue française recevaient aussi un rendement significativement inférieur à celui des élèves des écoles de langue anglaise. Faut-il se questionner sur l'écart entre le rendement en lecture des élèves des populations francophones et anglophones et sur les facteurs qui l'expliquent ? Afin de tenter de répondre à cette question, la présente étude compare les pratiques pédagogiques et évaluatives des enseignants des trois populations suivantes: a) enseignants des écoles de langue anglais de l'Ontario, b) enseignants des écoles de langue française de l'Ontario et c) enseignants des écoles de langue française du Québec. Cette comparaison s'effectue à l'aide des données recueillies à l'aide du questionnaire à l'intention des enseignants administré lors du PIRLS 2001.

Recension des écrits

Plusieurs études ont démontré que les pratiques des enseignants déterminent de manière significative l'apprentissage de la lecture (Block, Oakar, & Hurt, 2002 ; Dreher, 2003 ; Juel & Minden-Cupp, 1999 ; Morrow, Tracey, GeeWoo, & Pressley, 1999 ; Taylor, Pearson, Clark, & Walpole, 2000). Selon leurs résultats, la différenciation de l'enseignement (Gambrell & Mazzoni, 1999), un environnement littéraire riche (Cole, 2003) et faisant appel à plusieurs types de textes variés (Ivey &

Broaddus, 2001) et la priorité accordée à l'enseignement de la lecture en classe (Giasson, 1994; 2003) comptent parmi les pratiques exemplaires dans la réussite des élèves. En ce qui concerne l'évaluation de la lecture, les approches dites *alternatives* favorisent l'apprentissage de la lecture chez les élèves. Ces approches comprennent par exemple le questionnement oral, le questionnement écrit à réponses élaborées, les questions faisant appel à des stratégies cognitives de haut niveau, les mimes et l'entretien individuel avec l'élève. Perçues comme étant signifiantes aux yeux des élèves (Allington, Johnston, & Day, 2002), ces stratégies leur donnent l'occasion de réfléchir sur leurs apprentissages (Farr, 2004; Guterman, 2002) et permettent à l'enseignant de vérifier le niveau d'analyse, d'inférence et de prédiction. (Taylor, Pearson, Peterson, & Rodriguz, 2003). Par ailleurs, les approches plutôt *traditionnelles*, comme le questionnement écrit à choix multiples ou à réponses courtes, se limitent souvent à la mesure du rappel d'information et privilégient les tâches fermées plutôt qu'ouvertes (Shepard, 2000).

Cadre de référence

Les renseignements issus de la recension des écrits dans le domaine ont permis d'identifier les pratiques pédagogiques et d'évaluation potentiellement déterminantes sur le rendement des élèves en lecture. Bref, en ce qui concerne les pratiques pédagogiques, celles-ci comprennent les quatre catégories suivantes : 1- différenciation de l'enseignement, 2- temps accordé à l'enseignement de la lecture, 3- ressources et 4- types de textes (courants et littéraires). Les variables entourant les pratiques d'évaluation regroupent les approches traditionnelles (p.ex, questionnement écrit à réponses courtes) versus celles alternatives (p.ex., portfolio, réponses élaborées).

Méthodologie

L'échantillon comprend les enseignants des écoles de langue française de l'Ontario (n = 80), ceux des écoles de langue anglaise de l'Ontario (n = 115) et ceux des écoles de langue française du Québec (n = 100) ayant participé au PIRLS 2001. Il faut noter qu'aux fins de l'enquête, le processus d'échantillonnage s'est fait auprès des élèves et non auprès des enseignants. Les items proviennent du questionnaire contextuel à l'intention des enseignants. Initialement, 42 questions ont été retenues selon les quatre catégories des pratiques pédagogiques du cadre conceptuel et selon les deux groupes de pratiques d'évaluation. L'application d'analyses factorielles exploratoires et de mesures d'association (Cramer's V) aux données des trois populations donné lieu à la sélection de 22 questions réparties selon les cinq catégories suivantes : différenciation de l'enseignement, ressources, temps consacré à la lecture, fréquence d'utilisation d'évaluations alternatives et importance accordée aux épreuves formelles. Par la suite, des analyses de correspondances multiples ont contribué à la mise en relation de ces variables entre elles lors de la comparaison des pratiques des trois populations.

Résultats

Les résultats des trois analyses révèlent des résultats semblables entre elles. Le tableau en annexe présente les résultats des analyses descriptives des pratiques pédagogiques et évaluatives des enseignants des trois populations. Trois grandes observations se dégagent du tableau. Premièrement, les enseignants des écoles de langue anglaise de l'Ontario privilégient davantage la différenciation de l'enseignement que ceux des écoles de langue française de l'Ontario et du Québec. En d'autres mots, un plus grand nombre d'enseignants des écoles de langue anglaise favoriseraient l'enseignement en petits groupes et proposeraient du matériel de lecture adapté aux besoins des élèves. Deuxièmement, une plus grande proportion d'enseignants des écoles de langue française de l'Ontario et du Québec,

comparativement à leurs homologues anglophones de l'Ontario, disent employer des manuels scolaires ou des textes séquentiels quotidiennement. Inversement, les enseignants des écoles de langue anglaise de l'Ontario auraient tendance à utiliser davantage une diversité de textes et de ressources variées comme des romans, des récits, des livres d'histoire, des manuels instructifs ou encore des descriptions explicatives. Enfin, le temps accordé à l'enseignement de la langue, de la lecture ou à des activités de lecture diffère significativement d'une population à l'autre. Les enseignants francophones du Québec attribuent le plus de temps aux activités d'enseignement du français tandis que les enseignants francophones de l'Ontario réservent le moins d'heures à l'apprentissage hebdomadaire du français ou de la lecture. Par ailleurs, les enseignants anglophones de l'Ontario constituent la plus importante population accordant plus de cinq heures par semaine à l'enseignement spécifique de la lecture.

Au plan de l'évaluation, les pratiques des enseignants des trois populations diffèrent également. Les enseignants anglophones privilégient davantage les approches alternatives d'évaluation. En effet, ils utilisent davantage le portfolio, assignent plus régulièrement des tâches à réponses élaborées et ont plus souvent recourt aux rencontres avec les élèves pour discuter de leurs lectures et leur travail accompli. En revanche, les enseignants des écoles de langue française, tant de l'Ontario et que du Québec, accordent une plus grande importance, que les enseignants des écoles anglophones de l'Ontario, aux épreuves formelles. En effet, plus de la moitié d'entre eux disent accorder beaucoup d'importance aux tests en classe, aux épreuves diagnostiques et aux enquêtes à grande échelle.

Discussion

Au Canada, la responsabilité de l'éducation revient à la province. Pourtant, les pratiques des enseignants francophones de l'Ontario semblent rejoindre davantage celles des enseignants du Québec que celles des enseignants de l'Ontario. L'étude suscite donc plusieurs questions : Pourquoi les populations des écoles de langue française du Québec et de l'Ontario adoptent-elles des pratiques semblables malgré un programme de formation et un curriculum distincts ? Pourquoi les pratiques pédagogiques de ces deux populations francophones tiennent-elles peu compte des

approches de lecture préconisées dans les écrits ? Y aurait-il un effet culturel ou linguistique biaisant les résultats du PIRLS ?

En tentant de répondre à ces questions, il faut d'abord considérer certaines conditions relatives aux analyses. D'abord, les valeurs V de Cramer obtenues dans l'étude sont relativement faibles (valeur maximale = 0,37), de même que les taux de variances expliqués par les dimensions des analyses de correspondances multiples (valeur maximale = 25,18 %). De plus, lorsqu'il manque jusqu'à 40 % de données dans des échantillons à faibles effectifs, le processus d'imputation s'applique difficilement sans nuire à la validité des résultats. Les analyses se sont donc basées uniquement sur les données brutes complètes.

Le PIRLS présente également des limites au plan du nombre, du contenu et de la formulation questions et de leurs échelles de fréquence ou d'importance. Par exemple, le questionnaire contextuel ne compte que trois items portant sur l'évaluation, ce qui limite les thèmes potentiels à analyser. Il ne sonde aucunement les enseignants, entre autres, au sujet du type de textes choisis aux fins d'évaluation ou au sujet du niveau de participation des élèves aux décisions relatives aux pratiques d'évaluation. Le contenu de certaines des questions laisse également à désirer. Comment les enseignants interprètent-ils, par exemple, les deux options suivantes à la question 14 : « j'utilise les mêmes documents avec tous les élèves parce que tous sont du même niveau » et « j'utilise les mêmes documents avec tous les élèves peu importe leur niveau en lecture, et je les fais travailler au même rythme » ? Que veut dire pour eux le terme « niveau » ? Finalement, le choix et la formulation des échelles occasionnent aussi quelques difficultés. Comment les enseignants des trois populations réagissent-ils face aux choix suivants « peu ou pas d'importance » et « un peu d'importance » en réponse à certaines questions d'attitudes ?

Qu'il s'agisse d'un problème de traduction, d'adaptation, d'interprétation ou de pertinence, ces difficultés se multiplient sans doute sous l'effet des différences culturelles, linguistiques et sociologiques des populations comparées. Il y aurait donc lieu de questionner en personne les enseignants des diverses populations linguistiques afin de mieux comprendre leur interprétation du questionnement PIRLS et leur réalité scolaire. Par exemple, le fait que plusieurs enseignants du Québec choisissent d'occuper des postes dans les écoles de langue française en Ontario peut expliquer en

partie les résultats obtenus dans l'étude et les questionnaires contextuels ne posent pas de telles questions. Aussi faut-il comparer des données des autres populations francophones et anglophones qui participent au PIRLS et vérifier dans quelle mesure les écarts et les tendances s'observent également dans les autres enquêtes internationales comme la Troisième enquête en mathématiques et en sciences (TEIMS). Malgré les limites inhérentes à ce type d'analyse, la présente recherche sensibilise les chercheurs et les praticiens à la présence de certaines tendances en ce qui concerne les pratiques pédagogiques en lecture et aux difficultés reliées aux enquêtes comparatives du rendement des élèves. Ce ne sont que quelques pistes de recherche qui méritent une exploration en profondeur.

Références

Allington, R. L., Johnston, P. H. & Day, J. P. (2002). Exemplary fourth-grade teachers. *Language Arts,79*, 462-466.

Block, C.C., Oakar, M.& Hurt, N. (2002). The expertise of literacy teachers : A continuum from preschool to grade 5. *Reading Research Quarterly, 37*, 178-206.

Bussière, P., Cartwright, F., Crocker, R., Ma, X., Oderkirk, J. & Zhang, Y. (2001). *À la hauteur : La performance des jeunes du Canada en lecture, en mathématiques et en sciences : Étude PISA de l'OCDE — Premiers résultats pour les Canadiens de 15 ans*. Ottawa, ON : Ministre de l'Industrie.

Cole, J.E. (2003). What motivates students to read ? Four literacy personalities. *The Reading Teacher, 56*,326-36.

Dreher, M. J. (2003). Motivating teachers to read. *The Reading Teacher, 56* ,338-40.

Farr, R. (2004). Putting it all together: Solving the reading assessment puzzle. In R.B. Ruddell & N.J. Unrau (Eds.), *Theorical models and processes of reading, fifth edition*. (pp.4.4). Newark, DE: International Reading Association.

Gambrell, L.B. & Mazzoni, S.A. (1999). Principles of best practices : finding the common ground. In L.B. Gambrell, L.M. Morrow, S.B. Neuman, & M. Pressley (Eds.), *Best practices in literacy instruction*. (pp.11-21). New York, NY: The Guilford Press.

Giasson, J. (1994). Lire avant tout. In L. Saint-Laurent, J. Giasson, C. Simard, J.J. Dionne et E. Royer, et coll. *Programme d'intervention auprès des élèves à risque. Une nouvelle option éducative*. (pp.81-91). Boucherville, Québec : Gaëtan Morin Editeur.

Giasson, J. (2003). *La lecture : de la théorie à la pratique*. 2ième édition. Boucherville, Québec : Gaétan Morin Éditeur.

Guterman, E.(2002). Toward dynamic assessment of reading : applying metacognitive awareness guidance to reading assessment tasks. *Journal of Research in Reading, 25*, 283-298.

Ivey G. & Broaddus, K. (2001). "Just plain reading": A survey of what makes students want to read in middle school classrooms. *Reading Research Quarterly, 36*, 350-377.

Juel, C. & Minden-Cupp, C. (1999). *Learning to read words. Linguistic units and strategies* CIERA report. University of Michigan : Ann Arbor.

Ministère de l'Éducation du Québec. (2003). *Le programme international de recherche en lecture scolaire (PIRLS) 2001: Résultats obtenus par les élèves Québécois de dix ans*. Québec: Direction de la sanction des études.

Morrow, L. M., Tracey, D. H., Woo, D. & Pressley, M. (1999). Characteristics of exemplary first-grade literacy instruction. *The Reading Teacher, 52*, 462-476.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher,29*,4-14.

Taylor, B. M., Pearson, D. P., Clark, K. F. & Walpole, S. (2000). Effective schools and accomplished teachers : Lessons about primary-Grade reading instruction in low-income schools. *The Elementary School Journal, 101*, 121-165.

Taylor, B.M., Pearson, P. D., Peterson, D. S. & Rodriguez, M.C. (2003). Reading growth in high-poverty classrooms : the influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal, 104*, 3-27.

Variables et leurs questions	O.Ang.	O.Fran.	Q.Fran.	V Cramer
Différenciation				
(11A) J'enseigne la lecture à toute la classe (toujours ou presque)	21%	**45%**	35%	.251
(11B) Je crée des groupes de même niveau (souvent)	**30%**	15%	20%	.209
(14) J'utilise des documents différents, selon le niveau des élèves en lecture	**45%**	8%	4%	.372
Matériel utilisé				
(12A) J'utilise des manuels ou des textes séquentiels (tous les jours ou presque)	29%	**44%**	**65%**	.244
(12F) J'utilise divers ouvrages pour enfants (tous les jours ou presque)	**56%**	24%	22%	.272
(13C) J'utilise des ouvrages plus longs, divisés en chapitres (tous les jours ou presque)	**43%**	9%	9%	.391
(13B) J'utilise d'autres récits (une à deux fois par semaine)	**42%**	5%	27%	.357
(13 f) J'utilise des descriptions explicatives (non roman) (une à deux fois par semaine)	**47%**	22%	26%	.249
Temps consacré à l'enseignement de la lecture et de la langue par semaine				
(7A) Je consacre entre 8 heures et plus aux activités d'enseignement de la langue	43%	32 %	**69%**	.324
(9A) Je consacre 5 heures et plus aux activités d'enseignement de la lecture	**63%**	34%	**58%**	.329
Approche alternative de l'évaluation				
(27h) J'évalue la performance en lecture en rencontrant les élèves pour une discussion sur les lectures et le travail accompli (une ou deux fois par mois)	**54 %**	32%	27%	.280
(27c) Je pose des questions appelant des réponses d'un paragraphe (une fois par semaine et plus)	**43%**	20%	18%	.322
(28) Les portfolios représentent un moyen important de mon évaluation des élèves en lecture	30%	27%	21%	.274
Approche traditionnelle de l'évaluation				
(26a) J'accorde (beaucoup d'importance) aux tests diagnostiques	10%	**57%**	**45%**	.311
(26b) - aux tests en classe (beaucoup d'importance)	35%	**77%**	**79%**	.301
(26c) - aux épreuves provinciales ou régionales (un peu d'importance)	31%	**58%**	**46%**	.306

ANNEXE

Tableau 1 - Résultats descriptifs des pratiques pédagogiques et évaluatives des enseignants des trois populations (PIRLS)

DATA ENVELOPMENT ANALYSIS (DEA) : UNE NOUVELLE MÉTHODE D'ÉVALUATION DE LA PERFORMANCE SCOLAIRE.

Djily Diagne, Université de Genève

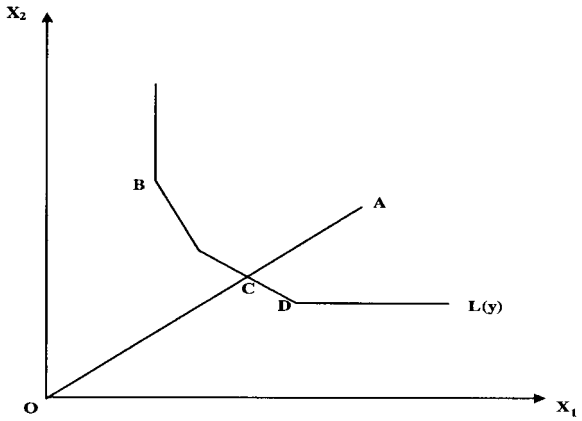
Résumé :
Cette communication a pour but de présenter la méthode d'évaluation de la performance scolaire connue sous l'acronyme DEA (Data Envelopment Analysis). Issue de la programmation linéaire, la méthode DEA a été développée par Charnes, Cooper et Rhodes en 1978. Elle est particulièrement idoine pour évaluer la performance dans des domaines complexes comme celui de l'éducation. La méthode DEA permet de déterminer la frontière efficiente de production ou frontière de *meilleures pratiques* d'un ensemble « d'unités de décision ou DMUs (écoles, facultés, départements, etc.). L'efficience est mesurée sous un angle empirique, à partir de données observées d'inputs (ressources mobilisées) et d'outputs (résultats réalisés), en tenant compte de l'environnement plus ou moins favorable dans lequel opère les DMUs. Les DMUs les plus performantes sont celles qui assurent le meilleur service avec des ressources minimales. Elles peuvent ainsi servir de référence aux autres DMUs et une analyse de leurs caractéristiques peut apporter à celles qui sont moins efficientes des informations susceptibles d'améliorer leur niveau d'efficience. Les résultats ainsi obtenus peuvent servir de point de départ à une procédure de benchmarking.

Présentation

Le graphique 1 illustre la construction par la méthode DEA de $L(Y)$ qui représente les combinaisons minimum d'input par unité d'output ou la frontière de production. Les établissements scolaires correspondant aux points B, C, et D qui se trouvent sur l'isoquant sont techniquement efficients. Par contre, l'établissement correspondant au point A ne l'est pas. Son inefficience est représentée par le segment CA correspondant aux proportions d'inputs qui pourraient être réduites sans diminution du montant de l'output. Il est en effet possible de produire le même niveau d'output avec une diminution de tous les inputs dans la proportion $\frac{CA}{OA}$. Ainsi, le degré de performance (efficience technique) de l'établissement A est mesuré par :$ET = \frac{OC}{OA}$ ($0 \leq ET \leq 1$). En d'autres termes, la frontière efficiente est constituée des unités affichant des scores égaux à 1 ; pour les autres DMUs, le score est compris entre 0 et 1.

La méthode DEA cherche donc à estimer la frontière de production par une courbe enveloppe formée des segments de droite joignant les entités efficientes (d'où sa dénomination : *data envelopment analysis*). D'un point de vue calcul, la construction de la frontière d'efficience repose sur la résolution, pour chaque établissement scolaire, d'un problème de programmation linéaire. En effet, le modèle CHARNES, COOPER et RHODES (1978) ou CCR se base sur la maximisation de la somme pondérée des outputs rapportée à la somme pondérée des inputs

Graphique1. Principe de la méthode DEA



Contrairement aux méthodes traditionnelles, la méthode DEA ne requiert aucune hypothèse a priori sur le processus de production des unités de décision. Cet avantage est important lorsqu'il s'agit de l'éducation car les résultats scolaires sont tributaires de nombreux facteurs (environnement familial des élèves, effets des pairs, motivation des élèves, etc.) dont la spécification n'est pas aisée du point de vue du calcul économique traditionnel. Un autre avantage de la méthode DEA est qu'elle permet facilement d'intégrer plusieurs inputs et outputs dans l'évaluation, ce qui permet de mieux tenir compte de la nature multidimensionnelle de l'éducation. On sait, en effet, que l'output du système éducatif est divers et ne se résume pas aux seuls résultats scolaires. L'analyse DEA se distingue d'une analyse de tendance centrale comme la technique de régression du fait que la frontière d'efficience est déterminée du point de vue de la meilleure pratique. Elle permet de ce fait de comparer chaque école à la plus performante de l'échantillon et non à une valeur moyenne.

Durant les deux dernières décennies, les applications de la méthode DEA au champ éducatif ont été nombreuses. Les premiers travaux ont eu lieu aux Etats-Unis : CHARNES, COOPER et RHODES (1978), BESSENT et al. (1982, 1984), FÄRE et al. (1989) et RAY (1991). Cette dernière étude est innovatrice car le traitement des variables socio-économiques dans les modèles DEA a été amélioré. Contrairement aux études précédentes, Ray propose en effet de ne tenir compte dans l'analyse DEA que des variables sous le contrôle des écoles. Ainsi, l'efficience de chaque école est calculée uniquement avec les inputs qu'elle contrôle. L'étape suivante consiste alors à régresser les taux d'efficience obtenus sur les facteurs d'environnement qui ne sont pas sous son contrôle. L'idée de l'auteur est d'isoler, pour les écoles inefficientes, l'effet des conditions externes, d'une part, et celui de la mauvaise gestion, d'autre part. Plusieurs auteurs ont par la suite adopté cette démarche en deux étapes. Par exemple, KIRJAVAINEN et LOIKKANEN (1998) utilisent un modèle Tobit après l'analyse DEA pour tenter d'expliquer les déterminants de l'efficience des écoles secondaires finlandaises. Elles trouvent des taux d'efficience moyens variant entre 82 et 84% selon les modèles, et constatent que le niveau d'éducation des parents joue un rôle déterminant comme facteur explicatif des écarts de performance des écoles.

BRADLEY et al. (2001) utilisent également la méthode DEA et un modèle Tobit pour évaluer l'efficience technique des écoles secondaires anglaises. Les taux d'efficience moyens obtenus sont compris entre 83 et 75%. Ces auteurs trouvent que la concurrence entre établissements scolaires améliore le niveau d'efficience des écoles analysées. Ce dernier résultat est confirmé par WALDO (2002) qui étudie la performance des écoles secondaires suédoises à l'aide de la méthode DEA.

Le propos de cette communication est de présenter la première application de la méthode DEA au système éducatif suisse, et en particulier aux écoles de maturité suisse romandes.

Références

Bessent, A. M. & Bessent, E. W. (1980), Determining the Comparative Efficiency of Schools through Data Envelopment Analysis, *Educational Administration Quarterly* 16, 57-75.

Bessent, A. M., Bessent, E. W., Kennington, J. & Reagan, B. (1982), "An Application of Mathematical Programming to Assess Productivity in the Houston Independent School District", *Management Science* 28, 1355-1367.

Bonesrønning, H. & Rattsø, J. (1994), "Efficiency variation among the Norwegian high schools: consequences of equalization policy", *Economics of Education Review* 13, 289-304.

Bradley, S., Johnes, G. et Millington, J. (2001), "The effect of competition on the efficiency of secondary schools in England", *European Journal of Operational Research*, 135, 545-568.

Charnes, A., Cooper, W.W. & Rhodes, E. (1978), "Measuring the efficiency of decision making units", *European Journal of Operational Research* 2, 429-444.

Färe, R., Grosskopf, S. & Weber, W. L. (1989), "Measuring school district performance", *Public Finance Quarterly* 17, 409-428.

Kirjavainen, T. & Loikkanen, H.A. (1998), "Efficiency Differences of Finnish Senior Secondary Schools: An Application of DEA and Tobit Analysis", *Economics of Education Review* 17, 377-394.

McCarty, T. & Yaisawarng, S. (1993), Technical efficiency in New Jersey school districts, in Lovell *et al.* (eds.), *The Measurement of Productive Efficiency*, 271-287, Oxford University Press, New York.

Ray, S.C. (1991), "Resource-use efficiency in public schools: A study of Connecticut data", *Management Science* 37, 1620-1628.

Waldo, S. (2001), "*Municipalities as Educational Producers – An Efficiency Approach*", Mimeo, Lund University.

UTILISATION D'UN MODELE DE REPONSE A L'ITEM POUR LA VALIDATION D'UNE ECHELLE DES ATTITUDES PROFESSIONNELLES DES MEDECINS GENERALISTES

Etienne Le Bihan[1], Cédric Baumann[23], François Alla[23], Michèle Baumann[1], (1) INtegrative research unit on Social and Individual DEvelopment (INSIDE), Université du Luxembourg, (2) Service d'Epidémiologie et d'Evaluation Cliniques, CEC – INSERM, Centre Hospitalier Universitaire de Nancy, (3) Ecole de Santé Publique, EA 4003, Faculté de Médecine, Université de Nancy, France.

Mots-clés : Echelle d'Attitudes Professionnelles des Médecins Généralistes, Modèle de Réponse à l'Item, Fonctionnement Différentiel d'Item.

Résumé :

Nous avons développé une échelle des attitudes professionnelles des médecins généralistes destinée à évaluer le niveau de communication thérapeutique dispensé durant les consultations. La construction de l'échelle a abouti à un questionnaire constitué de 15 items à 10 modalités de réponse allant de "Jamais" à "Toujours". Après avoir étudié les propriétés métriques de l'échelle selon une méthodologie classique, nous avons employé un modèle de réponse à l'item (le Generalized Partial Credit Model, GPCM) pour explorer un éventuel fonctionnement différentiel d'item (FDI) selon le genre des individus. Le modèle a été construit dans un cadre bayésien en utilisant un algorithme MCMC. La détection du FDI est basée sur le test de rapport de vraisemblance, en comparant le modèle de référence où tous les paramètres associés aux items sont communs aux hommes et aux femmes, à un modèle incluant deux ensembles de paramètres distincts selon le genre des individus pour un item particulier. La démarche a été appliquée aux 338 observations de l'échantillon de validation ne comportant pas de données manquantes. Parmi les 15 items de l'échelle, seul l'item 6 ("Mon médecin m'informe sur les effets secondaires des médicaments") présente un FDI significatif au niveau de signification de 5% ($p = 0.016$). Cet item est plus discriminant chez les hommes. De plus, il semble que, à un niveau de communication thérapeutique égal, les hommes ont une probabilité supérieure aux femmes de répondre aux catégories les plus élevées. Finalement, nous avons décidé de ne pas écarter cet item car 1) le FDI est modéré (si nous ajustons le test pour les comparaisons multiples, celui-ci n'est plus significatif au niveau 5%) et 2) nous n'avons aucune interprétation qualitative de cette différence.

275 mots.

Introduction

Contexte

Le développement d'une échelle d'attitudes des médecins généralistes a pour objectif de mesurer le niveau de communication thérapeutique dispensé par les médecins généralistes durant les consultations.

Selon l'Organisation Mondiale de la Santé, l'éducateur thérapeutique est "celui qui permet aux patients et à leur famille d'acquérir et de conserver les capacités et les compétences qui les aident à vivre de manière optimale leur vie avec leur maladie" (Organisation Mondiale de la Santé [OMS], 1998).

Dans le but d'améliorer l'état de santé des populations et de limiter les dépenses de santé, ce concept fait l'objet de nombreux programmes de politique de santé publique pour lesquels l'échelle des attitudes professionnelles des médecins généralistes à l'égard de la communication thérapeutique (GPPATC = General Partitionner Professional Attitudes towards Therapeutic Communication) a pour objectif de fournir un outil d'évaluation.

Construction des items du questionnaire et développement de l'échelle *

Le développement de l'échelle a fait l'objet d'une précédente publication (Baumann, Baumann, Aubry, Alla, 2005). Il s'est appuyé sur le recueil du discours de 40 usagers lorrains d'un centre de médecine préventive, 21 médecins généralistes et 22 pharmaciens d'officine de l'Est de la France. L'analyse de contenu a permis de retenir les items candidats parmi lesquels des experts ont sélectionnés les items génériques du questionnaire. Au final 15 items ont été retenus dans le questionnaire (tableau 1). A chacun des items, dix modalités de réponse sont proposées variant de "jamais" à "toujours".

Tableau 1 - Les items de l'échelle des attitudes professionnelles des médecins généralistes.

No	Item
1	Mon médecin prend le temps de m'écouter
2	Mon médecin fait le nécessaire pour gagner ma confiance
3	Mon médecin m'explique à quoi sert le traitement
4	Mon médecin tient compte de mes préférences pour la prescription (cachet, gouttes etc.)
5	Mon médecin me donne l'impression qu'il me respecte
6	Mon médecin m'informe sur les effets secondaires des médicaments
7	Mon médecin insiste sur les médicaments importants
8	Mon médecin parle avec moi des difficultés que j'ai à suivre le traitement
9	Mon médecin m'explique les choses avec des mots simples
10	Mon médecin me propose des nouveaux traitements
11	Mon médecin écrit lisiblement l'ordonnance
12	Mon médecin me laisse poser mes questions
13	Mon médecin me motive pour suivre mon traitement
14	Mon médecin me donne des conseils de prévention (alimentation, activité sportive…)
15	Mon médecin me donne l'impression de connaître son métier

* Michèle Baumann, assistante professeure en sociologie de la santé, responsable scientifique du programme, a obtenu le soutien financier de la Caisse Nationale d'Assurance Maladie, Paris, France.

Validation « classique »

Echantillon

L'échantillon pour la validation statistique du questionnaire se compose de 393 sujets recrutés dans un Centre Médecine Préventive en Lorraine en 2003. Les sujets ont été invités à remplir un deuxième questionnaire 15 jours après la première passation. L'âge moyen des sujets de l'échantillon était de 46.8 ans. La parité hommes / femmes est quasi-parfaite (50,4% d'hommes versus 49,6% de femmes).

Résultats

Les résultats de la validation classique montrent que l'échelle est unidimensionnelle puisque le premier facteur de l'Analyse en Composantes Principales représente 49.8% de l'inertie totale (autres facteurs : 8.8% à 1.2%). Tous les items sont bien représentés sur ce facteur. La cohérence interne est élevée (alpha de Cronbach = 0.92) et reproductibilité satisfaisante (Coefficient de corrélation intra-classe = 0.74 [0.66 ; 0.82]). La figure 1 montre cependant un effet seuil important pour certains items, ce qui suggère que l'échelle est peu discriminante pour des valeurs d'attitude élevées.

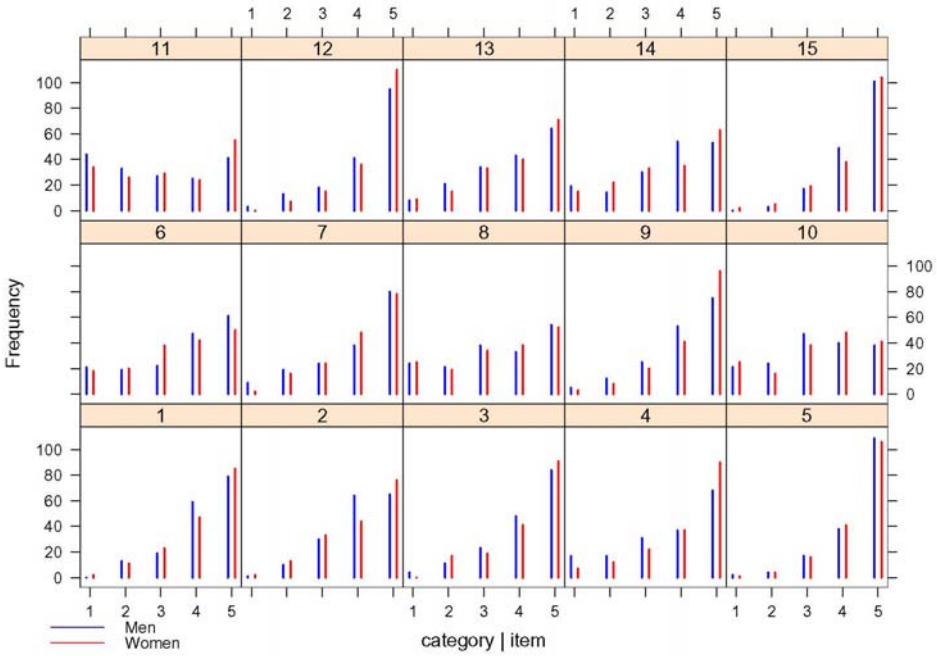


Figure 1 - Distribution des fréquences de réponse aux items.

Recherche d'un Fonctionnement Différentiel d'Item Hommes/Femmes

Définition

Un Fonctionnement Différentiel d'Item (FDI) se produit lorsque, pour un niveau d'attitude donné, la réponse attendue est différente selon le groupe d'appartenance de la personne qui a répondu au questionnaire (Penfield & Lam, 2000).

Le Generalized Partial Credit Model

Afin de tester la présence éventuelle d'un FDI, nous nous sommes appuyés sur un modèle de réponse à l'item, le Generalized Partial Credit Model (Muraki, 1992). Selon l'équation 1, le modèle associe la probabilité pour un individu i de répondre à la catégorie k de l'item j sachant qu'il a répondu à la catégorie k ou k-1, à 3 paramètres :

- θ_i : mesure de l'attitude thérapeutique du sujet i,
- b_{jk} : « difficulté » pour passer de la catégorie k-1 à la catégorie k,
- a_j : discrimination de l'item j.

$$\text{Logit}(P_{ijk|k-1,k}) = a_j\,(\theta_i - b_{jk}) \qquad (1)$$

Démarche

En supposant que le modèle est correct, les paramètres $\{a_j, b_{jk}\}$ résume totalement les propriétés métriques de l'item j. Ainsi, le FDI d'un item particulier j se traduit par le fait que les paramètres $\{a_j, b_{jk}\}$ diffèrent selon le groupe auquel l'item est proposé. Notre problématique se résume alors à tester l'égalité des 2 ensembles de paramètres $\{a_j^H, b_{jk}^H\}$ et $\{a_j^F, b_{jk}^F\}$, (H = Hommes, F = Femmes).

Envisageons les deux modèles GPCM suivants :

1. $\{a_j, b_{jk}\}$ pour tout j : tous les paramètres sont communs aux hommes et aux femmes,
2. $\{a_{j'}^H, b_{j'k}^H, a_{j'}^F, b_{j'k}^F\}$; $\{a_j, b_{jk}\}$, pour tout $j \neq j'$: les paramètres de l'item j' sont supposés différents entre les hommes et les femmes, les autres paramètres sont communs aux hommes et aux femmes.

Le modèle 1 étant un cas particulier du modèle 2, ces deux modèles sont emboités. Nous pouvons donc les comparer par le test du rapport du maximum de vraisemblance : sous l'hypothèse que le modèle 1 est correct, on démontre que

$$-2\{\text{Log}(L_1)-\text{Log}(L_2)\} \sim \chi^2(n_2-n_1) \qquad (2)$$

où L_1 et L_2 désignent respectivement les vraisemblances des modèles 1 et 2, n_1 et n_2 étant les nombres de paramètres respectifs des modèles 1 et 2.

Mise en oeuvre

La mise en œuvre s'appuie sur les algorithmes MCMC dans un cadre bayésien. Tous les modèles ont été ajustés grâce au logiciel WinBUGS (Spiegelhalter, Thomas, Best, Lunn, 2003). Le GPCM étant sur-paramétré, nous avons imposé aux paramètres associés à la première catégorie d'être nuls, quelque soit l'item. De même, la contrainte classique selon laquelle la distribution des θi est centrée et de variance 1 a

été retranscrite en posant une loi a priori centré réduite pour ces paramètres. Notons que cela n'est pas strictement équivalent, ce qui a pour conséquence que les estimations des paramètres ne sont pas immédiatement comparables à ceux fournis par les logiciels classiques. Néanmoins, les estimations sont liées entre elles par des relations linéaires dans lesquelles interviennent la moyenne et la variance de la distribution estimée des θ_i.

Nous avons posé des lois a priori non-informatives sur les autres paramètres : lois normales centrées de variance 100 et limitées aux valeurs positives pour les paramètres de discrimination, lois normales centrées de variance 100 pour les paramètres de c_{jk} (la paramétrisation du modèle a été modifiée en posant $c_{jk} = b_{jk}\ a_j$ afin d'améliorer la convergence des algorithmes MCMC).

Données

Les modèles ont été ajustés sur les observations recueillies durant la première passation. Les observations comportant au moins une donnée manquante ont été écartées. Au final l'échantillon est composé de 338 observations.

Pour des raisons techniques, les réponses appartenant à des catégories adjacentes ont été regroupées 2 par 2.

Résultats

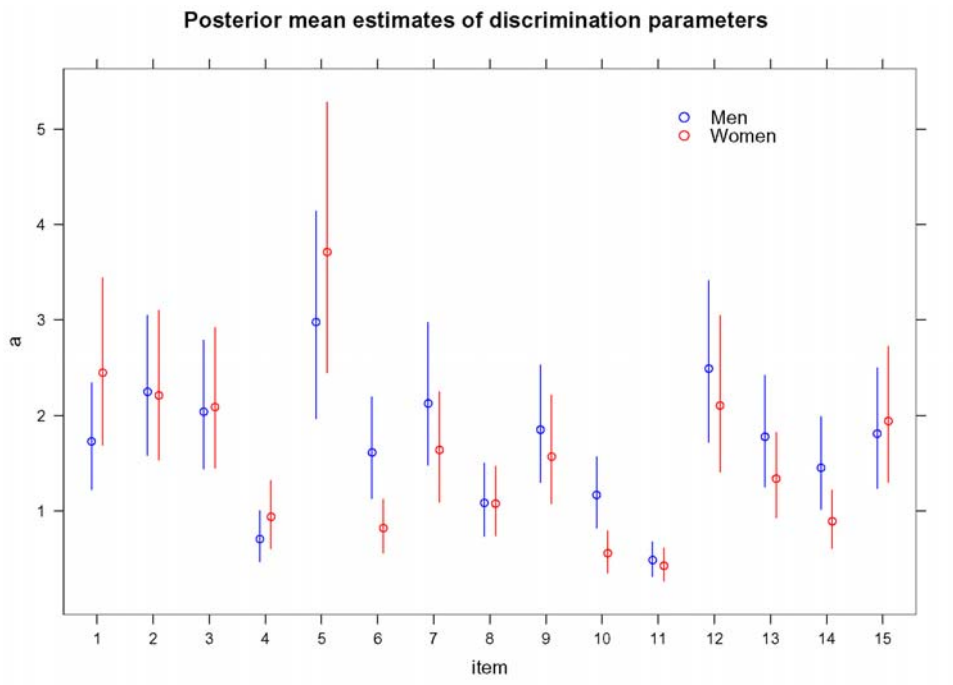


Figure 2 - Estimations et intervalles de crédibilité des paramètres de discrimination. Ces estimations proviennent de l'ajustement d'un modèle avec deux ensembles distincts de paramètres selon le sexe des individus pour chacun des items.

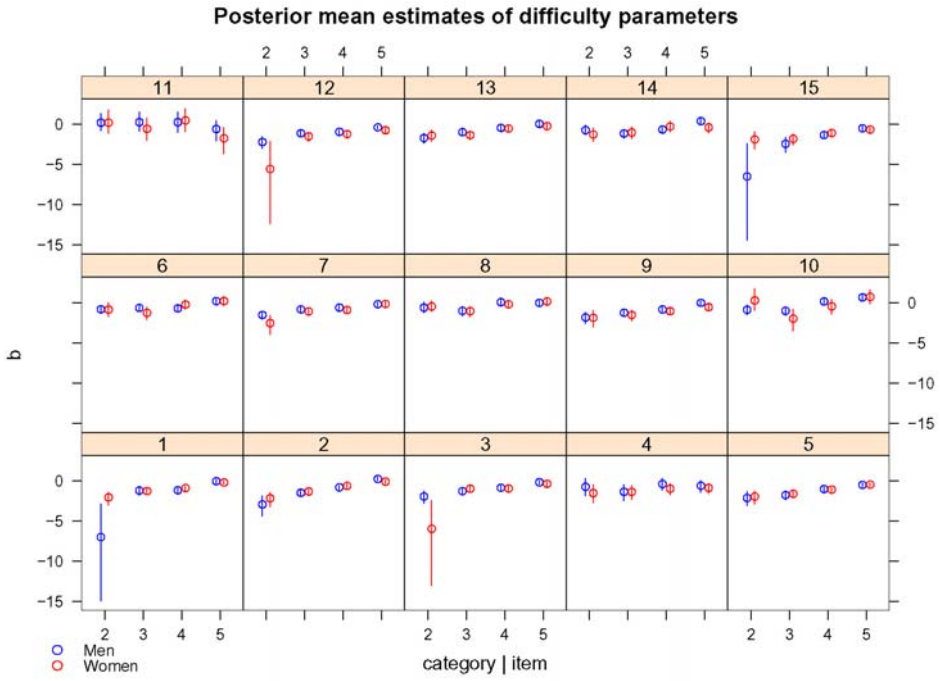


Figure 3 - Estimations et intervalles de crédibilité des paramètres de difficulté. Ces estimations proviennent de l'ajustement d'un modèle avec deux ensembles distincts de paramètres selon le sexe des individus pour chacun des items.

Les figures 3 et 4 montrent respectivement les estimations des paramètres de discrimination et de difficulté des items lorsque les ensembles de paramètres sont distincts chez les hommes et chez les femmes pour tous les items. On constate que les paramètres de discrimination sont proches chez les hommes et les femmes pour la plupart des items, excepté les items 6 et 10 pour lesquels les intervalles de crédibilité ne se chevauchent pas. On peut observer sur la figure 4 de très larges intervalles de crédibilité associés à des estimations fortement négatives pour la catégorie 2 des items 1, 3, 12 et 15 chez les hommes ou les femmes. Ceci s'explique aisément par le fait qu'aucun individu du sexe correspondant n'ait répondu à la première catégorie de ces items. Aucune différence notable n'est constatée sur les autres estimations des paramètres de difficulté.

Tableau 2 - Tests du FDI de chacun des items de l'échelle d'attitudes des médecins généralistes. Chaque ligne correspond à un modèle : la première correspond au modèle (1) où l'ensemble de paramètres associés à chacun des items est commun aux hommes et aux femmes, les suivantes correspondent au modèle (2), j' prenant successivement les valeurs indiquées dans la première colonne "Item". Les valeurs affichées dans la colonne "p" sont les probabilités pour que l'item indiqué en colonne 1 ne présente pas de FDI.

Item	-2LogL	Nombre paramètres	$\chi 2$	ddl	p
	10347.84	75			
1	10341.94	80	5.90	5	0.316
2	10343.63	80	4.21	5	0.520
3	10348.25	80	0.41	5	0.995
4	10341.87	80	5.97	5	0.309
5	10350.71	80	2.87	5	0.721
6	10333.97	80	13.87	5	0.016
7	10348.66	80	0.82	5	0.976
8	10352.85	80	5.01	5	0.414
9	10344.16	80	3.68	5	0.597
10	10341.26	80	6.58	5	0.254
11	10349.97	80	2.13	5	0.831
12	10343.41	80	4.43	5	0.489
13	10350.45	80	2.61	5	0.761
14	10342.15	80	5.69	5	0.338
15	10338.7	80	9.14	5	0.104

Seul le test du FDI de l'item 6 est significatif au niveau de 5% (tableau 2). Il s'agit de l'item "Mon médecin m'informe sur les effets secondaires des médicaments". L'observation des courbes caractéristiques de cet item chez les hommes et les femmes (figure 5) montrent une discrimination plus élevée chez les hommes. Lorsque l'attitude est proche de zéro, on constate aussi que la probabilité de répondre à la catégorie 4 est nettement plus élevée que celle de répondre à la catégorie 3 pour les hommes, alors que ces deux probabilités sont du même ordre chez les femmes.

Finalement, nous avons décidé de ne pas rejeter cet item pour les raisons suivantes :

- pas d'interprétation qualitative du FDI de cet item,
- le FDI est modéré, d'autant plus que nous n'avons pas tenu compte des comparaisons multiples (si on ajuste le test pour les comparaisons multiples par la méthode de Bonferroni, le test n'est plus significatif). On peut donc penser que son influence sur l'échelle est négligeable.

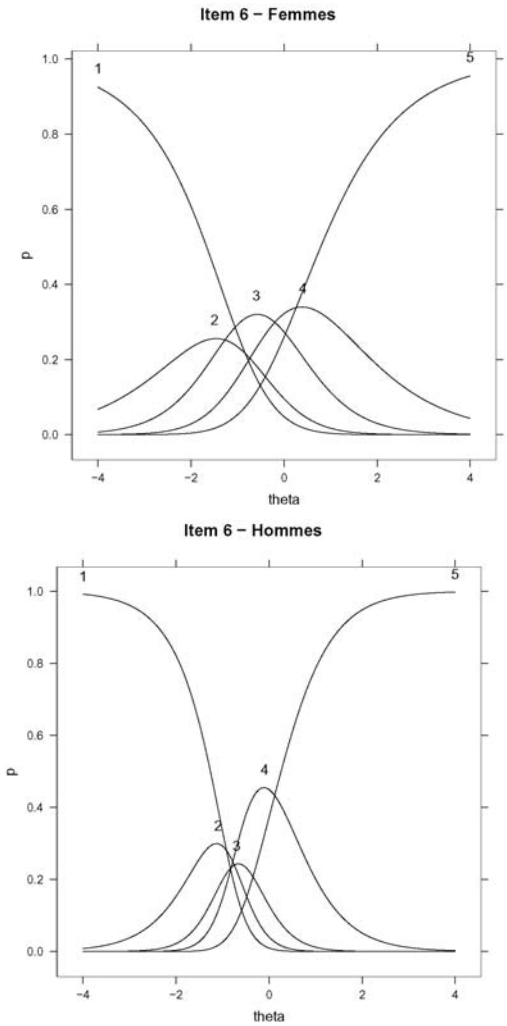


Figure 4 - Courbes caractéristiques de l'item 6 chez les femmes (à gauche) et chez les hommes (à droite).

Conclusion

Nous n'avons pas évoqué ici les aspects touchant à la métrique auxquels il faut pourtant être particulièrement attentif lorsque l'on cherche à mettre en évidence un éventuel fonctionnement différentiel d'item. En effet, si l'on suppose que les items ont des propriétés différentes selon les groupes auxquels ils sont soumis, cela revient à chercher à comparer les résultats obtenus par différents groupes à des échelles différentes. Dans notre cas, il est évident que les items de l'échelle présentent peu ou pas de FDI. Les modèles étant ajustés en supposant que tous les items sauf un ne présentent pas de FDI, l'influence de la métrique est assurément négligeable. Lorsque

cela n'est pas le cas, il existe différentes méthodes permettant de lier les échelles entre les groupes (Kim & Cohen, 1992). Citons la démarche de Lord qui consiste à tester la présence de FDI sur chacun des items, puis à écarter les items dont on suspecte qu'ils ont un FDI et à estimer les paramètres θ_i, et enfin à recommencer les tests de FDI en imposant que les paramètres associés aux individus soient égaux aux θ_i précédemment estimés.

La méthode est simple et efficace. Son principal désavantage est la durée d'ajustement des modèles (plusieurs heures). Sachant qu'il faut ajuster autant de modèles que l'échelle contient d'items plus un, cela rend la méthode pratiquement inutilisable en routine. Par contre, pour peu que l'on vérifie les conditions de convergence du modèle, l'utilisation des algorithmes MCMC assurent une qualité des estimations qui supporte un statut de méthode de référence. Plusieurs extensions sont envisageables :

- La prise en compte des données manquantes par augmentation des données,
- Modèle hiérarchique : les paramètres θ_i sont supposés tirés d'une population plus large,
- Prise en compte des données répétées (plusieurs observations étalées dans le temps chez les mêmes personnes) : dans ce cas, on peut imaginer de poser un modèle mixte sur les θ_i conjointement au modèle de réponse à l'item.

Références

Baumann M., Baumann C., Aubry C., Alla F. (2005). Echelle des attitudes professionnelles des médecins généralistes et des pharmaciens d'officine favorisant l'observance thérapeutique. *Revue Médicale de l'Assurance Maladie*, 36(1):23-33.

Kim S.-H., Cohen A. S. (1992). Effects of linking methods on detection of DIF. *Journal of Educational Measurement*, 29 (1):51-66.

Muraki E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16:159-176.

Organisation Mondiale de la Santé (1998). *Education thérapeutique du patient. Programmes de formation continue pour professionnels de soins dans le domaine de la prévention des maladies chroniques.* Bureau Régional pour l'Europe. Copenhagen : ed. OMS, 16.

Penfield R. D., Lam T. C. M. (2000). Assessing differential item functioning in performance assessment: Review and recommendations. *Educational Measurement: Issues and Practice*, 19(3):5-15.

Spiegelhalter D. J., Thomas A., Best N. G., Lunn D. (2003). *WinBUGS Version 1.4 User Manual*. Cambridge : MRC Biostatistics Unit.

UNE ALTERNATIVE EXPLORATOIRE A L'AJUSTEMENT PONDÉRAL POUR COMPENSER LA NON-RÉPONSE DES ÉLÈVES.

Christian Monseur, Université de Liège

Contexte de la recherche

Les enquêtes à large échelle en éducation sont souvent confrontées au problème de la non-réponse, problème susceptible d'engendrer des biais dans les résultats. La non-réponse peut apparaître à trois niveaux dans ce type d'études : (i) une école échantillonnée refuse de participer à l'enquête, (ii) un étudiant ne peut, pour diverses raisons, participer à l'évaluation et (iii) un élève participant refuse de répondre à une question.

Jusqu'à présent, la non-réponse des écoles et élèves a été compensée par une procédure d'ajustement pondéral. Cette procédure néanmoins postule que les écoles et les élèves qui n'ont pas participé présentent des caractéristiques semblables aux écoles et aux élèves qui ont participé.

Objectif de la recherche

Cette étude se propose d'analyser la non-réponse des élèves et notamment :

(i) de mettre à l'épreuve le postulat d'équivalence des populations participantes et non participantes d'élèves en analysant les données de PISA 2000;

(ii) d'étudier l'efficacité de la procédure d'ajustement pondéral pour la non réponse des élèves en fonction des caractéristiques structurelles des pays ;

(iii) de comparer cette efficacité par rapport à une alternative exploratoire :

a. qui exploite les possibilités offertes par les modèles de réponses à l'item associés à une méthode d'estimation selon une procédure de vraisemblance marginale et

b. qui consistent à générer des scores possibles (ou *plausible values*) ;

(iv) d'étudier les contraintes pratiques à la mise en œuvre de cette alternative exploratoire dans le cadre de *surveys* internationaux en éducation.

Résultats de la recherche

L'analyse des données du *Formulaire de Suivi des Elèves* de PISA 2000 a mis clairement en exergue l'existence de taux de participation différenciés selon le sexe et selon le grade fréquenté par l'élève âgé de 15 ans. A titre d'exemple, au Portugal, la participation des élèves de sexe masculin est inférieure de 5% à la participation des élèves de sexe féminin. Dans ce même pays, le taux de participation des élèves de 15 ans qui fréquentent le grade 7 s'élève à 76% alors que le taux de participation des élèves de 15 ans qui fréquentent le grade 10 s'élève à 88%. Ces résultats sont par ailleurs cohérents avec la littérature et d'autres recherches, à savoir que les élèves à plus faibles performances ont tendance à s'absenter davantage.

Par rapport au second objectif, les simulations effectuées permettent de démontrer le lien qui existe entre efficacité de l'ajustement pondéral et variance école. Dans les pays à faible variance écoles comme en règle générale les pays scandinaves, l'ajustement pondéral ne permet pas de réduire le biais introduit par la non réponse des élèves. Plus la variance école augmente, plus l'ajustement pondéral permettra de réduire le biais introduit par la non-réponse des élèves.

La simulation permet également de démontrer la supériorité de l'approche alternative à l'ajustement pondéral. Cette approche alternative sera d'autant plus efficace que les informations complémentaires obtenues par l'intermédiaire des professeurs ou de la direction d'écoles corrèlent avec la performance au test international.

Cependant, cette approche alternative ne permet plus de corriger les biais introduits par la non-réponse des élèves par rapport aux questionnaires contextuels, à moins que :

(i) deux séries de pondération soient fournies dans la base de données, une avant ajustement pondéral pour l'analyse des données cognitives et une après ajustement pondéral pour l'analyse des données contextuelles OU ;

(ii) les données manquantes dans les questionnaires contextuels soient également remplacées par des données générées sous l'hypothèse d'une distribution normale multi variée.

Bibliographie

Cette présentation se base sur l'article : Monseur, C. (2005). An exploratory alternative approach for student non response weight adjustment. *Studies in Educational Evaluation*, 31 (2005) 129-144.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

AVIS PEDAGOGIQUE, EVALUATION DES ENSEIGNEMENTS : QUELLES PRATIQUES, QUELS ENJEUX ?

Nathalie Delausnay, Institut des Sciences du Travail, Université libre de Bruxelles
Michel Sylin, Unité de Psychologie des Organisations, Université libre de Bruxelles

Mots-clés : évaluation pédagogique, université, haute école.

Résumé :
Le terme « évaluation » est polysémique. Selon les champs sémantiques qu'il recouvre, il s'apparente soit à de la gestion, soit à de l'excellence, soit à du management. Toutes les démarches d'évaluation dite pédagogique ne peuvent donc rassembler les mêmes pratiques et les mêmes objectifs.
Les différences de terrain qui mettent en œuvre un tel processus nuancent elles aussi les résultats de l'évaluation. Il en va ainsi des Universités et des Hautes Ecoles en Communauté française de Belgique, où l'obligation d'une évaluation externe n'existe que depuis peu de temps. Certaines Universités qui ont développé dès les années 70 des pratiques de prise d'« avis pédagogique » n'ont pas la même vision de l'évaluation que les Hautes Ecoles qui incluent souvent le développement de telles pratiques dans l'implantation de leur démarche qualité.
Mais si des différences de pratiques et d'enjeux existent effectivement, quelles sont-elles ? Pour tenter de répondre à cette question, nous avons effectué un recueil systématique des pratiques d'évaluation pédagogique dans une Université et une Haute Ecole de la Communauté française. Une fois mis en perspective en fonction des spécificités de terrain rencontrées, ce recueil nous a permis de situer l'évaluation pédagogique dans le cadre plus large d'une démarche qualité.
Il apparaît en effet que dans la plupart des Hautes Ecoles, c'est l'implantation d'une démarche qualité formalisée qui permet de mettre en place un processus d'évaluation pédagogique alors que dans les Universités où ces pratiques existent depuis longtemps, la démarche qualité doit impérativement être construite autour de tels éléments structurants. Serait-ce à dire que, dans le premier cas, c'est le sens donné aux démarches qualité qui déterminera le sens que prendront les évaluations pédagogiques et que, dans le second, la démarche qualité ne sera véritablement implantée qu'une fois intégré l'ensemble des éléments structurants préalablement existant ?

L'évaluation : un concept polysémique

L'évaluation est une notion assez courante, souvent utilisée lorsqu'on attribue une valeur, un prix ou un degré d'importance à une chose, un bien ou un service. Proche d'autres concepts comme la mesure, le contrôle ou le suivi, la notion d'évaluation recouvre un large éventail d'interprétations, surtout dans le domaine de l'éducation. Dennery souligne toutefois la différence qui existe entre évaluation et contrôle. Dans son acception française, le contrôle fait en effet référence à une notion de vérification

par rapport à une norme. Et si, dans l'évaluation comme dans le contrôle, il s'agit toujours de procéder à une comparaison par rapport à des critères précis, il n'y a pas forcément de connotation coercitive dans l'évaluation (Dennery, 2001).

Dans le Dictionnaire actuel de l'éducation, Legendre (2005) propose que le terme « évaluation » couvre des champs sémantiques différents selon l'emploi qu'on en fait. Il peut ainsi s'agir d'un processus cognitif comprenant la comparaison référent-référé et le jugement (Noizet et Caverni, 1978 ; Bloom, 1956 ; Guilford, 1967) ou d'un processus complexe incluant en plus la planification des étapes, la prise d'information (Linn et Grönlund, 1995 ; Vial, 1997, 2001) et la prise de décision (Stufflebeam, 1980). Pour Hadji (1997), évaluer c'est « se prononcer, c'est-à-dire prendre parti, sur la façon dont des attentes sont réalisées ». Le Ministère de l'Education du Québec (Meq, 2004) définit pour sa part l'évaluation comme une « démarche permettant de porter un jugement, à partir de normes ou de critères établis, sur la valeur d'une situation, d'un processus ou d'un élément donné, en vue de décisions pédagogiques ou administratives ».

Il est par ailleurs d'usage depuis les années septante de distinguer deux grands types d'évaluation : l'évaluation sommative et l'évaluation formative. La première consiste à établir un bilan des connaissances acquises par l'apprenant à la fin de son apprentissage en sanctionnant l'activité d'apprentissage de façon positive ou négative. L'évaluation formative a quant à elle pour but de mesurer, tout au long de l'apprentissage, les performances de l'apprenant en prenant en compte son niveau initial. Bloom et coll. (1971, cités dans Legendre, 2005) définissent pour leur part l'évaluation sommative comme l'ensemble des opérations d'évaluation menées à des fins de sanction et l'évaluation formative comme l'ensemble des opérations menées à des fins de régulation des apprentissages.

En d'autres termes, l'évaluation désigne tantôt une pratique conformative impliquant, directement ou non, une notion de gestion où le contrôle est présent de manière explicite, tantôt une pratique normative impliquant une notion de management plus adaptative aux contingences environnementales (Delausnay, Sylin et Wettendorff, 2005).

Les pratiques de l'évaluation pédagogique

Dans une perspective plus adaptée au domaine de l'enseignement supérieur et de la recherche, la Commission ministérielle de terminologie de l'éducation (France, 1992) définit l'évaluation comme l'« action de mesurer à l'aide de critères objectifs la valeur d'un enseignement, d'un programme, d'un apprentissage, la qualité d'une recherche… ». Dans cette optique, l'évaluation consiste ici à apprécier objectivement tous les aspects reliés directement aux apprentissages (rendement et caractéristiques des sujets ; programmes d'études ; enseignement ; évaluation ; gestion de l'enseignement et de l'apprentissage ; personnel et établissements) en vue d'assurer le meilleur système éducatif possible (Legendre, 2005). Cette approche rencontre de la

notion d'excellence telle qu'elle a été analysée il y a quelques temps déjà par Aubert et de Gaulejac (2001).

Toutes les démarches d'évaluation dite pédagogique ne recouvrent cependant pas ni les mêmes pratiques ni les mêmes objectifs. On parle ainsi d'évaluation certificative lorsque le processus d'évaluation débouche sur une décision de réussite ou d'échec relatif à une période d'apprentissage (De Ketele et Roegiers, 1993) ; d'évaluation continue quand il s'agit d'un processus cumulatif d'évaluations fréquentes et ordonnées à divers moments de l'apprentissage ; d'évaluation des programmes d'études lorsque le processus consiste à analyser l'état des programmes en fonction des objectifs visés et des ressources disponibles ; d'évaluation des apprentissages quand on évalue les compétences, connaissances, habiletés ou attitudes d'un élève dans le but de déterminer dans quelle mesure les objectifs éducatifs ont été atteints (Meq, 2004) ; d'évaluation de l'enseignement lorsqu'il s'agit d'apprécier les caractéristiques d'un enseignement, d'un ensemble d'objectifs ou d'un milieu scolaire à partir de critères précis et d'une méthodologie appropriée (Legendre, 2005). Dans le cadre de cette étude, c'est bien d'évaluation de l'enseignement dont il sera question…

L'évaluation de l'enseignement

Objet de l'évaluation de l'enseignement

Définir l'objet de l'évaluation de l'enseignement ne peut se faire en dehors du concept même d'enseignement. Selon Doyle (1983, cité dans Poissant, 1996), l'enseignement comprend trois aspects principaux : le processus, la substance et le résultat. Concernant le processus, plusieurs études ont permis de dégager cinq catégories de dimensions qui y sont liées (Bernard et Trahan, 1989, cités dans Poissant, 1996) : les capacités d'analyse et de synthèse du professeur, l'organisation du cours et la clarté des explications, l'interaction du professeur avec son groupe, l'interaction du professeur avec l'étudiant, le dynamisme et l'enthousiasme du professeur. L'évaluation de la substance comprend pour sa part des aspects tels que la connaissance de la matière enseignée, l'organisation et la structure du cours, la qualité des moyens d'évaluation. Le troisième aspect, celui du résultat, est quant à lui fort peu exploité. L'apprentissage de l'étudiant en tant qu'objet de l'évaluation de l'enseignement a en effet été peu étudié jusqu'à présent, sans doute parce qu'il n'existe pas à l'heure actuelle de théorie de l'apprentissage unanimement acceptée (Poissant, 1996). Sans doute également car l'apprentissage de l'étudiant comme résultat de l'enseignement présente certaines difficultés inhérentes à sa mesure. A ce sujet, Doyle (1983) mentionne différentes variables qui peuvent interférer avec l'apprentissage et qui échappent pourtant facilement à l'emprise du professeur, comme la motivation, l'intérêt, l'attitude ou les besoins de l'étudiant.

Sources d'information

Les étudiants constituent à l'heure actuelle l'une des sources les plus employées pour évaluer la qualité de l'enseignement. Une enquête effectuée par Seldin en 1984

montre qu'au début des années 80, 70% des universités et collèges américains avaient recours aux évaluations par les étudiants. Apparemment, cette tendance n'a fait que s'accroître depuis, même si l'évaluation par les étudiants reste la méthode d'évaluation la plus controversée... Les résistances des enseignants face à ce type d'évaluation sont souvent liées au fait qu'ils estiment qu'un étudiant ne peut porter de jugement sur un cours qu'il découvre progressivement et qu'il ne peut comparer à d'autres. Plusieurs études ont néanmoins prouvé l'importance du rôle et de la valeur des évaluations faites par les étudiants (Abrami et d'Apollonia, 1990 ; Marsh, 1987).

Une étude sur la perception des étudiants à l'égard de l'évaluation de l'enseignement a par ailleurs été réalisée auprès de 391 sujets de premier cycle universitaire et provenant de sept disciplines différentes (Thivierge et Bernard, 1996). Les résultats ont montré que si les étudiants jugent le processus important et croient qu'ils ont un rôle à jouer dans cette opération, ils doutent de son utilité parce qu'ils sont convaincus que les professeurs et l'institution y accordent peu de valeur.

Murray (1984) affirme pour sa part que si les étudiants sont une source privilégiée d'information, ils ne peuvent être l'unique source pour juger de la qualité de l'enseignement des professeurs. L'évaluation de l'enseignement devrait ainsi provenir non seulement des étudiants mais aussi des diplômés, des pairs, de spécialistes en enseignement et des enseignants eux-mêmes. Ainsi, le spécialiste en enseignement pourrait s'exprimer sur la qualité de l'enseignement à des fins formatives alors que le pair pourrait davantage intervenir à des fins de promotion en jugeant la pertinence des contenus enseignés (Bernard et al., 2000). L'enseignant pourrait quant à lui décrire sa situation d'enseignement, expliquer ses approches pédagogiques et mettre en évidence les variables qui ont pu influencer son enseignement (Bernard, 1992). Il semblerait par ailleurs que les enseignants aient une connaissance plus ou moins juste de la perception des étudiants à leur égard. Les corrélations entre les évaluations par les étudiants et l'autoévaluation de l'enseignant varient en effet, selon les études, entre 0,20 et 0,49 (Poissant, 1996).

Buts de l'évaluation de l'enseignement

D'après Poissant (1996), la majorité des auteurs s'entend pour dire que l'évaluation de l'enseignement poursuit 4 buts principaux : fournir un diagnostic et aider l'amélioration de l'enseignement (évaluation formative) ; fournir des informations pour les décisions administratives (évaluation sommative) ; fournir des informations aux étudiants pour leur choix de cours et fournir des informations aux planificateurs de programmes ; fournir des données pour la recherche en enseignement.

Si pour Murray (1984, cité dans Poissant, 1996), l'évaluation formative et l'évaluation sommative convergent car elles visent toutes deux l'amélioration de l'enseignement, on note qu'il n'en va pas toujours de même dans les faits, l'évaluation formative étant souvent délaissée au profit d'une simple sélection des enseignants qui obtiennent des scores élevés aux questionnaires d'évaluation. Centra (1979, cité dans Poissant, 1996) affirme à ce sujet que lorsque ces deux systèmes coexistent, ils devraient rester indépendants ; les établissements d'enseignement

devraient par ailleurs davantage mettre à profit les informations diagnostiques fournies par l'évaluation formative et développer des programmes de formation appropriés.

Du bon usage des évaluations par les étudiants

La lecture des principaux écrits traitant de l'évaluation de l'enseignement permet de relever un certain nombre de limites dont il faudrait tenir compte lors de la mise en place d'une telle procédure.

On note ainsi que, selon plusieurs auteurs, dix items généraux doivent suffire à cerner le concept d'enseignement, ce nombre pouvant monter jusqu'à vingt quand l'évaluation sert exclusivement à des fins d'amélioration de l'enseignement (évaluation formative). Les items doivent par ailleurs être assez spécifiques pour permettre aux enseignants de convertir ces informations en comportements concrets. L'emploi d'items spécifiques pose cependant des problèmes lorsqu'il s'agit d'une évaluation sommative. Abrami et d'Apollonia (1990) affirment à ce sujet qu'il n'existe pas de qualités intrinsèques pour décrire un bon enseignement car un enseignement prend toujours forme dans un contexte et des circonstances particuliers. Pour cette raison, l'enseignement ne peut être défini par un ensemble d'items spécifiques, si nombreux soient-ils.

Ainsi, la conception nécessairement limitée résultant d'une évaluation au moyen d'un questionnaire risque de mener à un point de vue partiel ou partial de l'enseignement (Braskamp et Ory, 1994). De fait, lorsque les questionnaires d'évaluation par les étudiants sont employés à des fins de promotion, ils devraient toujours être utilisés conjointement avec d'autres sources d'évaluation. Le système d'évaluation de l'enseignement dépasse en effet la seule prise de mesure au moyen des questionnaires d'évaluation par les étudiants et il importe donc que les opérateurs de système élargissent leur vue de façon à ne pas réduire l'évaluation à une seule composante (Poissant, 1996). Tout nouveau système d'évaluation devrait également bénéficier d'une période de rodage d'au moins une année, période durant laquelle aucune décision concernant le personnel ne devrait être prise.

Au niveau des résultats, la description des résultats transmis aux enseignants devrait inclure des explications quant aux chiffres présentés et à leur pertinence par rapport aux construits liés à l'enseignement et à l'apprentissage (Poissant, 1996). On constate par ailleurs souvent l'absence de liens entre des résultats négatifs et les prises de décisions, ce qui explique certainement le grand manque d'intérêt que l'on peut noter de la part des acteurs vis-à-vis de l'évaluation et plus particulièrement de la part des étudiants qui doutent souvent que leurs évaluations aient un quelconque effet sur les prises de décisions. Outre ce manque d'effet concret, les évaluations peuvent également souffrir d'un accroissement du niveau de frustration, les améliorations à mettre en place dépendant le plus souvent de ressources financières non disponibles. A tout cela s'ajoute le fait que les évaluations entraînent généralement une importante

surcharge de travail, le coût de l'évaluation en temps, moyens et personnels étant d'ailleurs rarement abordé de façon explicite.

Malgré ces limites, il faut souligner l'importance que peuvent avoir les résultats des évaluations comme levier à des changements dans la pédagogie (Demailly, 2001).

Les pratiques d'évaluation pédagogiques à l'ULB et à la Helb Ilya Prigogine

La définition des sources d'information et des buts de l'évaluation n'est pas chose aisée. C'est pour cette raison que les établissements d'enseignement tels que les Hautes Ecoles de la Communauté française ont à définir clairement dans leur projet pédagogique, social et culturel les modalités de cette évaluation. Ces modalités semblent par ailleurs différer sensiblement des pratiques d'évaluation mises en place par les Universités, où le terme d'« avis pédagogique » remplace souvent celui d'évaluation.

Signalons qu'en Communauté française de Belgique, l'évaluation fait partie des domaines où la liberté pédagogique de chaque réseau est garantie. Chaque pouvoir organisateur peut ainsi définir son type d'évaluation ainsi que les moyens mis en œuvre pour conduire cette évaluation et la façon d'en communiquer les résultats. Certaines universités ont, en Communauté française, développé un processus d'évaluation pédagogique dès les années 70, impliquant alors directement l'avis des étudiants dans la gestion de la carrière de leurs enseignants. Et si, pendant longtemps, il n'a pas existé en Communauté française de Belgique de processus d'évaluation externe aux établissements qui soit obligatoire, le décret créant l'Agence pour l'évaluation de la qualité dans l'enseignement supérieur en Communauté française de Belgique (2002) oblige désormais tous les établissements d'enseignement supérieur à évaluer la qualité de leurs formations.

Partant de là, les pratiques d'évaluation semblent présenter autant de spécificités que les terrains qui les mettent en œuvre. C'est pour cette raison qu'un premier volet de notre recherche-action consiste en un recueil systématique des pratiques d'évaluation pédagogique dans une Université (Université libre de Bruxelles) et une Haute Ecole (Haute Ecole libre de Bruxelles Ilya Prigogine) de la Communauté française.

Les avis pédagogiques à l'ULB

Dans la foulée de mai 1968, les étudiants de l'Université libre de Bruxelles ont acquis le droit d'émettre un avis pédagogique sur leur enseignement à l'université. Pour ce faire, une même structure a été mise en place dans toutes les facultés et instituts de l'ULB : la commission d'évaluation pédagogique. Composée d'un nombre de membres variable selon la taille des institutions concernées, cette commission fonctionne dans tous les cas avec des représentants des corps académique et scientifique et des représentants étudiants.

Outre la composition des commissions d'évaluation pédagogique, on note certaines similitudes dans le fonctionnement des différentes commissions et la mise en œuvre des avis pédagogiques. Il apparaît ainsi que les prises d'avis pédagogiques sont désormais obligatoires dans tous les dossiers relatifs au renouvellement des mandats d'assistant, aux promotions et nominations des enseignants. Un avis pédagogique peut en principe également être organisé dans toutes les facultés et instituts à la demande d'un enseignant ou d'un étudiant. Des procédures relativement semblables sont par ailleurs prévues en cas d'avis défavorables, les avis favorables n'entraînant pas de mesures concrètes immédiates. Il est à ce sujet coutume d'obliger les étudiants à justifier leurs positions négatives, les réponses incomplètes (car non justifiées) n'étant généralement pas prises en compte lors du dépouillement.

A côté de ces points communs, des divergences existent, qui portent le plus souvent sur les choix méthodologiques ou le fonctionnement même des prises d'avis pédagogiques et qui sont par ailleurs fortement liées à la taille et à la structure des facultés et instituts concernés. De grosses facultés telles que la faculté de médecine bénéficient ainsi d'un traitement entièrement automatisé des avis pédagogiques (du questionnaire en ligne au traitement des résultats) alors que dans les instituts de petite taille, les membres des commissions dépouillent les questionnaires manuellement et rédigent les rapports de résultats. Cette situation entraîne par conséquent de grandes différences dans la charge de travail des commissions d'évaluation pédagogique et dans les choix méthodologiques possibles… Selon les cas, la prise d'avis pédagogiques est en effet soit annuelle soit uniquement liée aux renouvellements de mandats et aux demandes de promotions. Les questionnaires, qu'ils soient présentés en version papier ou en version électronique, sont quant à eux composés soit d'une unique question d'appréciation globale soit d'une série d'items ayant trait à l'enseignant, au cours (structure, etc.), aux modalités d'évaluation et à la qualité des supports utilisés. A noter que le volet « modalités d'évaluation » ne se retrouve pas dans tous les canevas d'avis pédagogiques mais que nombreuses sont les commissions qui cherchent à l'y intégrer. Certaines facultés ont par ailleurs décidé de rendre les avis pédagogiques obligatoires (pas de feuille d'options, pas d'inscription aux examens) ou non anonymes alors que d'autres ont choisi d'insérer une question sur le pourcentage de fréquentation des étudiants aux cours évalués.

Au-delà des éléments de convergence et de divergence entre facultés et instituts de l'ULB, plusieurs questions se posent quant à la façon dont l'évaluation pédagogique peut ou devrait fonctionner. On remarque par exemple qu'il est très difficile pour les commissions d'évaluation pédagogique de maintenir les listes du personnel à jour sans aide logistique extérieure ou de rentrer les avis pédagogiques de plus en plus tôt dans l'année académique. Quand les résultats sont transmis à d'autres personnes que les enseignants/assistants eux-mêmes, les commissions d'évaluation pédagogique ne reçoivent habituellement aucun feed-back sur ce que ces personnes auraient éventuellement entrepris suite à la communication des résultats.

L'engagement des étudiants pose par ailleurs problème dans presque toutes les structures effectuant des prises d'avis pédagogiques « non sanctionnés », la crainte de ne pouvoir atteindre les quorums requis impliquant souvent des compromis

méthodologiques où la question reste posée de privilégier le quantitatif par rapport au qualitatif. Ce manque d'engagement s'expliquerait par des profils d'étudiants parfois très atypiques (étudiants de DES, qui ne restent qu'un an) ne s'impliquant pas de la même façon dans la vie de leur institution et par un manque de sanctions directes visibles lors d'avis pédagogiques défavorables. Il est toutefois étonnant de constater que les étudiants sont en général très peu sévères dans leurs appréciations globales sur les enseignants. Craindraient-ils l'une ou l'autre représailles ?

Au niveau des corps scientifique et académique, les avis sont mitigés. Certains enseignants sont en effet réellement impliqués dans une démarche d'amélioration continue alors que d'autres ne supportent pas d'entendre les critiques formulées par leurs étudiants. On constate également certaines variations entre divers groupes d'étudiants pour un même cours, ce qui n'aide évidemment pas à légitimer l'outil… Il faut par ailleurs noter le peu de reconnaissance générale des activités pédagogiques à l'ULB. Un mauvais pédagogue peut en effet être « sauvé » par une importante activité de recherches, ce que les étudiants et certains enseignants acceptent difficilement. Dans les deux cas de figure, les commissions d'évaluation pédagogique ont un rôle important à jouer pour responsabiliser les étudiants face aux avis pédagogiques et sensibiliser les enseignants à la notion d'amélioration continue de la qualité de l'enseignement.

Pour tenter de répondre à certaines de ces questions ou pour faire taire des critiques formulées en interne, des commissions d'évaluation pédagogique ont entrepris de renouveler entièrement leur système de prise d'avis pédagogiques, notamment sur base de conseils de spécialistes en éducation ou d'une plus large consultation des étudiants. On note dans ces cas une forte tendance à mélanger dans un même questionnaire des éléments d'évaluations sommative et formative. Concrètement, cela se traduit par une ou deux questions générales dont les résultats sont versés aux dossiers des personnes évaluées et une série d'items plus spécifiques dont les résultats sont uniquement transmis aux personnes directement concernées par l'avis pédagogique. Etonnamment, une commission s'est vue obligée de réaliser le travail inverse, c'est-à-dire de simplifier un outil auparavant très riche d'un point de vue qualitatif mais où l'interprétation des résultats ne permettait pas de rendre un avis clair en terme de favorable ou défavorable, en terme de résultat final chiffré…

Une réflexion est également menée par certaines commissions sur la façon de réaliser les avis pédagogiques dans les années terminales, les étudiants étant plus difficiles à impliquer une fois leurs études achevées. Le moyen le plus simple pour les joindre est en effet de leur envoyer un questionnaire par courrier mais on compte à peine 1/3 de réponses dans ce type d'enquête... Or, la dernière année est en général beaucoup plus axée sur des aspects liés à la vie professionnelle et il importe donc de voir dans quelle mesure le programme d'études de la dernière année répond aux exigences du monde professionnel.

De façon générale, on note que les membres des commissions qui désirent se lancer dans une réflexion plus large sur la qualité des outils d'évaluation manquent souvent de temps pour le faire, les outils étant le plus souvent utilisés à l'identique depuis des années… Il n'existe par ailleurs aucun moment de rencontre entre les

présidents et membres des différentes commissions d'évaluation pédagogique, ce qui ne les encourage pas à remettre leurs pratiques en question. De même, il n'y a apparemment aucune concertation entre les acteurs concernés par les prises d'avis pédagogiques et ceux concernés par les exercices d'évaluation de la qualité de l'enseignement à l'Université, ces pratiques étant de toute évidence considérées comme deux choses totalement à part.

L'évaluation des enseignements à la Helb Ilya Prigogine

En Communauté française de Belgique, les Hautes Ecoles restent dans beaucoup de cas des agrégats de structures qui tendent à préserver leur autonomie et leur intégrité d'avant leur regroupement. Ainsi, elles vivent avec la crainte de voir leur identité mais aussi leur objet institutionnel disparaître. Au niveau de l'évaluation des enseignements, cette situation engendre le plus souvent une grande disparité dans les pratiques mises en place par l'un ou l'autre enseignant, sans concertation avec le reste des enseignants dans la section ou les autres sections de l'établissement. La Helb (Haute Ecole Libre de Bruxelles) Ilya Prigogine n'échappait pas à cette constatation jusqu'à l'arrivée du « décret qualité » de 2002 obligeant chaque établissement d'enseignement supérieur de la Communauté française à mettre en place une démarche qualité formalisée dans un certain nombre de sections.

Dans ce contexte, et dans la mesure où la section électronique de la Helb Ilya Prigogine était directement concernée par les premiers exercices d'évaluation demandés par l'Agence Qualité, la Commission Qualité Electronique a choisi de mettre en place une première évaluation globale des enseignements dispensés dans les trois années d'études de la formation. Pour ce faire, la Commission Qualité Electronique a réalisé, en concertation avec les délégués étudiants, un questionnaire composé de 13 items assez généraux puisqu'identiques pour les trois années d'études. La Commission a par ailleurs choisi de reprendre tous les cours dans un même questionnaire, à l'inverse de certaines évaluations qui ne portent que sur un cours à la fois. Ces choix méthodologiques ont été fait pour parvenir à une vision globale de la qualité de l'enseignement en électronique d'abord ; pour pouvoir suivre l'évolution des étudiants d'une année à l'autre ensuite ; et enfin, pour ne pas cibler directement les enseignants mais bien les enseignements. La Commission a également choisi de diffuser le questionnaire via un outil de questionnaires en ligne créé en interne, ce qui permet aux étudiants de choisir les cours pour lesquels ils désirent ou non répondre aux questions, et ce pour chaque question individuellement. Cette possibilité et une division semestrielle de l'évaluation devaient en principe alléger au maximum l'exercice d'évaluation.

Une version papier des questionnaires a néanmoins été utilisée pour les évaluations du second semestre. Les taux de réponse pour les premières évaluations ne dépassaient en effet pas les 50% par année, ce qui s'explique notamment par le fort taux d'absentéisme et la difficulté de joindre les étudiants pour leur transmettre leurs codes d'accès et les sensibiliser à l'importance de l'évaluation. L'ensemble du dispositif d'évaluation sera par conséquent réenvisagé pour l'année académique à venir, sans doute sur base des résultats de cette recherche-action et des réflexions

menées en parallèle dans d'autres sections de la Haute Ecole, en kinésithérapie notamment.

Outre cette évaluation « à chaud » des enseignements, trois autres questionnaires « à froid » ont été créés par la Commission Qualité Electronique. Le premier est destiné aux étudiants diplômés ces dix dernières années ; le deuxième s'adresse aux employeurs en électronique et a pour principaux objectifs de compléter l'étude de l'insertion professionnelles de nos étudiants diplômés et de juger de l'adéquation de nos enseignements avec les exigences des employeurs ; le dernier porte sur les conditions de vie et de travail des étudiants de la Haute Ecole dans son ensemble.

Il faut enfin signaler que tous les outils de questionnement créés dans le cadre de l'exercice d'évaluation effectué dans la section électronique de la Helb ont été conçus de façon à pouvoir adapter des canevas communs aux spécificités de chaque formation et ce afin de dépasser les singularités de chaque parcours étudiant et/ou professionnel pour promouvoir dans la Haute Ecole la mise en place de mesures adaptables à tous. C'est donc bien l'exercice d'évaluation de la qualité effectué dans la section électronique qui permet d'introduire ce type d'outils dans toutes les autres sections de la Haute Ecole.

Avis pédagogiques et évaluation des enseignements…

Si le terme d'évaluation pédagogique pourrait en première analyse convenir aux pratiques d'évaluation mises en place à l'ULB et à la Helb Ilya Prigogine, une étude plus approfondie des réalités de terrain permet de constater que le raccourci n'est pas vraiment approprié.

Certes, de par les caractéristiques des structures concernées, on pourrait penser que le mode de fonctionnement est assez similaire. On a en effet des deux côtés un ensemble d'entités qui, regroupées dans une même institution, peuvent mettre en place des évaluations pédagogiques finalement assez libres par rapport au canevas commun proposé. On note par ailleurs que, dans les deux cas également, le nombre d'étudiants concerné par ces évaluations fait que le grand défi lors de la construction des outils de questionnement consiste à trouver un juste équilibre entre richesse des informations récoltées et nombre d'avis rentrés. Mais force est de constater que les parallélismes ne vont guère plus loin…

Côté ULB, on est en effet dans un système vieux de plusieurs dizaines d'années alors que le système n'est pas encore mis en place dans l'ensemble des sections de la Helb. L'utilisation qui est faite des résultats change elle aussi selon l'institution puisque qu'elle sert essentiellement d'évaluation sommative à l'ULB alors que c'est surtout d'évaluation formative dont on parle à la Helb. Les différentes commissions qui travaillent à la création des questionnaires d'évaluation des enseignements à la Helb ont par ailleurs décidé de ne jamais utiliser les questionnaires d'évaluation par les étudiants comme seul outil d'évaluation, raison pour laquelle d'autres outils ont été créés en parallèle… Enfin, la Helb place résolument son évaluation des enseignements dans le cadre plus large de l'évaluation de la qualité alors qu'à l'ULB,

le lien semble difficile à établir. Serait-ce là la plus grande différence entre avis pédagogiques à l'Université et évaluation des enseignements dans une Haute Ecole ?

Les évaluations pédagogiques et les démarches qualité

Nous l'avons déjà signalé, tous les établissements d'enseignement supérieur de la Communauté française sont aujourd'hui soumis aux exercices d'évaluation prévus par la nouvelle Agence Qualité créée en 2002. Qu'il s'agisse d'une Université ou d'une Haute Ecole, les indicateurs qualité englobent un volet « évaluation des enseignements », et ce quelles que soient les formations concernées par l'évaluation.

Parler de qualité de l'enseignement, c'est parler des interactions entre l'institution et les ressources dont elle dispose, entre l'enseignant et les méthodes pédagogiques qu'il développe, entre l'enseignant et l'engagement des étudiants. A ce titre, les avis pédagogiques et les évaluations des enseignements permettent d'obtenir de précieuses informations quant à la perception des étudiants vis-à-vis des enseignants, des cours et des modalités d'évaluation. Il est toutefois primordial de bien choisir les items en fonction des informations que l'on désire obtenir. Demander si Monsieur Y est un bon enseignant est un peu limite en termes de démarche qualité !... Les évaluations pédagogiques devraient donc toujours être inclues dans une panoplie plus large d'activités qui, ensemble, permettront de juger de la qualité des enseignements dispensés.

Dans la mesure où l'évaluation de l'enseignement s'intègre dans une évaluation de tous les aspects de l'établissement, elle engage les établissements d'enseignement dans un processus continu, progressif, systématique et constamment amélioré d'évaluations, même si celles-ci ne sont pas toujours perçues favorablement. Dans cette perspective, l'évaluation constitue véritablement un outil stratégique de développement de la qualité pour l'ensemble des acteurs, des composantes, des disciplines (Demailly, 2001).

En guise de conclusion

Le recueil des pratiques d'évaluations pédagogiques a montré qu'à la Helb, comme dans la plupart des Hautes Ecoles de la Communauté française, c'est l'implantation d'une démarche qualité formalisée qui permet de mettre en place un processus d'évaluation pédagogique alors qu'à l'Université, la démarche qualité n'est pas forcément construite autour de tels éléments structurants.

Serait-ce à dire que, dans le premier cas, c'est le sens donné aux démarches qualité qui déterminera le sens que prendront les évaluations pédagogiques et que, dans le second, la démarche qualité ne sera véritablement implantée qu'une fois intégré l'ensemble des éléments structurants préalablement existant ? Quoi qu'il en soit, on est bien face à deux visions différentes de l'évaluation des enseignements avec d'une part, une évaluation formative et normative impliquant une notion de

management de l'institution et, d'autre part, une évaluation sommative et conformative où le contrôle est présent de manière explicite.

Il n'empêche qu'au-delà de cette constatation, les pratiques d'évaluation pédagogique restent paradoxales à plus d'un titre… L'obligation de résultats apparaît en effet souvent comme une reproduction du management de l'entreprise privée ou plus exactement comme une importation du modèle libéral anglo-saxon d'établissements scolaires concurrents, régulés par le marché de la demande parentale (Demailly, 2001). Dans ce contexte, est-ce que les évaluations peuvent encore contribuer à un changement positif du système scolaire ? L'évaluation peut-elle contribuer à l'invention d'un nouveau mode de régulation du système scolaire qui serait participatif, coopératif et citoyen ?

Si la question de l'évaluation de l'enseignement supérieur risque de ne devenir qu'une affaire de contrainte, elle peut aussi être entendue comme une occasion de donner des résultats à comprendre moins comme des jugements sur les personnes enseignantes que comme une appréciation des enseignements dispensés (Collette et Lelard, 2001). Et si l'évaluation de l'enseignement parvient un jour à passer d'une conception individuelle ou individualisante de l'évaluation à une conception collective, alors elle parviendra sans doute à agir sur les attitudes spontanées des acteurs concernés, à changer leur regard sur leur métier et à faire évoluer leur pratique dans le sens de la qualité de l'enseignement (Demailly, 2001).

Références

Abrami, P.C. & d'Apollonia, S. (1990). The Dimensionality of ratings and their use in personnel decisions. Student ratings of instruction: Issues for improving practice. *New Directions for Teaching and Learning, 43*, 97-111.

Aubert, N., & de Gaulejac, V. (1991). *Le coût de l'excellence.* Paris: Seuil.

Bernard, H. (1992). *Processus d'évaluation de l'enseignement : Théorie et pratique*. Laval: Éditions Études vivantes.

Bernard, H., Postiaux, N. & Salcin, A. (2000). Les paradoxes de l'évaluation de l'enseignement universitaire. *Revue des sciences de l'éducation, Vol. XXVI, n°3*, 625-650.

Bloom, B.S. & al. (1956). Taxonomie des objectifs pédagogiques. In *Domaine cognitif*, traduit par M. Lavallée, Éducation Nouvelle, Montréal, 1969.

Braskamp, L.A. & Ory, J.C. (1994). *Assessing faculty work : enhancing individual and institutional performance.* San Francisco: Jossey Bass.

Collette, S. & Lelard, N. (2001). L'évaluation par les étudiants, une valeur légitime ? In M. Lecointe & A. Aubert-Lotarski (Eds), *Evaluation et valorisation institutionnelles*. Toulouse: Presses Universitaires du Mirail.

Delausnay, N., Sylin, M. & Wettendorff, I. (2005). *Les démarches qualité dans l'enseignement supérieur en Communauté française de Belgique : entre normalisation et conformation*. XXIIe congrès de l'AIPU, Genève, 12-14 septembre 2005.

Demailly, L. (2001). Evaluer les politiques éducatives. Bruxelles: Editions De Boeck Université.

Dennery, M. (2001). *Evaluer la formation*. Paris: ESF.

De Ketele, J.-M., & Roegiers, X. (1993). *L'évaluation : approche et démarches*, Louvain-la-Neuve: Diffusion universitaire CIACO.

Doyle, K.O. Jr. (1983). *Evaluating teaching*. Toronto: Lexington Books.

Guilford, J.-P. (1967). *The Nature of Human Intelligence*. New York: Mc Graw Hill.

Hadji, C. (1997). *L'évaluation démystifiée*. Paris: ESF.

Sous la direction de Legendre, R. (2005). *Le dictionnaire actuel de l'éducation*. Montréal: Guérin.

Linn, R.L. & Grönlund, N.E. (1995). *Measurement and assessment in teaching*. Englewoods Cliffs: Prentice Hall.

Marsh, H.W. (1987). Students' evaluations of university teaching : Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11(3)*, 255-388.

Murray, H.G. (1984). The Impact of formative and sommative evaluation of teaching in North American Universities. *Assessment and Evaluation in Higher Education*, *9*(2), 117-132.

Noizet, G. & Caverni, J.-P. (1978). *La psychologie de l'évaluation scolaire*. Paris: PUF.

Poissant, H. (1996). *L'évaluation de l'enseignement universitaire*. Montréal: les Editions Logiques.

Stufflebeam, D.L. & al. (1980). *L'évaluation en éducation et la prise de décision*. Traduction de Jules Dumas. Victoriaville: Editions NHP.

Thivierge, A. et Bernard, H. (1996). Croyances des étudiants à l'égard de l'évaluation de l'enseignement. *Mesure et évaluation en éducation*, *18*(3), 59-84.

Vial, M. (1997). *L'auto-évaluation comme auto-questionnement*. En question. Cahier n°12. 143-200.

Vial, M. (2001). Evaluation et régulations : situation, modélisation et perspectives. In G. Figari, M. Achouche & V. Barthélemy (Eds.), *L'activité évaluative : nouvelles problématiques, nouvelles pratiques*. Bruxelles: De Boeck. 68-78.

MAIS POURQUOI LES ÉVALUATEURS N'UTILISENT-ILS PAS LES GRILLES D'ÉVALUATION?

Marie Noëlle De Theux, Université Catholique de Louvain

Jean Marc Braibant, Université Catholique de Louvain

Benoit Raucent, Université Catholique de Louvain

Pascale Wouters, Université Catholique de Louvain

Thérèse Bouvy, Université Catholique de Louvain

Mots-clés : évaluation critériée, apprentissage par projet, formation d'ingénieur

Résumé:

Les futurs ingénieurs de notre université sont amenés, dès la première année, à gérer des projets. Parmi les modalités d'évaluation, ils ont à les défendre devant un jury de 3 personnes. Des consignes précises et des grilles d'évaluation devraient rendre les évaluations plus cohérentes et plus équitables. Toutefois, si les consignes sont respectées, les grilles de critères sont peu utilisées. Cette question est analysée auprès des différents évaluateurs.

Cette communication prévoit, après une présentation des modalités d'évaluation du projet et de la grille de critères, de se focaliser sur les évaluateurs eux-mêmes : leur démarche au cours des jurys de projet et leur perception de la grille. Nous aboutirons à des propositions d'action pour une meilleure exploitation de celle-ci au cours des évaluations à venir.

COMMENT EVALUER EQUITABLEMENT LES PROJETS ? L'EVOLUTION DES CONCEPTIONS DE L'EQUITE AU SEIN D'UNE GRANDE ECOLE D'INGENIEURS.

Caroline VERZAT, Ecole Centrale de Lille, caroline.verzat@ec-lille.fr

Mots-clés : évaluation individuelle et collective, projets, équité, processus de changement.

Résumé :

Cette communication analyse le changement de dispositif d'évaluation des projets d'une grande école d'ingénieurs française.

En s'appuyant sur l'histoire des modifications du système, l'historique des notes, des observations participantes de jurys et des témoignages d'étudiants, la communication décrit les symptômes du problème puis le changement de conception de l'équité sous-jacent à l'évolution des critères et du processus d'évaluation. Le système passe ainsi d'une conception globale de l'équité à une conception différenciée qui s'accompagne d'une professionnalisation de l'évaluation : outils d'évaluation des jurys plus fins, certification des compétences individuelles avant les jurys, démarches d'auto-évaluation et d'accompagnement des étudiants, formation des enseignants.

La communication analyse ensuite le processus social qui a permis ce changement au sein de l'école. En s'appuyant sur les théories sociologiques de la traduction et des dispositifs de gestion, elle montre que l'évolution des formes du dispositif d'évaluation peut s'analyser comme une traduction. C'est-à-dire que le dispositif d'évaluation focalise les controverses entre les acteurs porteurs de logiques opposées puis enrôle un nombre grandissant d'alliés et de porte-parole qui aboutissent au changement et à la légitimation de ses promoteurs.

La communication débouche sur des perspectives de recherche-action pour étudier la mise en œuvre du nouveau système.

Introduction

La pédagogie par projets se développe fortement depuis une quinzaine d'années dans les formations supérieures d'ingénieur. Les objectifs sont souvent multiples : apprendre à travailler en équipe, acquérir des savoirs et savoir-faire techniques et managériaux, développer un esprit d'initiative, réaliser une production satisfaisant des critères professionnels... Lorsqu'il s'agit de projets pluridisciplinaires en équipe avec un client réel, les critères et le processus d'évaluation deviennent délicats. Comment juger équitablement ? Sur quels critères de justice acceptables par tous les acteurs engagés dans le projet, qu'ils soient étudiants, enseignants des différentes disciplines, client, membre de la direction, techniciens accompagnants…? Faut-il juger le processus et/ou le résultat ? Les différents types de connaissances doivent-ils avoir le même poids dans l'évaluation ? Qui doit évaluer ? Faut-il évaluer le travail du groupe et/ou celui des individus ? Quelle est la place des évaluations formative / certificative ?

Bien que des choix aient été faits dès l'origine en 1992, ces questions se sont reposées progressivement au cours de l'histoire des projets à l'Ecole Centrale de Lille. Chaque année 45 équipes de 5 ou 6 étudiants démarrent un projet d'innovation technologique en partenariat avec un client réel. Ce projet dure 2 années scolaires et

pèse 1/3 de la note globale de la scolarité de 1ère et 2ème année (28 ECTS). Dans ce contexte, le problème de l'équité est devenu prégnant au bout de 8 années de fonctionnement et il a conduit à l'élaboration progressive d'une réforme qui fait évoluer les représentations sur la manière de juger équitablement un projet.

Cette communication cherche à expliquer l'histoire de cette évolution à partir d'une investigation de type sociologique sur la construction sociale de la légitimité du système d'évaluation. Notre investigation prend appui sur l'histoire des modifications du système relue par différents types d'acteurs, l'historique des notes, des observations participantes de jurys, des témoignages d'étudiants. Cette communication mettra successivement l'accent sur trois points : 1) les symptômes du problème de l'équité : qu'est-ce qui fait que le système d'évaluation en place n'est plus perçu comme légitime ? 2) Les conceptions de l'équité en concurrence : quelles différences entre le système initial et le système d'arrivée ? 3) La démarche du changement dans le système d'action de l'école : un processus d'enrôlement et de traduction.

1. Les symptômes du problème de l'équité.

Le déclencheur : la dérive des notes

L'importance du problème de l'équité provient de la conjonction de plusieurs éléments :

- le poids des projets dans le curriculum : les projets représentent 300h par élève-ingénieur sur 2 ans, 28 ECTS. Ils ont un poids énorme dans la scolarité : une mauvaise note globale peut conduire au redoublement et limite considérablement les choix d'option en 3ème année ainsi que les possibilités de départ à l'étranger très prisées par les étudiants,
- le caractère très complexe des apprentissages à évaluer : apprentissages individuels et collectifs, attentes en termes de processus et de résultat, domaines pluridisciplinaires, apprentissages en termes de connaissances techniques mais aussi de méthodes, de savoir-faire et de savoir-être en équipe et avec le partenaire,
- la spécificité des jurys ; chaque équipe de projet ayant un projet différent est évaluée par un jury ad hoc,
- l'échelle de notation sur 3 positions (A-Excellent / S-Satisfaisant / I-Insuffisant).

Le problème de l'équité a été identifié progressivement à partir du constat de la dérive des notes des jurys. En effet, les notes collectives et individuelles attribuées par chaque jury en fin de projet, sont progressivement devenues de moins en moins discriminantes : Au bout de 9 ans, 80% des équipes et des étudiants sont jugés excellents. Les courbes suivantes montrent la dérive progressive des notes.

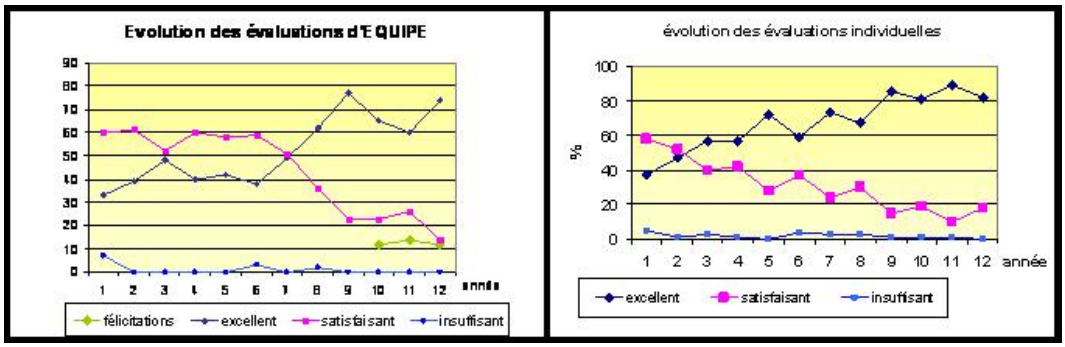


Figure 1 : Evolution des notes des jurys (évaluations collectives et individuelles)

Ceci a conduit à une augmentation significative des contestations sur l'équité de la part des 20% d'étudiants s'estimant injustement sanctionnés : « *Les autres équipes de projet ont eu A, alors qu'on sait bien parce qu'on les voit tous les jours, qu'ils ont beaucoup moins travaillé que nous* », « *C'est parce qu'on a eu un jury plus sévère que les autres* », « *C'est injuste, parce qu'on ne peut plus aller à l'étranger ou bien choisir l'option qu'on veut, tout ça parce qu'on est tombés sur ce jury là*». Cette récrimination est devenue quasi-systématique en cas de S, sauf lorsque l'équipe n'avait pas atteint les objectifs qu'elle s'était fixés et était consciente d'avoir trop peu travaillé.

Par ricochet se sont multipliées les controverses entre les professeurs pendant et surtout après les jurys, créant des débats au sein des professeurs réunis pour les bilans de fin d'année : le niveau S est interprété par certains comme Satisfaisant (ce qu'on attend normalement de l'équipe, c'est bien) et par d'autres comme une Sanction. Les trois niveaux commencent à être contestés alors qu'ils avaient été considérés comme une avancée significative vers la professionnalisation en 1992, proposée par des consultants en management de projet Les manières de juger de certains professeurs commencent à être commentées en aparté tant par les élèves que par les professeurs. Ce qui a des incidences notables sur le choix des encadrants par les promotions suivantes, averties par les élèves des promotions antérieures « *C'est mieux de choisir ce professeur comme encadrant parce qu'avec lui, on a toujours A* » mais aussi sur le choix des professeurs co-encadrants une équipe avec un de leur collègues . « *Tu comprends, je ne vais pas co-encadrer un projet avec X, parce qu'on ne peut jamais avoir raison contre lui au jury* ».

A la recherche des causes de la dérive : les controverses

Plus généralement, les mécanismes de cette dérive conjuguent plusieurs facteurs :

- effet d'amélioration des capacités a l'entrée des étudiants du fait de l'arrivée de pédagogie projets dans les formations antérieures à l'école,
- effet d'apprentissage des étudiants qui se passent des conseils de promotion en promotion,
- effet d'expérience des professeurs qui au fil des années, cernent mieux les difficultés de cette pédagogie, se concertent entre eux et guident mieux les étudiants,

- effets de comparaison entre les jurys qui conduit les professeurs et les étudiants à se situer par rapport à la norme de la promotion.

La teneur des débats au cours des réunions de bilans des professeurs montre une évolution. Dans les années 1 à 7, l'augmentation du niveau des notes est nette les premières années mais n'est pas flagrante par la suite. Les premières années les premiers facteurs (amélioration des capacités des étudiants, effet d'apprentissage et d'expérience) justifient l'augmentation du niveau et confirment le choix de la pédagogie projet, gage de qualité de la formation. Par contre la tendance de la courbe à dériver devient nette à partir de l'année 8, au moment où le nombre de A devient dominant. L'effet de comparaison prend alors de l'ampleur et s'emballe. Il mobilise les tentatives d'explication de la dérive.

A partir de 60% de A, il apparaît difficile de résister à la pression de la norme tant pour les étudiants (on l'a vu plus haut) que pour les professeurs. Les professeurs ayant été plus sévères que la moyenne estiment qu'ils ont été trop exigeants par rapport à leurs collègues. En effet :

- Soutenir une exigence supérieure est difficile à assumer vis-à-vis des étudiants car il est long et douloureux d'annoncer une note perçue comme une sanction. Il faut prendre le temps du feedback afin que l'étudiant en comprenne les raisons objectives et discerne les voies d'amélioration possibles,
- C'est aussi difficile de tenir tête aux collègues en jury en jouant le rôle du « méchant » dans une culture de jury axée sur le consensus au bénéfice de l'étudiant. Certains jurent de ne pas se faire reprendre au piège l'année suivante et d'être plus coulants,
- Enfin c'est toute une posture éthique qui exige un effort de distanciation par rapport à son implication dans le projet, ce qui est difficile dans un accompagnement aussi long et aussi engageant vis-à-vis des partenaires professionnels. Même s'il n'y a pas d'obligation de résultats, la plupart des acteurs (étudiants et enseignants) se sentent engagés. Comme le souligne J.M.Barbier (2000), les partenariats génèrent une généralisation du recours aux ressources cognitives et affectives de l'ensemble des acteurs. Tous sont appelés sous forme d'échanges, de verbalisations, de communications, de gestion d'information, à intervenir sur l'ensemble du processus dont ils sont rendus individuellement et collectivement responsables. Or beaucoup d'enseignants ont du mal à opérer cette distanciation : « *Si « mes » projets ont obtenu de moins bonnes notes, c'est que je suis mauvais…* ».

Il n'y a pas eu de sondage ces années-là pour mesurer le degré de légitimité du système d'évaluation mais il apparaissait assez nettement dans les débats et dans les comportements que les critères de justice et le processus d'évaluation devenaient de plus en plus contestés. Ces controverses ont poussé le responsable de l'activité projet à demander de l'aide à un laboratoire de recherche en éducation, ce qui s'est traduit en année 9 par un travail de DEA avec des entretiens et une observation participante de 12 jurys (Padoani-David, 2002) .Une autre observation participante de 24 jurys a eu lieu en année 12 avec des collègues chercheurs en pédagogie. Les citations et les observations notées dans cette communication sont issues de ces travaux.

En prenant un peu de recul à la lumière d'éléments de littérature sur l'évaluation, on comprend mieux la remise en cause des fondements du système d'évaluation.

Une remise en cause des fondements du système d'évaluation

Les vertus de l'évaluateur telles qu'elles sont énoncées par R. Pardiau, cité par D. Lahaye (2000), semblent de plus en plus délicates à mettre en œuvre dans le contexte des projets en partenariat. (Les partenariats réels avec des entreprises étaient recommandés au départ mais pas systématiques, ils deviennent progressivement dominants jusqu'à atteindre 95% en année 12).

- L'indépendance institutionnelle et financière n'est pas parfaitement réalisée du fait que les partenaires clients participent au jury et que la rétribution des enseignants (en heures de service) est en partie liée au temps effectivement passé,
- L'impartialité pose problème du fait des différences entre les sujets des projets (plus ou moins complexes, plus ou moins appliqués…). Certains critères et points de vue apparaissent plus importants que d'autres selon les projets et les jurys. L'observation en année 9 note que 70% des jurys ont contourné les critères d'excellence,
- La confidentialité on l'a vu, commence à poser problème,
- La transparence est relative car les critères sont assez généraux donc interprétables,
- La non inférence est globalement respectée. Le climat des jurys reste serein dans la très grande majorité des cas. Quelques conflits explicites entre le partenaire et les enseignants émergent à partir de l'année 9. Leur conséquence sur l'évaluation est alors très mal vécue soit par les étudiants (S perçu comme injuste) soit par les enseignants (A contesté),
- L'abstention qui oblige à renoncer à toute velléité opportuniste de conquête d'une place pour soi-même. Tous les évaluateurs ne respectent pas cette règle.

De plus, l'auteur souligne l'application de ces vertus nécessite un large consensus sur le sujet et sur le système d'évaluation en vigueur. Plus loin, Robichon et Josenhans (2004) suggèrent que la crédibilité d'un système d'évaluation est liée au respect des critères de validité (mesure considérée comme pertinente et utile) et de fiabilité (possibilité de reproduire autant que possible les mêmes résultats lorsqu'une modalité d'évaluation donnée est appliquée aux mêmes individus ou groupes d'individus). Or le consensus sur la validité des critères n'existe plus tout à fait et la variabilité des jurys (donc la non fiabilité) fait problème ici.

En se référant aux modèles de l'évaluation de Bonnol et Vial (1997), on peut se rendre compte que plusieurs postures et conceptions de l'évaluation existent selon les différents acteurs des jurys :

- L'évaluation « comme mesure du résultat », où l'évaluateur se pose comme un juge est clairement le fait des partenaires extérieurs, mais aussi de la majorité des encadrants scientifiques pour qui le critère majeur de réussite du projet est le fonctionnement fiable du démonstrateur,
- L'évaluation « comme gestion », se focalise sur les procédures et la plus-value qu'elle apporte à la pratique. Cette conception est plutôt le fait des encadrants en management mais aussi de certains encadrants scientifiques, qui vérifient

l'application d'une démarche rigoureuse par l'équipe. Mais cela nécessite un suivi rapproché de la progression de l'équipe,

- L'évaluation « comme problématique de sens » en émergence, centrée sur la signification multiréférentielle de la question de la valeur de ce qu'on fait. Cette conception existe chez les accompagnants du retour d'expérience, qui sont rarement présents au jury.

Enfin les controverses rejoignent les questionnements sur l'évaluation des projets en entreprise (Picq, 1999). Pour T. Picq, passer de « l'évaluation des performances à l'évaluation des contributions » nécessite quatre dimensions, qui mettent en tension deux dilemmes récurrents dans l'évaluation à l'école :

- S'il est nécessaire d'évaluer la production effectuée en fonction des objectifs fixés, il faut aussi examiner la qualité des processus mis en œuvre (moyens, démarches, méthodes…). Ce dilemme recouvre en grande partie les deux premières logiques d'évaluation ci-dessus.
- Il convient de rendre justice à l'apprentissage collectif (développement de la capacité à travailler ensemble) mais aussi au développement personnel (progression des compétences et satisfaction individuelle). Les enseignants sentent que les élèves n'ont pas tous participé au même degré dans l'équipe. Mais l'évaluation finale individuelle est négligée au jury car les enseignants ne voient pas tout. De fait les outils de suivi en entretiens individuels sont peu appropriés et inégalement utilisés. La plupart du temps le jury ne distingue pas les élèves.

Le choc pour les enseignants provient d'un travail de recherche qui aboutit à un film en année 11 sur la perception a posteriori par les étudiants des évaluations qu'ils ont reçues en projet. Dans ce film, le principal point qui émerge est la grande injustice des notes individuelles qui ne tiennent pas compte des différences d'investissement entre les élèves. En filigrane on sent dans ces témoignages étudiants que la dynamique de groupe est peu accompagnée : « *On a tous eu la même note mais certains ont bossé trois fois plus* » « *On a tous eu la même note, mais c'est pas très juste en fait, car moi par exemple, je ne méritais pas A.* » « *Il y a clairement un problème d'équité. Mais on a préféré ne pas dire que les autres ne bossaient pas pour ne pas rentrer dans le conflit* », « *De fait il faudrait noter l'investissement, mais c'est assez difficile à voir pour les profs, et on se protège : on n'est quand même pas des chiens entre nous* », « *la fiche d'entretien individuel, ça ne sert à rien, en fait, il n'y a pas de suivi. C'est le problème avec les élèves qui ne font rien.* », « *Il y a des fiches, mais on a tous les mêmes notes. De toute façon, les profs sont incapables de noter* ». P. Doucet (2004) montre à partir de son expérience à l'Université de Sherbrooke, qu'il est possible de mieux responsabiliser les étudiants sur leur implication dans la dynamique de projet en instaurant trois dispositifs cohérents : une définition par l'équipe de ses règles de fonctionnement, un accompagnement par des enseignants formés au coaching et un système d'auto-évaluation par les étudiants avec un impact sensible sur les notes individuelles.

Ainsi à partir de 8 ans de fonctionnement et du constat de la dérive des notes, tous les éléments du système d'évaluation se sont trouvés progressivement questionnés par les professeurs puis par les étudiants : les postures des évaluateurs dans le jury, le processus formatif/certificatif et les critères de justice (processus/résultat, équipe/individus). Trois failles principales se font jour : l'expertise et l'impartialité

des jurys est inégale, les résultats atteints sont difficiles à comparer vu la diversité des projets, des « passagers clandestins » profitent de la note du groupe...Les fondements d'un jugement équitable ne sont plus partagés et font émerger une conception alternative. Nous allons maintenant expliquer les différences entre le système initial et le nouveau.

2. Les conceptions de l'équité en concurrence : de « l'équité globale » à « l'équité différenciée »

Pour bien comprendre les conceptions de l'équité, il faut d'abord expliquer brièvement sur quels dispositifs d'évaluation elles s'appuient.

Brève présentation des 2 systèmes d'évaluation

Le premier dispositif a été mis en œuvre entre l'année 1 et 12. Le deuxième a été élaboré à partir de l'année 11 et voté en année 12. Les premières phases de sensibilisation des enseignants et des étudiants ont démarré afin de lancer la mise en place générale à la rentrée prochaine.

Tableau 1. Présentation succincte des 2 dispositifs d'évaluation.

	Dispositif 1	**Dispositif 2**
Evaluation collective	**Certificative** par le jury : 3 niveaux (excellent, satisfaisant, insuffisant) 4 critères généraux 1 note globale résume les 4 critères **Formative :** Indicateurs de suivi d'équipe gérés par l'accompagnant en management	**Certificative** par le jury : 5 niveaux (0 à 4) 5 critères liés aux membres du jury et déclinés par niveaux (grilles de faits précis) Pas de note globale **Formative :** Indicateurs de suivi d'équipe gérés par l'accompagnant en management
Evaluation individuelle	**Certificative** par le jury : 3 niveaux (excellent, satisfaisant, insuffisant) 1 note globale **Formative :** 1 fiche de suivi individuel gérée par l'accompagnant scientifique	**Certificative :** validation finale au jury 5 niveaux (0 à 4) 3 types de compétences à géométrie variable selon les rôles tenus par l'élève dans le projet : - travail en équipe (5 critères tous obligatoires) - gestion de projet (1 à 3 compétences possibles à définir) - expertises techniques (1 à 3 possibles à définir) **Formative :** 3 entretiens de définition + suivi avec les encadrants pertinents pour chaque type de compétence Auto-évaluation sur les compétences de chacun sur le travail en équipe.

Analyse des conceptions de l'équité sous-jacentes.

L'équité est un sentiment de justice qui fait appel à des principes de justice variables selon le contexte, les cultures, les personnes. Dans toute institution éducative, il existe des représentations dominantes, considérées comme légitimes sur les critères d'évaluation justes que sur les postures et le processus qui rendent justice aux étudiants.

Dans le système 1, l'équité est produite par la comparaison finale des résultats des groupes entre eux. Elle est globale.

En ce qui concerne les critères, il s'agit de rendre justice au groupe de projet, par rapport aux autres groupes de l'école, sans rentrer dans la dynamique du groupe, c'est-à-dire la manière dont il produit ses apprentissages et en évitant de différencier les individus. Le critère dominant est l'atteinte du résultat. Une note globale efface les notes intermédiaires sur les critères, c'est cette note globale qui focalise les débats du jury et que les étudiants retiennent.

En termes de processus et de postures des évaluateurs, le travail est surtout certificatif. Le dispositif formatif est très inégal et peu contrôlé. On fait surtout confiance au jury dans son jugement collégial et souverain, conforme aux pratiques de décision délibératives à l'université. Le fait que tous les points de vue arrivent à converger sur un consensus garantit que le jugement est cohérent et juste. Les acteurs dominants du jury sont les encadrants scientifiques. En cas de litige, il est d'usage (comme dans l'ensemble des jurys de l'école) que le consensus se fasse au bénéfice des étudiants, donc sur la note maximale. Les évaluateurs exigeants se sentent déviants.

Ceci est cohérent avec une vision de l'ingénieur spécifique du système français des grandes écoles. C'est l'ingénieur généraliste "qui sait tout faire" et que tous les étudiants sont capables de devenir à quelques exceptions près. La majorité des étudiants peuvent être jugés excellents, parce que le système de sélection à l'entrée est extrêmement sévère (système français élitiste des concours en fin de classes préparatoires), donc ils sont « déjà très bons ».

Dans le système 2, l'équité est produite par l'appréciation fine du résultat spécifique et de la progression de chaque groupe et de chacun. Elle est différenciée.

Au niveau des critères, c'est beaucoup plus objectivé et complexe. On abandonne la note globale. Il s'agit de rendre justice à la manière dont chaque projet atteint un résultat différent, en fonction de ses objectifs liés à la nature du sujet et de la relation avec son partenaire. A chaque critère et à chaque niveau, correspond une définition précise d'objectifs définis par des faits et activités attendus objectivables et mesurables. Mais aussi il faut rendre justice à la manière dont chaque étudiant contribue à l'effort collectif. On met ainsi en lumière les compétences spécifiques et les axes de progression de chaque étudiant, lequel apparaît alors fort différent de son voisin.

Au passage, le processus d'évaluation devient plus spécialisé : Les jurys sont outillés de manière fine (grilles de critères par niveau). Les critères sont liés aux points de vue spécialisés de chaque membre du jury. On attend moins du jury un consensus qu'un équilibre entre des jugements experts sur des domaines disjoints. Les compétences individuelles sont certifiées individuellement par des experts différents en amont du jury. La progression est accompagnée grâce à un suivi individuel au

cours de 3 entretiens dont les objectifs sont différents : 1) définition des compétences, 2) suivi et axes de progrès, 3) évaluation. On fait aussi la place à plusieurs étapes d'auto-évaluation de la dynamique du groupe par les étudiants, autrefois ignorée.

Ceci est compatible avec une vision de l'ingénieur plus différenciée. Les compétences attendues de l'élève ingénieur se traduisent par différentes figures possibles de l'ingénieur. On constate aussi une hétérogénéité croissante des étudiants à l'entrée, notamment du fait de l'intégration d'étudiants étrangers. Cette hétérogénéité se manifeste par des capacités très inégales d'intégration dans les équipes au démarrage des projets et des dynamiques de groupe très variées au sein des équipes. (Verzat, Bachelet, 2006)

Ce changement de logique constaté aussi dans d'autres contextes (Doucet, 2004) s'inscrit dans un mouvement plus large d'évolution vers une culture de la professionnalisation (Barbier, 2003, Lahaye 2001), où l'équité ne peut être garantie que si l'évaluation formative progresse à côté de l'évaluation certificative (Lenoir, 2004).

Il faut cependant bien noter que c'est un changement en cours, loin d'être achevé. Le nouveau système est porteur d'une nouvelle logique, mais il n'est pas encore rentré dans les mœurs. On peut faire l'hypothèse que des ajustements et des compromis entre les deux logiques vont se produire. Cette hypothèse est justifiée par l'analyse de la manière progressive dont ce changement a pu être introduit. C'est le troisième point que nous examinons.

3) La démarche de changement dans le système d'action de l'école : un processus d'enrôlement et de traduction

Dans un premier temps, regardons l'histoire de ce changement. Elle a été rédigée à partir des compte-rendus de recherche-action et des réunions, auxquels j'ai participé depuis l'année 7, puis relue et commentée par 8 enseignants chercheurs représentant différents points de vue dans l'école.

L'histoire du changement.

Etape 1 (année 8) : émergence du problème, début des controverses

La perception de la difficulté commence en année 8 quand 60% des projets sont jugés excellents. Le responsable des projets s'inquiète de la tendance de la courbe. Quelques professeurs critiquent alors le système : Pourquoi avoir seulement 3 positions ? Cette échelle proposée par des consultants a été soutenue par la direction des études parce que jugée plus professionnelle. Ne devrait-on pas revenir au bon vieux système des notes de 0 à 20 ? Les controverses se développent entre les professeurs. Il y a deux tendances : les scientifiques qui aimeraient bien revenir à l'échelle 0-20 et les « managers » qui se réunissent chaque mois pour suivre l'avancement en gestion de projet, et dont un bon nombre estime que ça ne résoudrait aucun problème et que cela reviendrait à revenir en arrière.

Etape 2 : (année 9) recherche-action en sciences de l'éducation – création d'un super-jury sans modification des jurys existants

Une étudiante en DEA en sciences de l'éducation est recrutée pour étudier le problème. Elle aboutit à lever le voile sur le déroulement réel des jurys : quelques rapports de force entre les membres du jury pour faire valoir un critère plus qu'un autre et peser sur la note globale mais le plus souvent consensus sur la note maximale pour éviter les conflits. Face à la difficulté d'objectiver des critères et de définir des règles d'arbitrage entre les points de vue, la solution suggérée par la chercheuse est de proposer la création d'un « super-jury » unique pour attribuer des félicitations aux meilleurs projets de la promotion. L'astuce est de permettre aux jurys de nominer certains projets sans créer vraiment un 4ème niveau. Les jurys peuvent ainsi distinguer certains projets sans pour autant pénaliser les autres.. A la fin de l'année 9, le responsable des projets (issu de l'entreprise) est nommé ailleurs. Il demande à un enseignant-chercheur en management de prendre sa succession.

Etape 3 (année 11) groupes de travail en partenariat avec des experts externes, film des élèves, sondage des enseignants – préparation d'une réforme

Le super-jury mis en place dès l'année 10 est un succès. Les étudiants se prennent au jeu du concours, l'école organise une communication externe vis-à-vis des partenaires d'entreprise et les encadrants des jurys se sentent flattés et non désavoués. Malheureusement au bout de 2 ans, la courbe que les félicitations avaient atténuée continue de dériver.

Le nouveau responsable des projets décide de traiter le problème. Il trouve des soutiens en interne (nouveau directeur des études + jeunes enseignants-chercheurs intéressés par la pédagogie) mais aussi en externe par une collaboration de recherche avec des experts en pédagogie active. Un travail de réflexion conjointe sur les finalités et les modalités d'évaluation des projets se met en place, ainsi qu'un film réalisé à partir de témoignages d'étudiants sur leur évaluation passée. Il met en lumière l'iniquité des évaluations individuelles et la difficulté des professeurs à voir réellement ce qui se passe dans les équipes. Un sondage est organisé au sein des professeurs aboutissant à un consensus sur l'inadéquation des critères actuels. L'ensemble de ces travaux aboutit à des propositions collectives de réforme du système d'évaluation. Cela nécessite une révision du règlement interne des études et examens donc un passage par les instances règlementaires de décision CEVU (Conseil des Etudes et de la Vie Universitaire) puis du CA (conseil d'administration).

Etape 4 (année 12) institutionnalisation, échec puis passage de la réforme,

Une première présentation est faite au CEVU. Mais l'accueil est réservé : quelques scientifiques influents disqualifient certaines propositions. Le directeur de l'école cherchant à ménager les deux parties conclut en repoussant le vote mais en demandant d'aller plus loin dans les propositions. Un facteur d'environnement favorise cette remise à plat générale. La direction souhaite se préparer à l'accréditation européenne. Elle demande de s'ajuster aux échelles d'évaluation du système européen de transfert de crédit (ECTS).

Le CEVU désigne alors un groupe de travail comprenant le responsable des projets, des enseignants-chercheurs volontaires (« managers » et « scientifiques ») et des élèves. Ce petit groupe repense alors l'ensemble du dispositif d'évaluation autour d'une logique de certification de compétences qui sans déposséder les jurys de leurs prérogatives, professionnalise leur travail en amont.

Ayant pris la mesure de l'échec au CEVU précédent, le groupe travaille la communication avec les acteurs concernés. Des mèls constants, des réunions, des tests avec les uns et les autres permettent d'affiner le nouveau système et sa présentation. Le groupe négocie les compromis nécessaires pour faire accepter cette innovation (pas une usine à gaz mais une grille objective facile à comprendre qui va simplifier le travail du jury, enfin 5 niveaux distincts que les scientifiques attendaient depuis longtemps, une vraie évaluation individuelle…etc). De même tous les appuis mobilisables au sein de la direction pour défendre le système sont sollicités afin de mettre au point les supports de présentation et préparer les objections potentielles aux prochaines instances officielles de décision.

A la réunion suivante du CEVU, le système est voté à l'unanimité. La direction remercie les étudiants ainsi que le responsable du département pour cet excellent travail. Le CA vote dans la foulée. Le système peut donc être mis en place. Reste à savoir ensuite comment il va passer dans les mœurs. Les premiers tests auprès de quelques équipes d'étudiants montrent que l'accueil étudiant est réservé au départ, mais plutôt favorable après l'avoir pratiqué une fois car cela aide à échanger sur les facteurs considérés comme essentiels de dynamique de groupe, que les étudiants n'abordaient pas spontanément entre eux. Les premières réunions d'information auprès des enseignants montrent que l'accueil est favorable. Les esprits étaient convaincus de la nécessité de changer. Mais ont-ils toujours perçu ce qui est en jeu ? A la lumière de l'analyse de cette évolution, nous parions qu'un travail de traduction est encore à produire.

Analyse sociologique de l'évolution au sein du système d'action de l'école.

Nous proposons d'analyser le système d'évaluation comme un dispositif de gestion d'une institution universitaire à travers lequel se construisent des rapports sociaux. (Benedetto-Meyer, Craipeau et alii, 2005). Comme le suggère Y. Lenoir (2004), le système d'évaluation est loin d'être neutre, il est idéologiquement marqué des valeurs du groupe dominant. Il rend visible les valeurs et les règles légitimes de la communauté éducative, qui à leur tour fixent le cadre normatif de l'activité des étudiants et des professeurs. Les critères d'évaluation deviennent les « abrégés du bien et du vrai » (Berry, 1983). Le dispositif est vivant, il fait l'objet d'un travail permanent d'argumentation sous forme de critiques ou de justifications de la part des différents groupes d'acteurs pour revendiquer ou asseoir leurs positions (Boussard 2001). Les controverses peuvent pousser au changement si l'équilibre des forces se modifie. Ainsi la sociologie de l'innovation (Alter, 2000) et de la traduction (Callon et Latour, 1986/1988) montrent que les innovateurs réussissent à modifier les pratiques et les représentations en enrôlant de plus en plus d'alliés à travers des dispositifs qui « traduisent » les divergences en une position commune. Lorsque celle-ci est reconnue par l'institution, elle devient légitime.

Le récit oppose deux groupes d'acteurs traduisant des postures d'évaluation, des valeurs mais aussi des pouvoirs et légitimités différentes.

D'un côté les « scientifiques » sont dominants dans l'institution (c'est une école d'ingénieurs). Ils défendent la logique traditionnelle globale d'évaluation. Lors de la mise en place des projets, ils ont dû accepter le système à 3 positions proposé par des consultants en management. Cette innovation leur a toujours paru « bizarre » mais elle a été acceptée car elle se situait dans un contexte de restructuration globale des études imposée par la direction en vue d'une plus grande professionnalisation. Mais les scientifiques sont attachés à la décision collégiale des jurys, conforme à leur pratique de décision délibérative, ainsi que la prééminence de l'évaluation finale certificative, cohérente avec la vision traditionnelle de l'enseignant expert et transmetteur de savoirs. Dans les jurys, ils ont la plus forte légitimité et tiennent le critère dominant (la réalisation technique) ce qui permet généralement de réaliser le consensus autour de leur avis.

De l'autre les « managers » sont marginaux dans l'institution (la gestion est une discipline récente au sein de l'école). Mais ils sont soutenus par la direction (nommés à la tête du département). Ils font émerger la logique différenciée qui leur permet de conquérir leur légitimité. Cette prise de pouvoir (pas forcément consciente et délibérée) se réalise graduellement par enrôlement et traduction dans un nouveau dispositif d'évaluation. Ainsi le premier responsable des projets réalise la courbe, qui fait émerger les controverses et pousse les encadrants présents aux réunions à se situer sur le problème (discussions puis sondage). Ceci permet de susciter des positions plus « managériales » au sein des encadrants scientifiques présents aux réunions. Les deux responsables intéressés par la pédagogie enrôlent le plus possible d'acteurs nouveaux autour du problème : la chercheuse en éducation, des experts d'une autre université (recherche-action puis réflexions croisées), les étudiants (film puis groupe de travail sur l'élaboration d'un nouveau système). Ils organisent les groupes de travail qui font émerger de nouvelles propositions et négocient un système acceptable par tous. Ils profitent notamment de l'opportunité des ECTS pour négocier un compromis avantageux sur les niveaux en contrepartie d'une professionnalisation de l'amont des jurys (certifications des compétences individuelles). Enfin, ils portent l'innovation auprès des instances de décision règlementaires de l'école, et engagent des porte-parole au sein de la direction, nécessaires à l'institutionnalisation.

Conclusion

Nous voyons donc que le changement de système d'évaluation implique un changement profond de valeurs sur le jugement équitable d'une équipe d'étudiants, passant d'une logique globale à une logique différenciée compatible avec l'évolution de l'enseignement vers une « culture de la professionnalisation » (Barbier 2003). Cette évolution portée par un groupe d'acteurs liés aux sciences du management et de l'éducation et soutenue par la direction a été graduellement négociée au travers des modifications du système d'évaluation. Mais l'aventure n'est pas finie. Car les nouveaux outils sont institués mais encore loin d'être appropriés par l'ensemble des membres des jurys et des étudiants. De nouvelles questions se posent :

- Les nouvelles grilles de jury sont maintenant reconnues comme utiles. Mais vont-elles transformer la pratique délibérative habituelle du jury ? Les évaluateurs vont-ils se sentir plus indépendants de l'avis des autres qu'auparavant ?
- Les entretiens obligatoires de définition, suivi et certification des compétences par plusieurs experts transforment totalement la pratique actuelle très inégale (et non contrôlée) des entretiens de suivi individuelle. Or les enseignants-chercheurs ont une disponibilité réduite, n'ont pas de pratique de la définition des compétences, et souvent peu tendance à la différenciation individuelle. La formation prévue va-t-elle être suivie ? Le système va-t-il leur paraître suffisamment simple à mettre en œuvre ? Va-t-il leur permettre de mieux comprendre l'intérêt d'accompagner les dynamiques de groupe à l'œuvre dans les équipes ?
- Les étudiants ont tendance actuellement à ne pas rentrer dans le processus de régulation de la dynamique de groupe et à demander rarement de l'aide aux encadrants par peur des conflits. Cela remet aussi en question leur représentation traditionnelle du professeur plus expert et juge que coach. Le système d'auto-évaluation prévu va-t-il leur permettre de mieux comprendre l'intérêt de cette régulation tant pour leur développement personnel et professionnel que pour la cohésion et l'efficacité de leur projet ? Va-t-il leur permettre de nouer avec leurs encadrants des relations de coaching plus que de contrôle des connaissances ?

Nous postulons à la lumière de l'histoire déjà réalisée qu'une nouvelle étape de traduction sera nécessaire à travers des outils opérationnels (supports du jury, supports d'entretiens, bulletins de notes, indicateurs de suivi…etc) prescripteurs de rôles et d'attitudes en évaluation, favorisant le passage de la logique de contrôle collégial final à une logique d'accompagnement individualisé sur la durée. Nous postulons aussi qu'émergeront de nouveaux acteurs intéressés et enrôlés pour défendre cette nouvelle logique : encadrants intéressés par la pédagogie active, chercheurs extérieurs invités, étudiants intéressés par le développement de l'investissement pédagogique de l'école, partenaires extérieurs intervenant dans la vie de l'école… Ces questions et ces hypothèses serviront de départ à un nouveau cycle de recherche-action.

Références

AKRICH M. CALLON M. LATOUR B. (1988) "A quoi tient le succès des innovations ? » in *Gérer et comprendre*, n°11 et 12

ALTER N. (2000) *L'innovation ordinaire*, Paris, PUF

BARBIER J.M. (2000) Tendances d'évolution de la formation et place du partenariat in *Apprentissages et évaluations, Education Permanente*, n° 143, 2000- 2, 43-54

BARBIER J.M. (2003) « L'émergence d'une culture de la professionnalisation », conférence d'introduction, in *Actes du 2ème colloque Questions de pédagogie dans l'enseignement supérieur*, Brest, 25-26-27 juin.

BENEDETTO-MEYER M. CRAIPEAU S. DARCOURT-LEZAT Y, GUEISSAZ A. METZGER J.L.(2005) « Que pourrait être une sociologie de la gestion ? » *actes de la journée du 11 mars 2005*, atelier issu du 1er congrès de l'Association Française de Sociologie, RT 30.

BERRY M. (1983) *L'impact des instruments de gestion sur l'évolution des systèmes humains*, Centre de recherche en gestion, Ecole Polytechnique, juin

BONNOL J.J. VIAL M. (1997) *Les modèles de l'évaluation*, Paris-Bruxelles, De Boeck

BOUSSARD V. (2001) « Quand les règles s'incarnent, l'exemple des indicateurs prégnants » in *Sociologie du travail*, 43, 533-551

CALLON M. (1986) « Eléments pour une sociologie de la traduction », in *L'année sociologique*, n° 36

DOUCET P. (2004) L'accompagnement des équipes de travail dans la pédagogie projet, in *Res Academica*, volume 22, n°2

LAHAYE D. (2000) « Doit-on professionnaliser l'évaluation ? » in Apprentissages et évaluations, *Education Permanente*, n°143,2, 135-143.

LENOIR H. (2004) « pour une éthique de l'évaluation » in *Education permanente*, n°158, mars, 51-72

PADOANI-DAVID G. (2002) *La formation de futurs cadres en entreprise, un dispositif pédagogique innovant en partenariat*, Mémoire de DEA, Université Charles de Gaulle, Lille 3.

PICQ T. (1999) *Manager une équipe projet*, Paris, Dunod.

ROBICHON M. JOSENHANS U. (2004) Comment évaluer les compétences à des fins de certification ? in *Education permanente*, n°158, mars, 87-98

VERZAT C. BACHELET R. (2006 à paraître) Les ingénieurs deviennent-ils managers grâce aux projets, in *Le métier dans l'organisation : enjeux de connaissances, de compétences et d'identités***,** sous la direction de JC SARDAS. Presses Universitaires de Laval

BESOINS ET QUALIFICATIONS DANS LE CONTEXTE DE LA FORMATION UNIVERSITAIRE DE L'ÉDUCATEUR/ ÉDUCATRICE GRADUÉ(E) ET DANS LE DOMAINE SOCIO-ÉDUCATIF AU LUXEMBOURG

Arthur Limbach-Reich, **Ulla Peters**, **Regina Gaitsch**, Université du Luxembourg

Mots-clés: formation universitaire, qualifications, secteur socio éducatif

Résumé :

Le lancement de la nouvelle formation du «bachelier en science sociale et éducative» (BPSSE) en 2005 a changé fondamentalement le système des formations dans le secteur socio-éducatif. L'étude présentée ici fait le point sur la formation des éducateurs gradués et met en évidence leur situation sur le marché du travail luxembourgeois (domaines d'activités, employabilité et chômage). Par ailleurs, les objectifs sont de: - comparer les compétences reçues lors de la formation avec les compétences nécessaires dans le travail socio-éducatif d'aujourd'hui et d'analyser l'épuisement professionnel des éducateurs et des éducatrices graduées.

1. Introduction et questions

L'accélération du changement social et la croissance de la diversification dans tous les aspects de la vie quotidienne montrent que le secteur social et éducatif est devenu indispensable pour le bon fonctionnement de notre société moderne. Des nouveaux défis sociaux émergent des échanges internationaux, de la mondialisation, des migrations et des changements démographiques, ils appellent de nouvelles réponses comme p.e. maîtriser l'hétérogénéité et éviter les nouvelles formes d'exclusion[1].

Pour le secteur social, ces défis se manifestent dans une croissance des populations cibles et d'une croissance des besoins spécifiques avec en même temps une diminution parallèle des ressources financières. Le plus part des structures et des

[1] Wolfgang Schroer selon Robert Castel qui parle de «neuve zones de la vulnérabilisât (Schröer, 2004).

formations construites au regard d'une situation socioculturelle et économique diverge aujourd'hui de plus en plus des besoins identifiés. Bien que les formations des futurs professionnels dans le secteur social du Luxembourg soient ajustées aux changements de la société, on peut observer un manque d'évaluation systématique de la correspondance entre la formation académique et les besoins du terrain ou des demandes provenant de la pratique[2.]

Pour faire face à cette absence, une équipe de recherche de l'Université du Luxembourg a tenté à l'aide d'une étude pluriannuelle d'analyser les besoins et qualifications du secteur social au Grand-duché. L'article présenté résume les résultats de la première phase du projet (2005 à 2006) comprenant l'élaboration du questionnaire standardisé et l'enquête auprès d'éducateurs gradués. Les questions principales qui ont animé cette recherche sont:

Où travaillent les éducateurs gradués? Dans quels domaines?

Est-ce que la formation reçue répond aux exigences de la pratique?

Quelles activités montrent des différences remarquables entre les besoins du terrain et les compétences et connaissances enseignées dans la formation dispensées aux éducateurs gradués du Luxembourg.

Par ailleurs, quelle est l'ampleur de l'épuisement professionnel ressentis par les éducateurs gradués/ éducatrices graduées? Quelle est la relation qui peut être mise en évidence entre compétences acquises et besoins professionnels?

La population concernée est constituée par les anciens étudiants et par les cadres des institutions sociales.

Nous présenterons tout d'abord la situation de la formation socio-éducative universitaire au Luxembourg, puis les principaux résultats et nous terminerons sur des perspectives pour de futures recherches.

2. La formation universitaire socio-éducative au Luxembourg

La formation universitaire a commencé en 2004 avec la création du Bachelier en sciences sociales et éducatives (BPSSE) à l'Université du Luxembourg.

[2] cf. pour le contexte allemand Schweppe, 2006, pp. 12.

Depuis 1990, la formation de l'Educateur gradué / Educatrice graduée (EG) était dispensée sur trois années (bac + 3), elle est reconnue au titre d'enseignement supérieur en 1997[3]. Mais, par exemple, dans les échelons de traitement salarial, les EG sont classés comme des personnes ayant un niveau de bac +2. Ce classement marque une nette différence avec les travailleurs sociaux ou assistants sociaux formés pour la plus part d'entre eux en France ou en Belgique. La formation EG n'est pas totalement comparable aux formations allemandes (Sozialarbeit, Soziapädagogik Diplompädagogik) par sa durée et son contenu, notamment les spécialisations en éducation institutionnelle et éducation spécialisée. Il en est de même dans les profils descriptifs, les EG parlent d'*éducateur gradué* et rarement de *pédagogue sociale*. Ce décalage terminologique s'explique historiquement.

En effet, l'émergence du profil de l'EG créé en 1990 est apparue à la suite de la formation de moniteur qui datait de 1973. Le fait que ces deux formations aient été dispensées à Fentange/Livange à IEES, au même institut, au même lieu, avec souvent les mêmes professeurs, a été à l'origine d'une réelle absence d'identité.

Les personnes inscrites en formation à l'IEES se définissaient comme des «élèves», aujourd'hui les étudiants de l'Université du Luxembourg en formation BPSSE se définissent également jusqu'à ce jour comme des « élèves ».

Malgré la fondation d'une association pour sauvegarder les intérêts des EG, l'APEG (Association Professionnelle des Educateurs Gradués (Beruffsverband vun de Sozialpädagogen Letzebuerg) créée en 1999, la discussion sur la représentation de l'EG dans le secteur social demeure. L'évolution des inscriptions dans la formation d'EG a commencé avec une moyenne d'environ 30 à 40 «étudiants». Depuis la deuxième moitié des 90 années, le chiffre a doublé et lors de la dernière année, il a atteint 120 étudiants. Au Luxembourg, le nombre de postes d'EG est aujourd'hui selon les estimations de l'APEG au alentour de 1000[4].

Le manque d'identités et la difficile séparation entre EG et éducateur tant sur le terrain professionnel en général qu'au niveau du discours des employeurs et des profils de la formation sont les principaux résultats d'une étude réalisée pour le Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche réalisée par

[3] Cf Haas, Cl. 2005

Ant & Pint 2001. Selon les auteurs, «Comme nous avons pu le constater à travers les différentes analyses que nous avons effectuées dans le cadre de cette étude, le concept de développement des compétences des éducateurs(trices) et des éducateurs(trices) gradué(e)s fait figure de clé de voûte de la problématique de ce secteur. Parmi les multiples définitions du concept de la compétence, celle qui trouverait la place la plus appropriée dans le contexte de cette étude est celle qui consiste à admettre que la compétence *est un ensemble stabilisé de savoirs et de savoir-faire, de conduites types, de procédures standards, de types de raisonnement, que l'on peut mettre en œuvre sans apprentissage nouveau*.»…«d'une part, malgré certaines attentes ou convictions, les formations professionnelles initiales proposées aux éducateurs(trices) et aux éducateurs(trices) gradué(e)s ne se distinguent pas sensiblement l'une de l'autre.» ([5].[6] Ant & Pint, 2001).

Pris comme point de départ, ce manque de délimitation de la formation d'EG montre qu'il était nécessaire de redéfinir la formation universitaire et basée cette formation sur des compétences professionnalisantes, en accord avec les standards internationaux et la situation du Luxembourg.

3. le projet «besoins et qualifications dans le domaine socio-éducatif au Luxembourg»

Au regard des recommandations du processus de Bologne, un projet de recherche a été construit pour évaluer dans les formations la présence des aspects du travail social, de la pédagogie sociale et de l'éducation spécialisée[7]. Ce projet interdisciplinaire de redéfinition des formations en relation avec les besoins des populations est une nécessité pour l'avenir du secteur social. Le texte ci-après présente les résultats

---

[4] Bressler, P. communication personelle 07/09/2006

[5] (Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche, 2001, p. 296)

[6] Ant & Pint se réfère ici sur Maurice de Montmollin qui a utilisé le terme compétence pareille que Weinert: „Kompetenz: Die bei Individuen verfügbaren oder von ihnen erlernbaren kognitiven Fähigkeiten und Fertigkeiten, bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können." (Weinert, 1999, cf. Klieme, 2003, pp. 53).

[7] cf. Symposium „quel avenir pour le secteur sociale? En février 2004

obtenus sur le profil des compétences, l'employabilité et les limites de la pratique professionnelle des éducateurs gradués.

4. Questions et hypothèses

Une des principales missions de notre enquête est de mettre à jour la situation de l'insertion professionnelle des éducateurs gradués sur le marché du travail. Quelques publications (cf. MCESR, 2001, Kurschat, 2002, WOXX 2006[8]) et quelques déclarations marquées du patronage dénoncent une situation difficile pour les jeunes éducateurs gradues à trouver du travail et un chômage certain.

D'un part, de nombreux reproches dénoncent l'inadéquation entre la formation de l'éducateur gradué et les demandes du terrain.

D'autre part, il semble que la profession de l'éducateur gradue comme toutes les professions sociales soit largement touchée par l'épuisement professionnel (burn out). Nous nous proposons d'examiner l'effet des divergences au niveau des compétences enseignées à l'université et des compétences nécessaires au travail de terrain sur l'ampleur de l'épuisement.

Un questionnaire standardisé sous forme écrite et auto-administrée a été élaboré puis distribué en 2005 aux anciens étudiants des promotions 1999 à 2005 ainsi qu'à un échantillon des membres de l'APEG.

4.1 L'échantillon et population

La population interrogée a été recrutée parmi les EG qui ont reçu leur diplôme entre 1999 et 2005 ($N_{prem.}$= 260); seuls 97 questionnaires ont été retournés. Le pourcentage de régression est de **37,31%**. De plus, 48 questionnaires ont été obtenus par l'APEG. Le taux de régression est ainsi passé à 38,36% (soit 145). Ce chiffre représente 17,43% de l'ensemble de la population (900 EG) formée au Luxembourg depuis 1973.

[8] WOXX Article du 03.11.2006 -874, n.n.

Le profil majoritaire de l'échantillon est: 97% de luxembourgeois; 81% sont des femmes; l'âge moyen est de 30,6 ans.

5. Principaux résultats

Trois types de résultats vont être présentés, il s'agit de ceux sur:

- L'insertion professionnelle
- Les qualifications
- Le risque de burn-out (épuisement professionnel)

5.1 Le passage de l'université à l`emploi

La recherche d'un poste de travail dure en moyenne 4,4 mois. Parmi les répondants (19%) ont trouvé, sans aucun délai d'attente, un poste correspond à leur formation d'EG ou équivalent.

Apres un délai de trois mois, 61% ont obtenu un poste, et après six mois, 79% ont été employés comme EG ou comparable. (Figure 1).

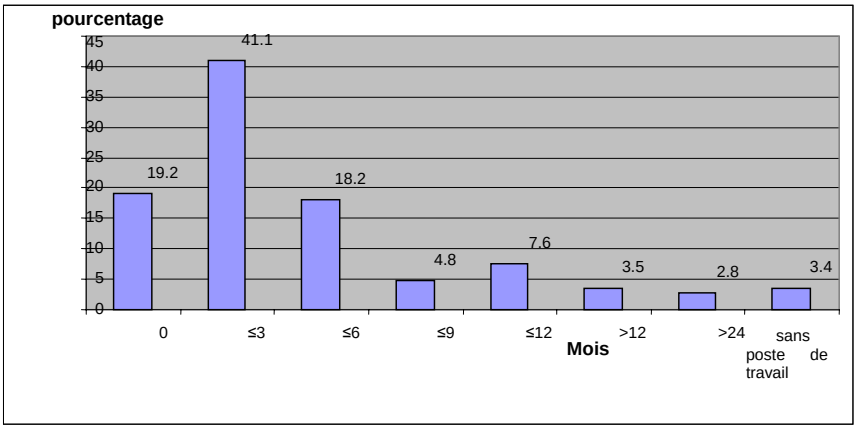


Figure 1 – Durée de la recherche d'un emploi après l'obtention du diplôme

Parallèlement à notre observation, nous savons qu'à la même époque parmi les anciens étudiants, 64% avaient trouvé un poste en tant qu'EG alors que 24 % avaient été classés comme éducateur, 12% en tant qu'assistantes pédagogiques et 3% autres.

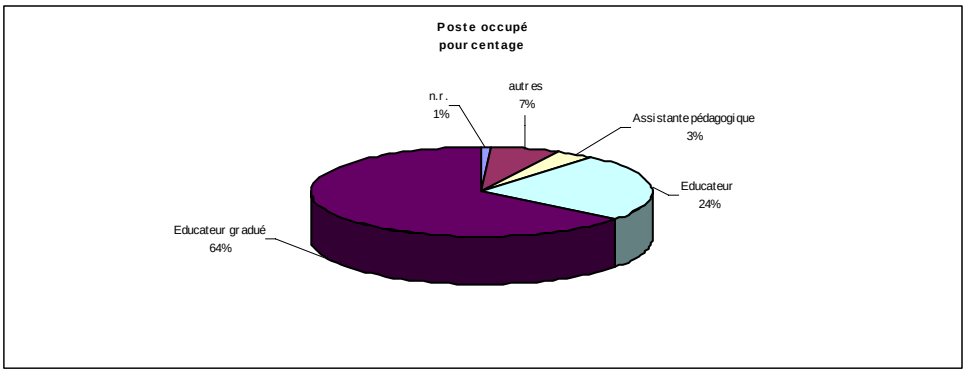


Figure 2 – Premier poste occupé

Parmi les personnes interrogées, 49% ont obtenu un contrat à durée déterminée et 43% un contrat à durée indéterminée (8% non réponses). Au moment de l'enquête en 2005, 62 % des EG travaillaient avec CDI et 19% avec un CDD. Les domaines d'activités professionnelles recensés sont les suivants (Figure 3):

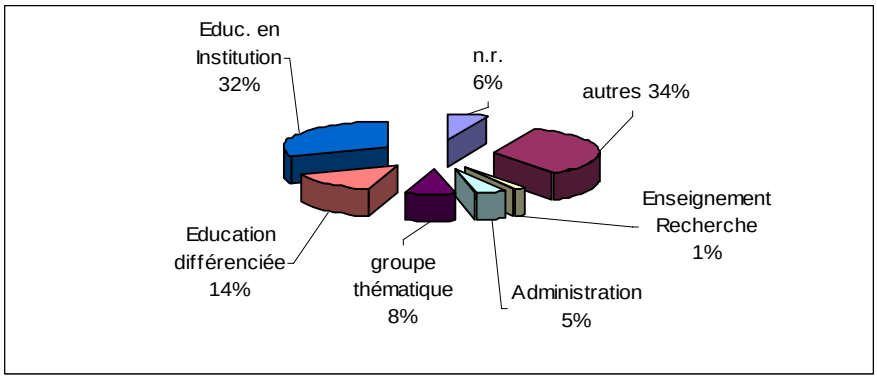
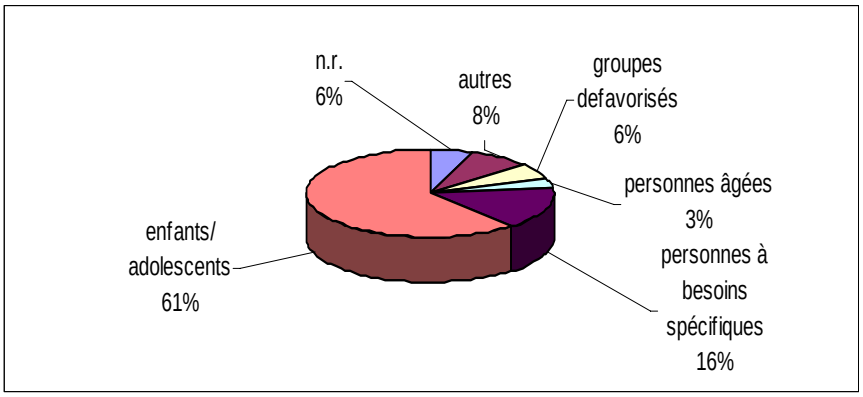


Figure 3 – Domaines d'activité

Les groupes concernés par le travail des EG en particulier: les enfants et les adolescents (61%), les personnes à besoins spécifiques ou personnes handicapées (16%), les populations défavorisées ou marginalisées (5%) et les personnes âgées (3%) (Figure 4)



Figur 4 – Groupe concernés par le travail des EG

5.2 Compétences reçues en formation et besoins socioéducatifs dans les secteurs

Nous avons demandé aux anciens étudiants d'apprécier la pertinence des compétences reçues en formation et celle des compétences nécessaires dans leurs domaines professionnels. Cette auto-évaluation montre un manque de compétence globale[9] (comparaison entre compétences nécessaires et compétences enseignées; x= -0,98). Les différences entre formation versus travail sont significatives (paired sample test t = 13,129 df=105 sig. = 0.00). Les compétences estimées comme les plus nécessaires au travail sont les suivantes:

- planification et programmation,
- interventions auprès d'un groupe;
- gestion d'une équipe,
- communication et réflexion,
- être capable d'être résilient.

Une échelle de notation mesurerait les écarts en 4 points. Si on regarde les compétences les plus enseignées, on observe que ce sont les compétences ancrées dans les disciplines (pédagogie, psychologie, sociologie, et pédagogie spéciale) qui ont obtenu les notes les plus élevées[10]

[9] Reliability: „Competences enseignés" (Come) Cronbachs Alpha =,8673 (N=100) et competences nécessaires (Comn) Alpha = ,8781 (N=94)

[10] Les échelles on prévue continuellement entre 1 à 6 des confirmations à des rejets pour chaque item .

Les différences entre les compétences enseignées et les compétences nécessaires dans le travail

sont importantes même si on regarde les deux plus grandes divisions séparément. $t_{(F)}$=5,673 ;

df=27; Sig.=0,000 et $t_{(T)}$: 9,129; df=54; Sig.=0,000.

Enfin sur les six dernières années, nous avons analysé le développement de la formation au regard de la cohérence des compétences enseignées et des compétences nécessaires (Figure 5)

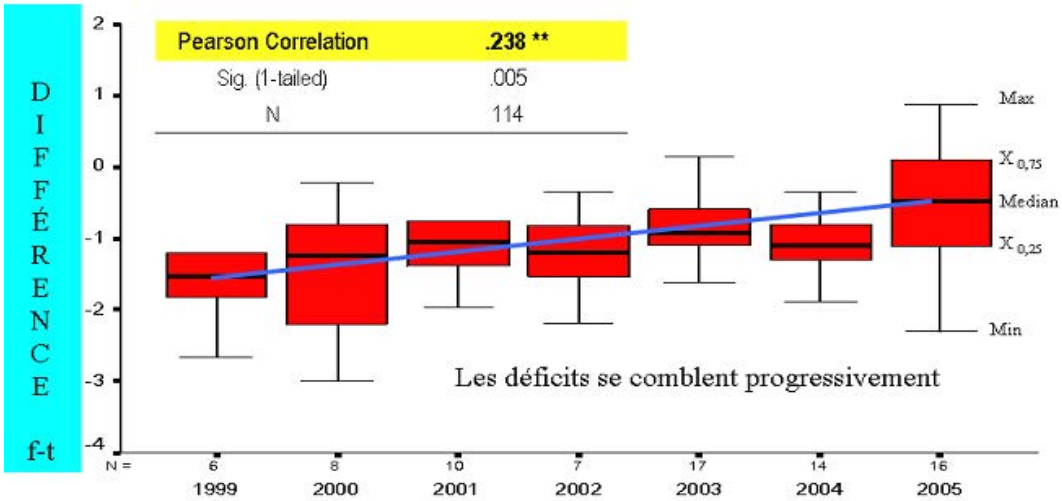


Figure 5 – la diminution des différences entre compétences enseignées et compétences nécessaires

Une analyse factorielle en composantes principales avec rotation varimax (logiciel SPSS) de toutes les variables (30) sur les différences entre les compétences enseignées versus nécessaires à l'exercice du travail a permis d'extraire neuf facteurs avec une valeur propre (Eigenwert) supérieure à 1. A l'aide de la valeur propre obtenue (screeplot) et en tenant compte de la pertinence, nous avons retenu sept facteurs qui expliquent 63,67 % de la variance totale qui (selon les Factor loading) sont: compétence d'agir; compétence personnelle; compétence en situations problématiques, compétence en relation avec la clientèle, compétence à gérer une équipe; compétence en communication; compétence juridique.

Au cours des six dernières années, on observe un développement différencié des compétences suscitées dans la formation de l'EG (figure 6).

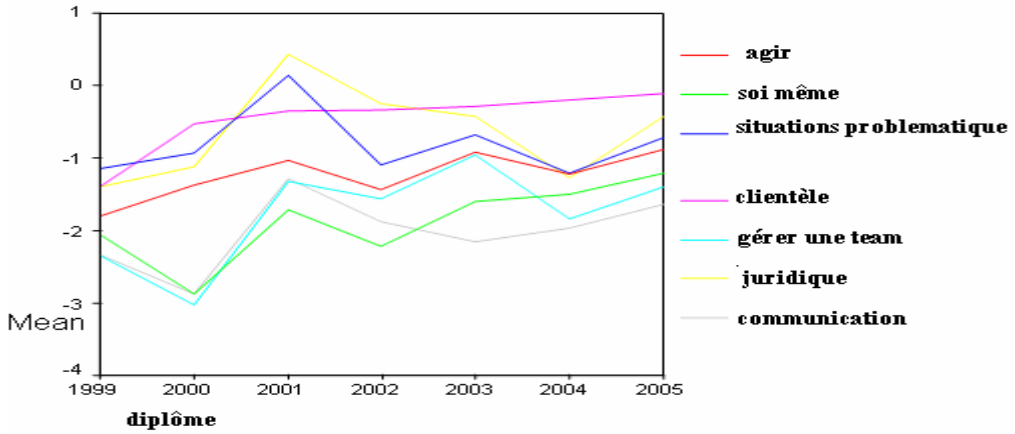


Figure 6 – Développement des compétences entre 1999 et 2005.

Les analyses montrent des variations différentes selon les compétences, mais globalement une amélioration entre formation et travail semble s'ajuster. Au niveau des compétences liées aux relations avec la clientèle, les compétences en pédagogie spéciale et en pédagogie sociale sont nettement en progression. Néanmoins, des différences de genre existent entre hommes et femmes.

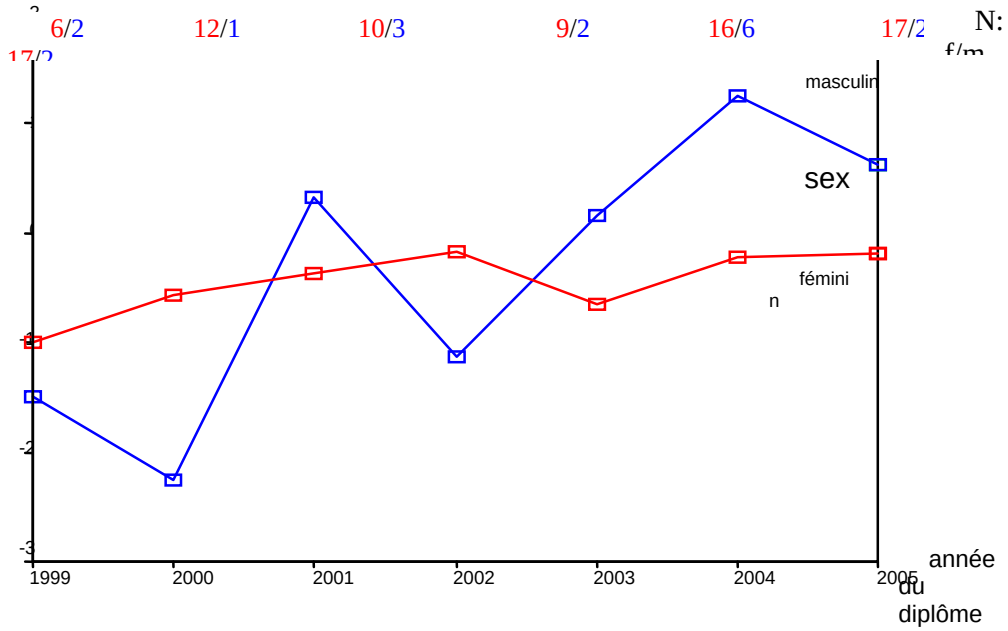


Figure 7 – Compétences auprès de la clientèle entre 1999 et 2005
et différences de genre

Le faible nombre d'hommes appelle une prudence quant à l'interprétation des résultats concernant les différences entre les sexes.

5.3 'épuisement professionnel (Burn-out)

A l'aide du questionnaire, les personnes interrogées ont pu mesurer leurs sensations d'être épuisées par le travail. Pour estimer l'ampleur de l'épuisement professionnel, nous avons utilisé le MBI (Maslach Burnout Inventory, Maslach, 1986). Les valeurs obtenues du «burn-out» ne montrent pas un grand épuisement professionnel chez les éducatrices graduées et chez les EG qui travaillent dans le secteur social et éducatif (voir tableau 2):

Tableau 2 - Données statistiques sur l'épuisement professionnel

	Epuisement emotionelle	Efficacité réduite (inversé)	Depersonalization	Σ Burn-out
N valid	131	120	131	112
Missing	14	25	14	33
Mean	**10.81**	**34.03**	**4.70**	**7.86**
Std. Deviation	7.66	6.82	4.10	1.85
Minimum	.00	2.00	.00	4.22
Maximum	40.00	48.00	19.00	13.90
Mean Maslach overall sample 1986	21.0	34.6	8.7	9.15
Mean SIB Limbach-Reich 2000	15.7	29.9	4.8	9.06

Les valeurs obtenues indiquent un surmenage professionnel global, qui comparée à des valeurs de référence issues d'études similaires de la littérature sont plus faibles.

Pour clarifier les liens entre la formation et l'épuisement professionnel, toutes les dimensions de la MBI ont été examinées à l'aide de corrélations. Des manques de compétences liées aux différences entre compétences enseignées et compétences nécessaires au travail ont été mis en évidence. L'épuisement émotionnel présente une corrélation marquée. (voir tableau 2) .

Tableau 3 - Correlation „épuisement émotionnel" avec manque de compétences

Burn-out-dimension		manque de compétences
épuisement émotionnel	Pearson Correlation	-.232*
	Sig. (1-tailed)	.010
	N	92

* Correlation est significative au seuil de 0.05.

Le signe négatif veut dire ici que le score d'épuisement émotionnel est bas si le manque des compétences augmente et inversement. La corrélation avec le manque de compétences enseignées (r=.269**; N=107; Sig.=.003) est très important. Un rapport stable existe entre la formation et le burn-out.

6. Bilan et résumé

Les résultats de l'étude présentée ici montrent qu'il est possible d'avoir une image comparée entre les compétences enseignées lors de la formation reçue et les compétences nécessaires sur le travail à partir des déclarations d'anciens élèves de la formation des EG gradue au Luxembourg.

En résume:

- Ces dernières années, on a constaté que les demandes du terrain en EG diplômés au Luxembourg sont plus importantes que les offres.
- Six mois après l'obtention du diplôme, presque 80% des EG ont trouvé un poste d'EG en adéquation avec le diplôme obtenu, parmi eux 20% n'ont jamais été à recherche d'un emploi. Ce qui est le garant d'une employabilité correcte.
- Seuls 27% déclarent avoir occupé un emploi ayant un niveau de qualification inférieure au diplôme reçu (e.g. comme éducateur). Mais cette situation a eu à une durée limitée qui n'a pas dépassé plus d'un an. Plus de 60% des EG ont obtenu un CDI (contrat à durée indéterminé).
- Conclusion: la situation de l'emploi pour les EG ne correspond pas à la situation problématique décrite par les media au Luxembourg
- Les domaines de travail, les institutions et les clientèles sont variés. Plus d'un tiers des EG interrogés travaillent dans des institutions et des centres d'accueils pour des enfants, avec des personnes handicapées et enfin dans des administrations en tant que chargés de mission. Jusque à présent, sont les EG qui travaillent avec des personnes âgées.
- Malgré la situation confortable de l'emploi des EG, les auto-évaluations des anciens EG ont montré qu'il y avait des carences entre les compétences

enseignées et les compétences nécessaires au travail, mais ce manque diminue peu à peu.

- Pour faire face aux enjeux actuels, les compétences souhaitées sont des compétences de bases ou transversales qui permettent de structurer et de programmer un projet. Les compétences socio pédagogiques et pédagogiques spécialisées restent très appréciées dans la formation.
- Concernant l'épuisement professionnel, le risque pour les EG est faible, une formation aux compétences nouvelles peut être une solution pour remédier à ce burn out.

Cette étude est à poursuivre en respect la structure de la formation universitaire imposée par le cadre de la reforme de Bologne, mais il faut y ajouter des approches plutôt qualitatives afin de mieux cerner les demandes et les attentes des étudiants à l'égard d'une formation universitaire. Par ailleurs, quels sont les nouveaux besoins des professionnels du secteur socio-éducatif pour apporter des prestations de services de qualité et des réponses adaptées aux changements sociodémographiques.

Références

APEG (2001). http://www.apeg.lu/docs/dossier30012001.pdf#search="apeg chiffres", 20.08.2006.

Deutsche Gesellschaft für Soziale Arbeit (2005): *Kerncurriculum Soziale Arbeit / Sozialarbeitswissenschaft für Bachelor- und Masterstudiengänge in Sozialer Arbeit, vorgelegt von der Deutschen Gesellschaft für Sozialarbeit (DGS) am 28.01.2005 erarbeitet von einer Arbeitsgruppe der Sektion „Theorie und Wissenschaftsentwicklung in der Sozialen Arbeit" mit den Mitgliedern Prof. Dr. Ernst Engelke, MSW Manuela Leideritz, Prof. Dr. Konrad Maier, Prof. Dr. Richard Sorg, Prof. Dr. Silvia Staub-Bernasconi.*

Haas, C. (2005). Ausbildung für soziale Berufe in Luxemburg. In F. Hamburger, S. Hirschler, G. Sander & M. Wöbcke (Ed.), *Ausbildung für Soziale Berufe in Europa.* Frankfurt am Main: ISS-Verlag.

Kuchartz, U., Lukas, H. & Skiba, E.-G. (1994). *Sozialpädagogisches Hochschulstudium und Berufstätigkeit. Am Beispiel der Absolventinnen und Absolventen des Dipomstudienganges an der FU Berlin.* Berlin: Walter May.

Kurschat, I. (2002). Das Warten geht weiter. *Woxx*, 18.01.2002, p. 624.

Labonté Roset, C. (2004). *Social Work Education and Training and the Bologna Process. Social Work & society*, Vol(2), 98-103.

MCESR (Ministère de la Culture, de l'Enseignement Supérieur et de la Recherche) (2001). *Etude sur les demandes en qualification du secteur socio-éducatif, réalisée par Marc Ant et Manon Pint.* Luxembourg: MCESR publication.

Schroer, W. (2004): *Soziale Arbeit im Übergang zum digitalen Kapitalismus*. in Bock K./Thole W. (Ed.) (2004): *Soziale Arbeit und Sozialpolitik im neuen Jahrtausen*d. Wiesbaden: VS-Verlag. p. 101-115.

Schweppe, C. (2006): *Studienverläufe in der Sozialpädagogik, Biographische Rekonstruktionen*. Weinheim, München: Juventa.

L'ÉVALUATION DE LA PORTÉE DE DISPOSITIFS INNOVANTS DE FORMATION CONTINUE D'ENSEIGNANTS DU PRIMAIRE ET DU SECONDAIRE SUR LE DÉVELOPPEMENT, LA MOBILISATION ET LE TRANSFERTS DE LEURS SAVOIRS ET COMPÉTENCES PROFESSIONNELLES

Annie Presseau, Stéphane Martineau, Université du Québec à Trois-Rivières

Problématique

Les attentes à l'endroit des enseignants sont de plus en plus élevées. Conséquemment, il apparaît essentiel de s'assurer que la formation qui leur est dispensée génère effectivement les résultats attendus (Bidjang et al, 2005). Depuis 2001, au Québec, la formation initiale des maîtres est marquée par un référentiel de compétences (MEQ, 2001), dressant un portrait précis de l'éventail de compétences que doit maîtriser tout enseignant et du degré de maîtrise attendu pour chacune des compétences. Se pose alors la question des moyens employés pour évaluer ces compétences. Selon Cauterman, Demailly, Suffys et Bliez-Sullerot (1999), la formation continue des enseignants est, pour sa part, mal évaluée. Trop souvent, le seul critère retenu est le degré de satisfaction des bénéficiaires de la formation ou encore l'intérêt qu'ils manifestent à la poursuivre. Une enquête sur la formation continue d'enseignants québécois (Laurin et al., 1999) révèle que ces derniers se disent généralement satisfaits des formations reçues. Si la satisfaction des participants peut contribuer à juger de la qualité d'une formation continue, il n'en demeure pas moins que cet élément, à lui seul, est nettement insuffisant pour attester de l'efficacité d'une formation. Sous-jacente à cette dernière, il y a l'idée que l'enseignant puisse construire des savoirs et développer des compétences, les mobiliser et les transférer. Or, qu'en est-il lorsque le soutien s'estompe puis disparaît? Cette question ouvre sur la nécessité d'envisager le recours à divers outils susceptibles d'informer sur les savoirs construits ainsi que sur les compétences transférables développées. À l'instar de Lebrun et al. (2005), nous déplorons le nombre restreint de recherches portant sur les pratiques réelles des enseignants, lesquelles nécessitent forcément la mobilisation et le transfert de savoirs et de compétences, par rapport au grand nombre de travaux qui rendent compte de pratiques déclarées, qui fournissent, à elles seules, un assez faible indice de la maîtrise de compétences professionnelles.

Objectif

Cette communication vise, d'une part, à décrire les outils d'évaluation que nous avons utilisés dans le cadre de dispositifs innovants de formation continue d'enseignants visant le développement, la mobilisation et le transfert de savoirs et de compétences professionnelles et, d'autre part, à poser un regard critique sur ces outils et à discuter des enjeux relatifs à la mobilisation des savoirs et des compétences et au transfert.

Méthodologie

Le type de recherche privilégié est la recherche-action. Celle-ci est considérée par plusieurs chercheurs comme une stratégie de développement professionnel importante et efficace (Guskey et Huberman, 1995; Savoie-Zajc et Dolbec, 1997). Quatorze enseignants (onze du primaire et quatre du secondaire) provenant de trois écoles ont participé à la recherche. Afin d'avoir accès aux pratiques déclarées et aux pratiques observées des participants, mais aussi aux savoirs en construction, un dispositif combiné d'outils a été utilisé: a) quatre entrevues semi-dirigées individuelles d'une durée d'environ une heure trente : une première lors de la phase diagnostique, deux pendant la recherche et une au terme de la recherche; b) cinq périodes d'observation non participante en classe, d'une durée de soixante-quinze minutes, à partir d'une grille co-élaborée avec les enseignants; c) des séances de rétroaction et d'explicitation à la suite des observations effectuées, d'une durée approximative d'une heure; d) des entrevues de groupe (focus group) (n= 4), e) l'élaboration ou la co-élaboration de réseaux conceptuels; f) les planifications; g) le matériel pédagogique développé ou utilisé par les enseignantes (dont le nombre varie entre cinq et vingt documents selon les enseignantes) ainsi que h) les enregistrements des discussions qui ont eu lieu lors des rencontres de travail (entre huit et quinze, selon l'école) au cours desquelles la co-élaboration ont été privilégiées des situations d'apprentissage requerrant, de la part des participants, la justification des choix envisagés. Le choix de ces outils s'inspire notamment des travaux de Allal (1999), Charlier et Charlier (1998), Durand (2000), Fullan, Galluzzo, Morris et Watson (1998), Guskey (2000) et Haskell (2001).

Résultats préliminaires

Certains outils paraissent particulièrement pertinents pour se prononcer sur la maîtrise de compétences. Tel est le cas des observations, qui plus est lorsqu'elles sont faites à partir d'une grille co-construite avec les participants. En revanche, la multiplicité des contextes d'observation ressort comme étant essentielle pour se prononcer sur le caractère transférable de ces compétences. Par ailleurs, l'observation semble insuffisante pour conclure avec certitude du degré de maîtrise de compétences, d'où la pertinence de recourir à des outils complémentaires, permettant de valider les perceptions. Par exemple, il aurait parfois été tentant de conclure à un faible degré de maîtrise de certaines compétences en nous limitant à l'observation, alors que l'explicitation des intentions sous-jacentes aux choix posés nous amène à tirer des conclusions beaucoup plus nuancées. À l'inverse, certains gestes professionnels pouvaient sembler illustrer un haut niveau de maîtrise de compétences alors qu'une explicitation des motifs à la base des choix s'est parfois révélée décevante, témoignant d'une intégration minimale de savoirs. Le caractère transférable des compétences s'illustre principalement grâce aux occasions multiples de mobilisation des compétences, mais également par le discours des enseignants, lors d'entrevues ou de rencontre de travail, où ils ont l'occasion de justifier les décisions qui président à leurs choix pédagogiques, de préciser les conditions de réutilisation. L'analyse de l'évolution des planifications, élaborées individuellement ou co-élaborées dans le cadre de la recherche-action fournit également des indices précieux sur le développement de l'autonomie professionnelle des participants, au fur et à

mesure que le soutien s'estompe. Finalement, le réseau conceptuel ressort comme étant un instrument puissant pour illustrer les changements conceptuels opérés en cours de formation. En revanche, il informe peu sur les compétences maîtrisées, mais, comme le mentionnent les participants, ils constituent un outil pertinent pour « faire le ménage » dans ses savoirs, ce qui peut s'avérer utile à la mobilisation de ces ressources à l'intérieur de compétences. Ce qu'il apparaît crucial de retenir, c'est la complémentarité des outils et la pertinence de recourir à ceux-ci sur une base récurrente, de manière à fournir un maximum d'indices des processus en jeu.

Quelques références

Bidjang, S., Gauthier, C., Mellouki, M. et Desbiens, J.-F. (2005). Les finissants en enseignement sont-ils compétents? Une enquête québécoise. Sainte-Foy : Presses de l'Université Laval.

Cauterman, M.-M., Demailly, L., Suffys, S. et Bliez-Sullerot, N. (1999). La formation continue des enseignants est-elle utile ? Paris : Presses Universitaires de France.

Charlier, É. et Charlier, B. (1998). La formation au cœur de la pratique. Analyse d'une formation continuée d'enseignants. Paris, Bruxelles: De Boeck Université.

Durand, M. (2000). Quand se tissent l'offre de formation et l'expérience des enseignants-stagiaires… Dans Carlier, G., Renard, J.-P. et Paquay, L. (dir.). La formation continue des enseignants (47-67). Bruxelles : De Boeck & Larcier.

Fullan, M. Galluzzo, G. Morris, P. et Watson, N. (1998), The Rise and Stall of Teacher Education Reform, American Association of Colleges for Teacher Education, Washington, Dc.

Guskey, T. R. (2000). Evaluating Professional Development. Thousand Oaks, CA: Corwin Press.

Ministère de l'éducation (2001). La formation à l'enseignement. Gouvernement du Québec : Québec.

ETAT DES LIEUX DES PRATIQUES FORMELLES ET INFORMELLES DES ACTEURS DE LA FORMATION.

Leclerc Claude, enseignant chercheur, IUFM Créteil, laboratoire ESCOL-ESSI

Mots clés : évaluation interne, approche coopérative de l'observation, formateurs réflexifs.

Résumé :

L'observation de la formation implique la construction d'un regard aux multiples focales. Evaluations externe et interne constituent ainsi des distances spécifiques par rapport à l'observation des formations. En effet les pratiques développées de l'évaluation, qu'elles soient internes ou externes réfèrent à des échelles (régionale, nationale, internationale) et porte sur des objets (établissement, plan de formation, etc.) dans lesquels les acteurs se trouvent peu ou pas associés alors même qu'ils constituent un objet essentiel de la mesure. Cette distance construite a pour effet de désimpliquer (Vial, 2000) les acteurs au regard notamment de l'adaptation attendue. Il devient dés lors nécessaire d'explorer les possibilités d'une évaluation (participative, collaborative, etc.) qui permettrait de réintroduire les acteurs dans le processus évaluatif comme condition des possibles transformations.

La présente communication vise à analyser les conditions de mise en œuvre d'un observatoire des formations en Iufm. Il s'agit, plus particulièrement, de mener une interrogation, à partir notamment de l'expérience engrangée à l'Iufm de Créteil, sur la position et la stratégie d'un observatoire des formations en institut de formation des enseignants. Deux lignes d'action se sont progressivement dessinées. La première dominante dans les travaux des observatoires existants s'inscrit dans celle développée par l'évaluation interne ou « auto-évaluation » recommandée par les multiples incitations ministérielles et européennes. Celle-ci relève désormais peu ou prou du livre des références (2003) préconisé par le comité national d'évaluation fixant désormais un cadre précis. La seconde, l'évaluation participative ou collaborative, tendrait à associer les acteurs de la formation (formés et formateurs) et à développer une démarche plus implicative dans le processus évaluatif. Elle se détache de la première par un certain nombre d'éléments que nous tenterons de mettre en évidence. Elle conditionnerait, dans une large mesure, la qualité des indicateurs susceptibles d'être construits et peut-être plus encore le degré d'implication des acteurs dans l'approche évaluative et les transformations des pratiques (Bressoux & Pansu, 2003).

Par ailleurs, l'évaluation des formations intervient dans un contexte général d'amélioration de la qualité en éducation, inscrite dans le processus de Lisbonne avec réalisation des objectifs fixés à l'horizon 2010. L'étude européenne sur l'évaluation de la qualité de la formation des enseignants en Europe (*L'assurance qualité de la formation des enseignants,* 2006) réalisée à la demande de la commission européenne par Eurydice, organisme européen de collecte d'informations sur l'enseignement,

confirme la montée en puissance de l'évaluation tant externe qu'interne des formations des enseignants parallèlement aux développements de nombreux référentiels et standards nationaux et internationaux. A terme l'objectif sous-jacent serait de favoriser l'articulation étroite entre évaluation externe et évaluation interne avec l'enjeu d'une possible subordination du cadre national au cadre international. Les auteurs du rapport confirment, en substance la propension repérée à une sur-réglementation allant en ce sens. Cette dernière se trouve, en quelque sorte, relayée par une pratique soutenue des standards dans le champ de la formation (Behrens, 2005) au regard de laquelle l'acteur est perçu à travers des formalisations de plus en plus élaborées. L'éventualité d'une inversion de la démarche à partir d'une conformité de la réalité aux standards eux-mêmes (De Ketele, 1993), bien que paradoxale, semble alors effectivement envisageable.

L'évaluation des formations ne porte pas en elle-même toutes les qualités ou les défauts qu'on lui prête selon les positions adoptées. Les objectifs, qui lui sont assignés, déterminent amplement ses possibles usages et par conséquent l'accueil qui lui sera fait par les différents acteurs concernés. Certaines pratiques institutionnelles, par trop descendantes, ont pour effet de renforcer les acteurs de la formation dans la distance adoptée à l'encontre des questions de l'évaluation alors même que cette dernière représente, d'un point de vue plus général, une véritable possibilité d'adaptation de la formation aux conditions de l'exercice du métier d'enseignant.

Un enjeu : l'acteur co-observateur

L'expérience engagée à l'Iufm de Créteil tend à explorer des formes d'implication des acteurs de la formation dans la démarche proposée. L'objectif consiste à favoriser, à l'intérieur même de l'institut, le glissement progressif de l'approche socio-organisationnelle dominante, vers une approche davantage socio-interactionnelle. Dans la première configuration, les acteurs tendent à renforcer leurs positions dans l'institution. L'expertise intervient fréquemment comme une justification de la décision. Elle permet en quelque sorte le dernier mot. Peu ou pas d'échange de savoirs interviennent alors. Les acteurs développent des stratégies d'évitement plutôt que d'adaptation (coordonnateurs, formateurs, tuteurs, conseillers, chef d'établissement, etc.) en vue de préserver les positions instituées. La convergence vers l'objectif transite essentiellement par les jeux institutionnels habituels. Les tensions générées, en raison notamment de l'évaluation implicite pratiquée par défaut, (figure 1), sont relativement fortes et renvoyées de surcroît vers les acteurs de la formation.

La seconde configuration favoriserait la mise en œuvre d'un modèle reconsidéré de l'organisation apprenante (Mallet, 1994). Les acteurs alors impliqués dans une telle interaction sont susceptibles de déplacer les positions instituées et de pratiquer, par adaptation aux situations rencontrées, l'échange des savoirs et des informations. Les rôles antérieurement dévolus sont travaillés par les dynamiques socio-interactionnelles (Morgan, 1999). La question d'une éthique de l'évaluation, sous-

jacente à toute situation d'analyse de la qualité, se trouve alors posée de manière co-constructive et non uniquement contractualisée. Cet aspect est décisif dans la mesure où l'évaluation de la qualité contribue, par certains aspects, à obscurcir les limites des rôles tenus. Chercheur, expert, consultant et acteurs divers de l'évaluation développent des postures voisines susceptibles de brouiller les enjeux véritables (Ardoino, 1990). Les tensions ne sont plus renvoyées vers les seuls acteurs, notamment en fin de formation, mais s'inscrivent dans un processus de redéfinition de l'offre de formation à partir des effets de retour opérer à froid (figure 1, deuxième configuration).

Les deux configurations exposées ci-après (figure 1) traduisent les enjeux et tensions entre deux approches de l'évaluation d'un établissement d'enseignement supérieur à forte visée professionnalisante. La première configuration (1,2,3) se décline du cadrage de la formation vers sa formalisation par le plan de formation, jusqu'à sa mise en œuvre pour s'achever à ce stade. Les régulations consistent, pour l'essentiel, à favoriser l'efficace de l'action entreprise. Les évaluations effectuées se centrent pour l'essentiel sur les formés et interviennent dans un cadre à dominante certificative.

La seconde configuration a pour effet de favoriser une implication des acteurs (formateurs/formés) à toute les étapes, hors celle du cadrage par ailleurs susceptible de subir lui-même quelques aménagements en rapport avec le positionnement de l'institution à son encontre. Le cadrage appartient à la sphère du politique au sens ou il s'agit effectivement de choix dégagés à partir du jeu démocratique. Les régulations s'exercent à différents niveaux mais également à partir d'une échelle des durées facilitatrice des « retours à froid » sur la formation dispensée. Les outils développés ont sur ce point un rôle décisif tant en terme d'implication des acteurs que de pertinence des données recueillies (entrée dans le métier, stabilisation de la pratique, contextes professionnels, etc.). L'ensemble de l'action de formation s'inscrit dans un processus favorisant les conditions d'une méta-évaluation. La mise en œuvre de cette dernière (Demailly, 2000) intervient comme le baromètre de la qualité méthodologique, éthique de l'évaluation elle-même.

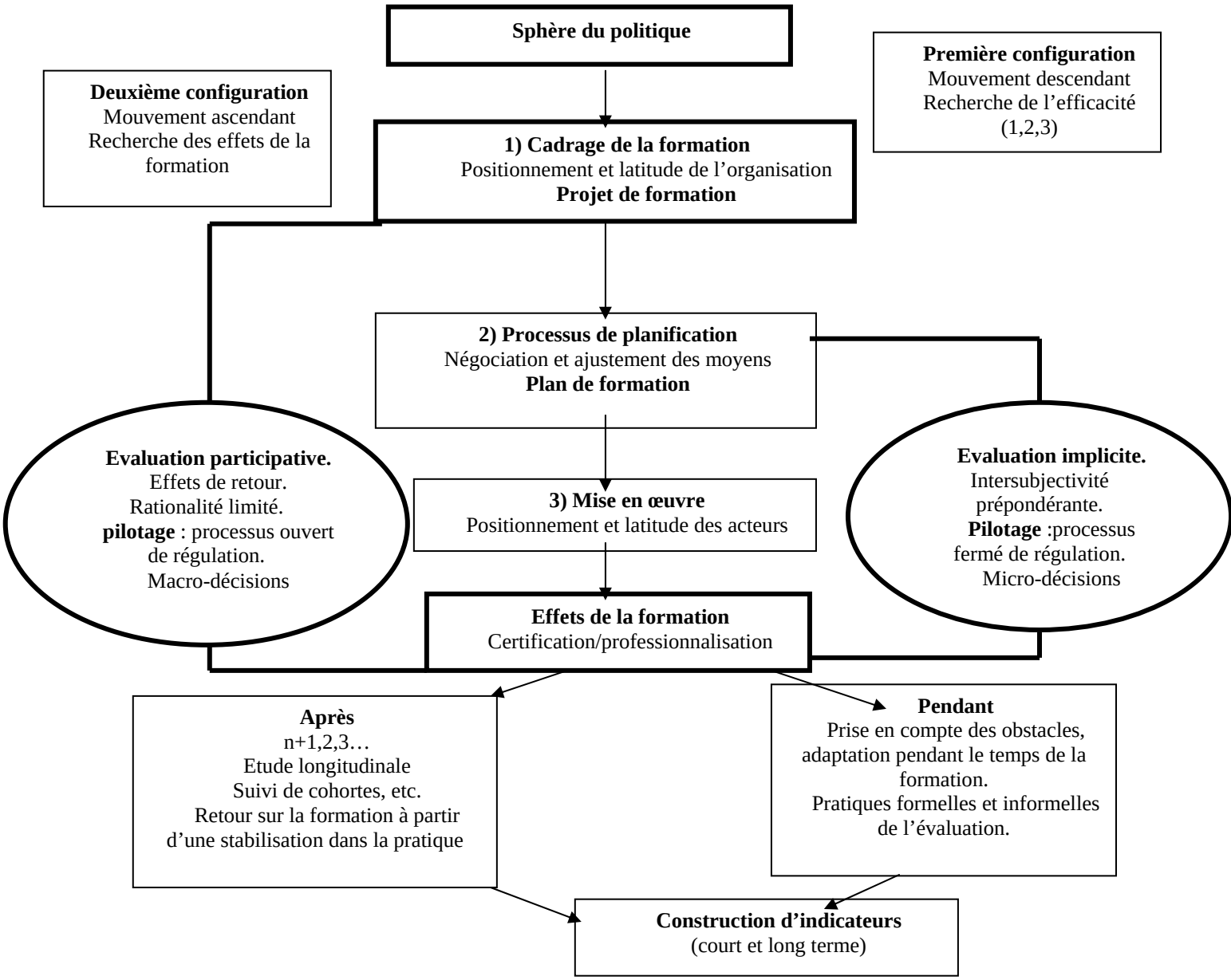


Figure 1 - L'organisation apprenante : entre approches socio-organisationnelle et socio-interactionnelle

Observer : entre formation et professionnalisation

L'observation de la formation nous semble devoir intégrer deux composantes décisives, en substance le contexte environnementale et celui de la formation elle-même. La centration sur la seule qualité de la formation au regard d'un cadrage programmatique (figure 1, première configuration) conduit inexorablement à ne rechercher que les éléments susceptibles d'accroître l'efficace de celle-ci tout autant qu'à circonscrire les éléments contraires. La qualité de la formation peut dés lors être envisagée sous deux angles distincts : celui d'une conformité la plus accomplie possible aux différents cadrages de la formation et ses multiples déclinaisons (cadrage ministériel, plans de formations, contenus disciplinaires, etc.) ou celui d'une adaptation aux différents contextes d'enseignement possibles (collège/lycée, rural/urbain, centre/banlieue, établissement ordinaire/établissement catégorisé, etc.). Nombre d'enseignants stagiaires mettent l'accent sur cette inadaptation de la formation aux réalités rencontrées dans la classe (Larivain, 2006). Ce rapport, après d'autres études, confirme l'intérêt porté par les enseignants en formation aux dimensions contextualisées du métier notamment à partir des stages en situation. Les stages occupent un espace, en termes d'implication, relativement considérable chez le futur enseignant et se trouvent souvent opposés à l'ensemble de la formation. La traduction d'une telle situation par le couple irréconciliable théorie/pratique renverrait à ce paradoxe d'une formation fortement centrée sur ses propres dispositifs et peu à une problématique plus globale des évolutions sociétales en termes d'adaptabilité aux réalités du métier. Lors de ces derniers mois nous avons rencontré, au titre de l'observatoire, un grand nombre d'acteurs de la formation. Par ailleurs nous nous sommes engagés avec un certain nombre de groupes de formation, dans un travail d'état des lieux de l'évaluation (évaluation des formés, évaluation de la formation) et de construction d'une approche (outils ouverts) susceptible d'induire une prise en compte au regard notamment des contenus de formation. Cela nous a conduits à identifier un certain nombre de points focaux en matière d'observation de la formation.

Un enjeu : construire un plan médian

L'articulation entre les facteurs externes à la formation (modifications sociétales, cadrage institutionnel, etc.) et les facteurs internes (les dispositifs de formation, les positionnements induits des acteurs, les cultures institutionnelles, etc.) constitue une réelle difficulté au regard de la prise en compte ou non des éléments évolutifs dans la formation. L'articulation entre les données de l'évaluation et celles de leur utilisation représente à la fois la dimension opératoire de l'évaluation mais également un réel obstacle. La production d'indicateurs est souvent détachée des conditions de leur réception par les acteurs. Cela nécessiterait que l'évaluation interne et la mise en œuvre des outils de celle-ci fasse l'objet d'une co-élaboration. L'élément effectivement implicatif se situerait de façon relativement décisive dans la forme de l'intervention choisie (figure 2) et moins dans l'objet même de l'évaluation. Dans le cadre des évaluations externes (conseil national de l'évaluation, inspection générale, OCDE, etc.), la réception des données de l'enquête risque fort d'être de peu d'incidence (Pacquay, 2005b). Les acteurs perçoivent une telle évaluation comme une

technique de pilotage parmi d'autres. Ils adoptent dès lors une posture du retrait et cela quelle que soit la validité informative des données. Ils en attribuent inéluctablement l'initiative à l'instance décisionnaire, l'hypothéquant d'intentions particulières. De ce dernier point de vue une réelle culture de la résistance s'est développée en raison notamment des multiples fluctuations intervenues dans le champ (Lessard & Meirieu, 2005).

Un inter-subjectivisme à déchiffrer

A défaut de toute évaluation et de tentative d'objectivation de la réalité se substitue une appréciation subjective de celle-ci à partir de modalités essentiellement implicites autorisant des stratégies pas toujours indemnes d'intentions spécieuses. Dans son ouvrage sur *L'évaluation en formation*, J.M. Barbier (1985) observe que dans les pratiques ordinaires des institutions interviennent trois formes d'évaluation : l'évaluation implicite, l'évaluation spontanée et l'évaluation instituée. Dans l'évaluation implicite, les effets sont d'une certaine manière « muets ». Tel enseignant stagiaire simule l'intérêt par sa seule présence. L'évaluation spontanée semble la moins étayée et la plus efficiente à la fois. Le jugement de valeur intervient à chaud. Les contenus sont moins décisifs que la relation entre le formateur et les formés (rôle de la valeur sociale). Les éléments de désirabilité (effet de halo) participent d'une telle évaluation (Beauvois, 1995). Sur le plan institutionnel elle conditionne un grand nombre de décisions en matière de gestion des ressources humaines. Dans l'évaluation instituée le jugement de valeur s'expliciterait à partir d'une formalisation de la démarche (bilan prévu, questionnaires, etc.) incitant les formateur à privilégier les éléments de la satisfaction plutôt que ceux de la transformation.

Une culture de la résistance

Un grand nombre de formateurs rencontrés expriment la crainte que l'évaluation de la formation glisse vers une évaluation des enseignements puis des enseignants. Une culture de l'évaluation toujours ancrée sur celle de l'enseignant (évaluation de l'élève et inspection de l'enseignant) et en quelque sorte alourdie par celle de la certification en seconde année rend difficile tout déplacement de ce point de vue. Le formateur, parce qu'il est un praticien de l'évaluation, n'en ignore aucun des travers dont celui d'une approximation dans laquelle la subjectivité à une part non négligeable. L'évaluation porte en elle le dévoilement des pratiques sans que celui-ci fasse l'objet d'un cadre éthique clairement identifié. La perception des enjeux par la plupart des acteurs concernés demeure souvent incertaine tout autant que les effets produits notamment dans le sens d'une amélioration de la formation. Les résultats aux concours constituent pour la première année le principal sinon l'unique indicateur.

Un certain nombre de points identifiés

A partir de l'analyse des différentes rencontres avec les acteurs (individuellement et en groupe) intervenues dans le cadre de l'action de l'observatoire, il se dégage un certain nombre de points de résistance récurrents. L'ensemble des acteurs, cela quelle que soit leur position dans l'institution, développe de multiples stratégies de contrôle ou d'évitement au regard de l'approche évaluative en tant que telle. Nous avons identifié deux attitudes dominantes. La première intéresse plus particulièrement les acteurs en charge du pilotage institutionnel. Elle tend à s'accorder avec le discours de l'évaluation des formations mais avec la visée sous-jacente d'une entrée dans les enseignements. La démarcation explicitée entre l'évaluation de la formation et celle des enseignements devient un enjeu intrinsèque de la démarche initialisée. Cela ne signifie nullement que la part de « l'évaluation spontanée » (Barbier, 1985) s'en trouvera réduite. La question des indicateurs, des standards et de leurs usages devient à ce stade prépondérante dans la mesure où ils devront, pour être efficients, se trouver partagés par l'ensemble des acteurs (Campanale, 2003). Cette condition suppose des formes de l'intervention implicatives (figure 2) quant à leur élaboration ou leur explicitation. La seconde attitude intéresse les formateurs qui compte tenu de la zone d'incertitude qui entoure les outils et les finalités de l'évaluation opèrent des renforcements des positionnements dans l'institution. Le dévoilement des pratiques de formation n'est pas moins complexe que celui des pratiques d'enseignement. A cela s'ajoute une défiance marquée au regard de la terminologie propre à l'évaluation (indicateurs, curriculum, compétence, qualité, performance, etc.) en raison des multiples acceptions que peuvent recouvrir les termes. Enfin toute approche évaluative est peu séparable de celle de jugement et de la valeur ou de la dévaleur sociale résultante.

Une dialectique à construire

L'évaluation participative : entre processus et résultat

Elle est porteuse de plusieurs enjeux dont celui décisif de l'adaptation. Elle repose sur l'implication des acteurs à partir, non d'une impulsion institutionnelle, mais d'une proposition de co-élaboration de la démarche envisagée. La participation des acteurs de la formation que sont les formateurs, les formés, interviendrait comme décisive dans un tel processus. Elle est susceptible de participer à l'élargissement de la professionnalité de l'enseignant et à celle du formateur « réflexif ». Une telle participation s'inscrit cependant dans une tension générée par la dimension certificative propre à la formation des enseignants stagiaires de seconde année mais également par l'exigence de plus en plus affirmée d'une nécessaire dimension formative (prise en compte des évolutions des enseignants stagiaires, personnalisation des parcours, etc.). Cette dernière implique de dissocier les deux dimensions mais également de reconsidérer les positionnements respectifs du formateur et de l'enseignant stagiaire. Les pratiques de références cristallisent sur ce plan des points de résistances d'autant plus légitimés que le cadre demeure certificatif. Cela peut conduire à s'interroger sur la spécification des outils destinés aux évaluations

certificatives (observations directes et indirectes, multiples pratiques du rapport) et à questionner le statut ambigu de l'évaluation formative en un tel contexte.

Le pilotage et les indicateurs

Il devient décisif de concevoir l'évaluation dans le dispositif (amont/aval) et non comme une intervention a posteriori (Landsheere, 1994). Co-construire les outils implique de concevoir la place des acteurs dans le dispositif de formation (conception systémique avec des relations ascendantes ou conception institutionnelle classique avec relations descendantes ?). Cela peut impliquer un travail en compréhension (sociologie, psychosociologie, etc.) portant sur la titularisation et les éventuels obstacles générés au regard de la professionnalisation. La construction et la stabilisation des indicateurs résultants constituent l'un des enjeux de la démarche de l'observation (Jumentier, 1996). Les indicateurs ont été jusqu'à présent amplement conçus dans le seul sens d'une aide au pilotage. Ils concernent pour l'essentiel les directions des institutions politiques et éducatives. De ce fait ce sont principalement les approches quantitativistes qui dominent en ce domaine en réduisant souvent la réception aux seuls décideurs (Vinokur, 2005).

Un outil ouvert : Le suivi de cohorte

Un tel outil, par sa mise en place même, semble de nature à modifier la relation entre l'institut de formation et les formés durant le temps même de la formation. La fidélisation des futurs enseignants implique ou révèlent une indispensable qualité dans la relation notamment avec les formateurs impliqués. Par ailleurs pérenniser la relation en vue de la stabilisation de l'effectif requiert également d'impliquer les formés dans l'approche évaluative. Une culture coopérative de l'observation est ainsi susceptible de se construire. Le suivi de cohorte constitue un exemple, parmi d'autres (questionnaire, étude de cas, entretiens, etc.) d'un outil susceptible d'être redéfini à partir de l'implication négociée des acteurs (Gérard, 2005). Les outils, au-delà de leur méthodologie propre, dépendent pour une grande par des formes de l'intervention préconisée.

Les formes de l'intervention

L'enjeu en termes de déroulement de la relation et d'exploration de la démarche évaluative dépend pour une grande part des formes de l'intervention mobilisées (évaluation, jugement, etc.). Nous avons tenté de synthétiser les possibles formes de l'intervention. Ces formes sont indicatives de la relation qui s'établit, de manière

générale, entre deux acteurs impliqués dans toute action *a fortiori* évaluative. Plus le degré de l'imposition se trouve élevé moins celui de l'implication paraît devoir l'être. Les formes de l'intervention sont ainsi susceptibles de se décliner de la co-construction vers celle de l'exclusion avec des effets inverses aux regards de l'exigence affichée.

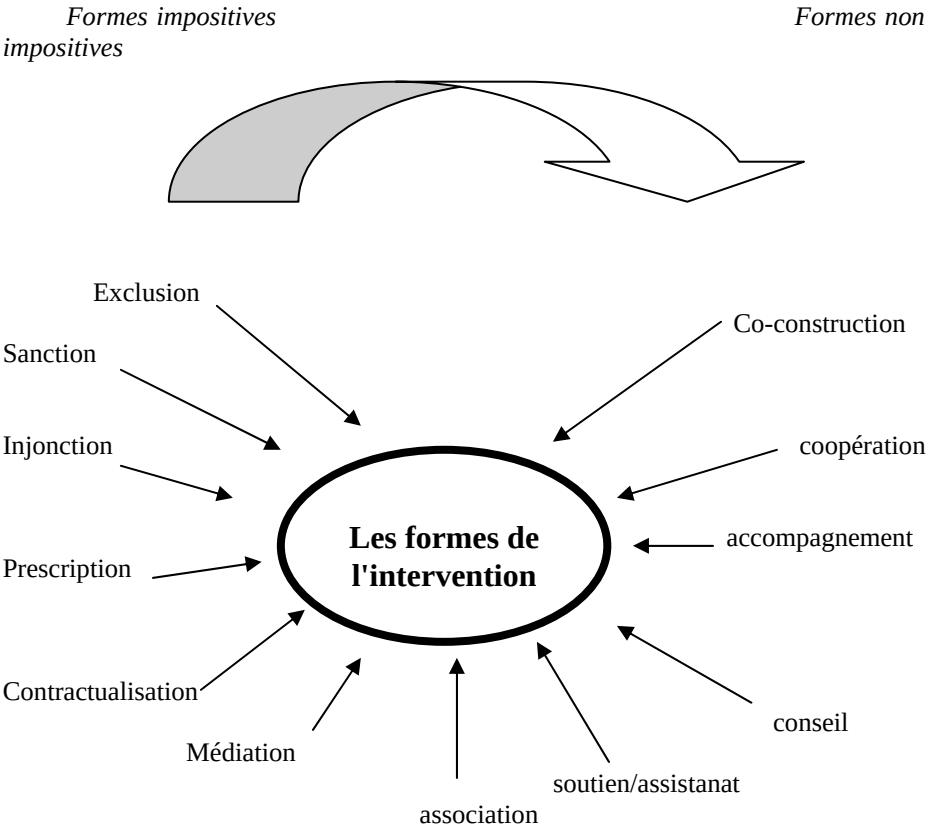


Figure 2 - L'arc de l'intervention

L'acteur institutionnel se trouve ainsi impliqué dans des formes de l'intervention définies par un cadre référentiel constituant en lui-même une formalisation effective de l'agir (Joas, 1999). Ce qui, par ailleurs, n'induit pas une inévitable conformité au rôle. Les acteurs, d'une manière générale, sont susceptibles d'entrer en braconnage (De Certeau, 1990), de construire une distance au regard du cadre d'intervention prescrit.

L'observatoire et les formes de l'intervention

co-construction

Une dialectique de l'observation qui s'articulerait principalement à partir de la proposition de formation et de sa réception à long terme par les formés. Il s'agit moins de relever les écarts que de repérer les dynamiques à l'œuvre au regard de la complémentarité de l'offre de formation (centres, acteurs, etc.) et des modalités d'appropriation par les formés. La co-construction induite d'outils de pilotage de la formation requiert la mise en œuvre d'une observation sur plusieurs années. Les contenus de formation sont ainsi susceptibles d'adaptation à partir de l'analyse du transfert opéré ou non.

Association

La question de la visibilité de l'observatoire semble se poser à plusieurs niveaux parmi lesquels :

- Celui de la mise à disposition de données, de résultats en direction de l'ensemble des acteurs de l'institut. Comment développer une culture de l'évaluation implicatrice des acteurs de la formation sans que soit envisagé d'opérer un retour co-construit en leur direction? L'effet de retour (Hadji, 1992) est susceptible d'influer sur les positionnements des acteurs tout autant qu'il étaye et confirme l'intérêt même de la démarche.
- Celui d'aide au pilotage de la formation auprès du conseil scientifique et pédagogique et du conseil d'administration. Il s'agit peut-être moins de s'assurer d'une conformité *stricto sensu* aux objectifs de formation du quadriennal que d'interroger le processus engagé. Cela peut passer par l'identification de lignes de résistance ou de transformation, le repérage de dispositifs innovants, l'effectivité de l'individualisation du parcours, etc. Cette possibilité renvoie aux condition de l'évaluation participative.
- Celui de l'autonomie de l'observatoire au regard des acteurs. L'identification par les acteurs de cette dernière participe pour une part de leur rapport à l'évaluation. Ce rapport pourrait être « travaillé » à partir de formes d'intervention implicatrices. (figure 2).
- Celui d'une éthique de l'évaluation co-construite. L'éthique constitue l'une des conditions de l'implication des acteurs. Son explicitation intervient comme un facteur déterminant de celle-ci.

Evaluation participative et éthique de la discussion

Une perspective à ouvrir

Les formes de l'intervention (figure 2), parmi celles les plus implicatives des acteurs, nécessitent que soient réalisées les conditions les plus favorables à l'approche socio-interactionnelle. De ce point de vue l'éthique de la discussion proposée par Habermas opère un renversement de la perspective Kantienne. La pertinence d'une règle d'action, de ce dernier point de vue, repose sur sa possible universalisation.

L'impératif Kantien implique que la règle découle du seul sujet : "Agis de telle sorte que tu puisses ériger la maxime de ton action en loi universelle". La loi ainsi devenue universelle se transforme alors en cadre fortement prescriptif impliquant une conformité désormais incontournable. L'approche d'Habermas propose que la validité d'une règle d'action découle d'une interaction/acceptation des personnes concernées par la règle d'action : « Toute norme valable doit satisfaire la condition selon laquelle : les conséquences et les effets secondaires peuvent être acceptés par toutes les personnes concernées » (Habermas, 1986, p. 86).

L'éthique de la discussion requiert en préalable l'accord alors que l'impératif Kantien l'envisage *a posteriori* par extension. Dans la première configuration, la discussion, le dialogue, l'interaction constituent le moyen et la fin. L'exigence éthique consiste à précisément ne pouvoir s'abstraire du dialogue comme définitoire de la finalité recherchée. Dans la seconde, la dimension sera invariablement monologale, en substance il s'agit d'imposer l'adhésion par la pertinence revendiquée de la règle d'action.

Une horizontalité décisive

L'éthique de la discussion selon Habermas implique une horizontalité de la démarche alors même que l'institution s'est toujours jusqu'à présent inscrit dans une verticalité revendiquée. Le principe de conformité aux normes dans l'évaluation des enseignants et, de façon plus concentrée encore, des enseignants stagiaires autorise peu de marge à l'acteur. Habermas (1983) propose de concevoir l'éthique non comme préétablie et relevant d'un principe transcendant mais comme résultante d'une communication à haute teneur réflexive. La définition proposée par J. Habermas de l'autoréflexion manifeste la globalité de l'agir communicationnel. Ce dont il s'agit tend à préserver le sujet de l'illusion consistant en une séparation du sujet et de l'objet alors même que le premier construirait en quelque sorte le second. C'est la perception par le sujet des formes dans lesquelles l'objet se manifestent qui établirait l'objet en question.

Le dialogue entre formés et formateurs ne porte pas en lui-même les vertus que souvent on lui attribue spontanément dès lors qu'il intervient dans le cadre d'une formation, elle-même sous-tendue par des enjeux de certification. La relation de pouvoir, notamment à travers la recherche d'une conformité aux rôles, n'est jamais très éloignée de la relation de formation elle-même intégrant des objectifs de plus en plus implicateurs de la personne (cadres d'analyse de la pratique centré sur le sujet). Il devient dès lors indispensable de garantir le dévoilement opéré. À défaut l'acteur peut rechercher à opérer des retours aux cadres et rôles antérieurs. De ce point de vue, le modèle du praticien réflexif semble s'imposer comme un modèle essentiellement projectif sinon prescriptif. L'évaluation participative serait susceptible de lui fournir ainsi un cadre d'exercice et de construction pertinent. A défaut, l'évaluation des formations risque de relever de formes plus prescriptives de l'intervention favorisant alors une désimplication des acteurs. Les adaptations s'opéreraient plus en raison des difficultés rencontrées, eu égard à la réalité du métier, qu'à partir d'une expertise développée dans l'institution de formation elle-même (Larivain, 2006).

Références

Ardoino, J. (1990). Les postures ou impostures respectives du chercheur, de l'expert et du consultant. *Les nouvelles formes de recherche en éducation*, 22-34.. Paris : Matrice.

Barbier, J.-M. (1985). *L'évaluation en formation*. Paris : PUF.

Beauvois, J.-L. (1995). La connaissance des utilités sociales. *Psychologie française*, 40, 375-387.

Berhens, M. (Ed.). 2006). Analyse de la littérature critique sur le développement, l'usage et l'implémentation de standards dans un système éducatif. Neufchâtel : IRDP.

Bernoux, Ph. (1995). *Sociologie des organisations*. Paris : Seuil.

Campanale, F. (2003). L'évaluation mutuelle interrogée par les principes d'efficacité et d'équité. *Actes du 16è colloque international de l'ADMEE-Europe* ? Liège, Belgique.

Bressoux, P.& Pansu, P. (2003). *Quand les enseignants jugent leurs élèves*. Paris : PUF.

Certeau M. (1990). *L'invention du quotidien*. Paris : Gallimard.

De Ketele, J.-M. (1993). L'évaluation conjuguée en paradigmes. *Revue française de pédagogie*, 103, 59-80.

Demailly, L. (2000). *Evaluer les politiques éducatives*. Bruxelles : De Boeck.

Gérard, F.-M. (2005). L'évaluation des compétences par des situations complexes. Actes du 18è colloque de l'ADMEE-Europe, Reims, France.

Habermas, J. (1986). *Morale et communication*. Paris : Éd. du Cerf.

Hadji, C. (1992). *L'évaluation des actions éducatives*. Paris : PUF.

Joas, H.(1999). *La créativité de l'agir*. Paris : Éd. du Cerf.

Jumentier, J. (Coordonné par) (1996). *Outils de pilotage dans les systèmes éducatifs.* CRDP de Lyon.

Landsheere, G. (de) (1994). *Le pilotage des systèmes de formation*. Bruxelles : De Boeck.

Larivain, C. (2006). *Les enseignants des écoles publiques et la formation. Interrogation de 1200 enseignants du premier degré.* Paris : Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche. Direction de l'évaluation, de la prospective et de la performance.

Leclerc C. (2004). Le formateur d'enseignants : entre directivité et réflexivité. In ASTOLFI J.-P., *Savoirs en action et acteurs de la formation (*pp 111-128). Rouen : Publications universitaires de Rouen.

Legendre Florence, 2005, *Evaluation de la formation IUFM de l'académie de Créteil par les professeurs des lycées et des collèges. Entre critiques et reconnaissances* ; rapport de recherche, iufm Créteil.

Lessard, C. & Meirieu P. (Ed.). (2005). *L'obligation de résultats en éducation. Evolutions, perspectives et enjeux internationaux.* Bruxelles : De Boeck.

Livre des références. Paris : Comité national d'évaluation.

L'assurance qualité de la formation des enseignants, 2006, Eurydice (ISBN 92-894-8628-7), (http://www.eurydice.org)

Mallet, J. (1994).*De l'organisation formatrice à l'organisation apprenante, en passant par les théories de la complexité.* Aix en Provence : Oméga formation conseil.

Morgan G.(1999). *Images de l'organisation.* Québec : Presses de l'Université Laval.

Pacquay, L. (2005b). Vers quelles évaluations du personnel enseignant pour dynamiser leur développement professionnel et leur implication vers des résultats ? *Recherche en formation*, 50, 55-74.

Vial, M. (2001). *Organiser la formation : le pari de l'auto-évaluation*. Paris : L'Harmattan.

Vinokur, A. (2005).Pouvoirs et mesure en éducation : avant propos. *Cahiers de la recherche sur l'éducation et les savoirs*, Hors série 1, 5-14.

REDEFINIR LES COMPETENCES ACQUISES PAR LES ELEVES A L'ECOLE PRIMAIRE POUR MIEUX COMPRENDRE LA STRUCTURE DES APPRENTISSAGES

Sophie Morlaix, IREDU/CNRS-Université de Bourgogne (Dijon-France)

Résumé :
Cette communication s'inspire d'une recherche récente (Morlaix, Suchaut, 2006) dont l'objectif principal est de comprendre comment les acquisitions des élèves se structurent notamment au cours du cycle III. En s'appuyant sur les évaluations mises en place par le ministère de l'Education nationale, ce travail engage la réflexion sur les compétences qui peuvent s'avérer déterminantes dans les acquisitions des élèves au CE2. Pour cela, nous reviendrons, dans un premier temps, sur la définition des compétences présentes dans les épreuves nationales, pour adopter par la suite une démarche beaucoup plus empirique, permettant d'isoler un certain nombre de variables essentielles à la réussite scolaire ultérieure.

Dans cette perspective, des données ont été collectées sur un échantillon d'environ 700 élèves à différents moments de la scolarité. Ces données concernent les résultats détaillés d'une même cohorte d'élèves aux évaluations nationales de 1999 pour le niveau du CE2 et de 2002 pour la classe de sixième ; les élèves ont également été testés en fin de cinquième en 2003 sur la base d'épreuves proches des tests nationaux. Des informations sur les caractéristiques socio-démographiques et scolaires des élèves ont également été recueillies.

I. Intérêts et limites des évaluations nationales

L'objectif de la mise en place des évaluations nationales est de fournir aux enseignants des outils standardisés de diagnostic des acquis des élèves. A différents niveaux de la scolarité, ces évaluations ont, en plus, le mérite d'évaluer un large ensemble de compétences. Par exemple, pour le niveau CE2, les évaluations répertorient 15 compétences en français et 27 en mathématiques. Chacune de ces compétences sont composées d'un nombre variable d'items, chaque item représentant une tâche à accomplir. Même si la dimension diagnostique reste prioritaire, ces évaluations peuvent constituer pour les chercheurs un outil précieux pour conduire, sous certaines conditions, une analyse fine des acquisitions des élèves. Les épreuves nationales ont en effet comme caractéristiques de balayer un ensemble de

compétences et de connaissances plus vaste que la plupart des tests mobilisés généralement dans les recherches françaises. Les exercices proposés sont conçus au plus près des programmes officiels et constituent à ce titre une référence commune pour l'ensemble des enseignants et des élèves.

Pourtant trois limites principales caractérisent les évaluations nationales : la première porte sur la mesure d'un niveau relatif de connaissances, la seconde sur les échelles de mesure utilisées pour valider l'acquisition d'une compétence, la dernière sur la définition des différentes compétences présentes dans les évaluations.

- **La mesure d'un niveau relatif de connaissances** : ces évaluations ne peuvent, en aucun cas, être considérées comme des normes à atteindre, mais seulement comme un diagnostic, en début de cycle, des réussites et des difficultés des élèves. Elles ne rendent que partiellement compte des compétences et des connaissances des élèves puisque la passation est collective et limitée à des exercices écrits. Il faut, de plus, être prudent dans l'interprétation des taux de réussite aux épreuves, la composition de celles-ci n'étant pas toujours les mêmes d'une année sur l'autre. Ainsi, les pourcentages de réussite peuvent varier parfois sensiblement suivant les années[1].

Ainsi, les chiffres doivent être « manipulés » avec précaution surtout si l'on veut effectuer des comparaisons d'un lieu à un autre et d'une année sur l'autre. Si, dans tous les cas, l'échelle de mesure retenue varie potentiellement de 0% à 100%, elle n'a toutefois pas de référence propre. Les concepteurs des épreuves s'attendent-ils à ce qu'un élève qui aurait assimilé la moitié du programme du cycle II obtienne 50% de réussite au test ? La réponse est bien sûr négative dans la mesure où un nombre infime d'élèves réussit la totalité des exercices présents dans les cahiers d'évaluation, ce qui ne veut évidemment pas dire qu'aucun élève français ne maîtrise totalement le programme du cycle II. On constate encore la difficulté à donner un sens immédiat aux chiffres en l'absence de norme absolue. En matière de pilotage local (école, circonscription, département, académie), les résultats des évaluations s'interpréteront

[1] Par exemple, en français, des écarts existent pour les scores globaux moyens mais aussi pour les différents champs évalués. La moyenne des scores globaux oscille entre 60,5% (en 2001) et 73,5% (en 2004). On voit bien qu'il est difficile d'accorder à tous ces chiffres une valeur unique car ils dépendent fortement de la nature des évaluations et les variations constatées ne renvoient pas forcément à des baisses ou à des hausses du niveau général des élèves.

principalement de façon relative en termes d'écarts avec les chiffres nationaux (ou académique) et d'évolutions de ces écarts au fil des années[2].

- **Les échelles de mesure utilisées** *:* dans les évaluations nationales, la réussite des élèves, est chiffrée de façon variable d'une compétence à l'autre selon le nombre d'items mobilisés pour évaluer les élèves. Pour prendre des exemples extrêmes, la compétence «*comprendre un texte…*» est évaluée à travers 12 items alors que la compétence «*construire un figure simple sur un quadrillage…*» ne concerne qu'un seul item. Selon le cas, la graduation de la réussite est donc très variable d'une compétence à l'autre allant de la simple dichotomie « échec / réussite » à une échelle en 12 scores. Les compétences sont donc mesurées de façon plus ou moins riche selon le nombre d'items utilisés dans les différents exercices de l'épreuve.

La seconde difficulté résultant des échelles de mesure insatisfaisantes concerne le seuil de réussite : à partir de quel score peut-on considérer que l'élève maîtrise la compétence ? Plus l'échelle de mesure comporte d'échelons, plus il est difficile d'apporter une réponse à cette question. Il serait bien sûr possible de fixer un seuil de réussite commun à toutes les compétences en adoptant les principes en usage qui reposent souvent sur l'arbitraire (75% est généralement le seuil qui est retenu), mais là encore, les dissymétries entre les différentes échelles rendent cette solution trop imparfaite, notamment pour les compétences qui comportent très peu d'items[3].

- **L'identification des compétences** : un troisième point de discussion se rapporte à la pertinence de l'identification de certaines compétences. Certains exercices, supposés évaluer une même compétence, peuvent comporter des items que les élèves réussissent différemment (indépendamment du fait que les items soient d'une difficulté variée pour une même compétence). Certains items mesurant la même compétence ne présentent aucune corrélation (nous allons y revenir). Il est donc

---

[2] Les comparaisons peuvent aussi être rendues délicates si la population scolaire varie au fil des ans. Pour avoir une mesure plus juste, il conviendrait de prendre en compte le caractéristiques sociales et scolaires des élèves pour calculer une mesure relative des performances, comme c'est le cas pour les IPES.

[3] Une autre solution consiste à avoir recours à des juges ou à des experts (des professionnels de terrain compétents dans le domaine) qui définiraient eux-mêmes et individuellement les critères d'acceptation de la performance. Plusieurs techniques statistiques peuvent ensuite être mobilisées pour exploiter les différents jugements et définir un score de réussite correspond à chaque compétence (Laveault, Grégoire, 1997). Cette méthode, même si elle présente des intérêts évidents, de par sa lourdeur de mise en œuvre, n'est pas envisageable dans notre étude (il serait nécessaire de disposer d'un nombre important de « juges », chacun devant s'exprimer sur toutes les compétences de l'évaluation nationale).

possible qu'il n'y ait pas dans les évaluations nationales une adéquation parfaite entre la définition des compétences et leur mesure à travers les items proposés. Autrement dit, il n'est pas certain que chaque item mesure parfaitement la compétence visée.

Eu égard à ces différents problèmes méthodologiques que pose l'approche par compétence, nous avons choisi d'adopter une perspective plus fine des apprentissages des élèves en se centrant sur les items. Même si cette approche comporte elle aussi des limites, notamment quand au degré de difficulté des items (Demeuse, Henry, 2004), elle présente certains avantages.

II. Redéfinir les compétences à partir des items présents dans les évaluations nationales

L'approche au niveau des items présente plusieurs avantages. Elle règle tout d'abord le problème de l'échelle de mesure puisque tous les items présentent le même barème de cotation : 0 pour une réponse erronée, 1 pour la réponse attendue ; on dispose donc d'une échelle commune de réussite pour toutes les compétences. De cela, découle un second avantage : la question du seuil de réussite ne se pose plus, puisqu'il est défini de façon objective par la réussite ou l'échec, sans possibilité de situations intermédiaires[4].

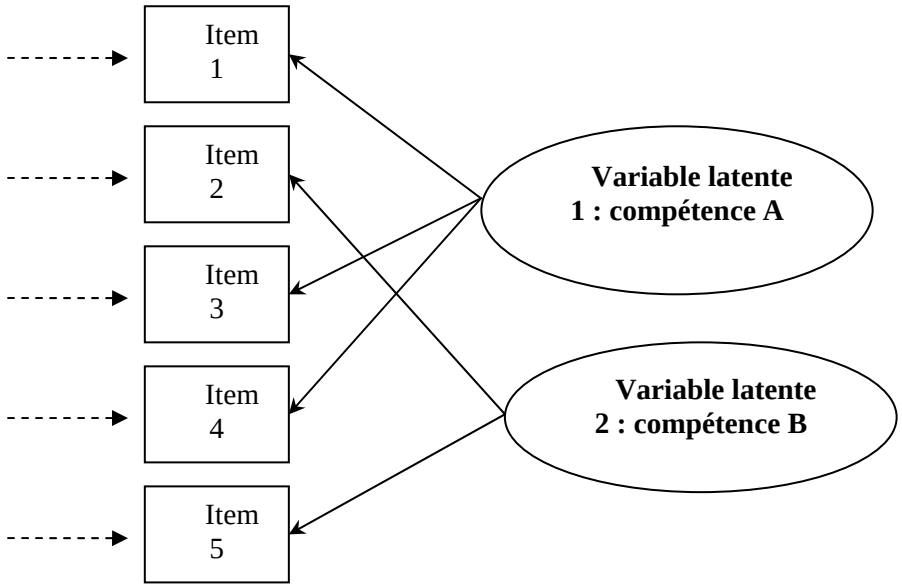
D'autres avantages à cette approche existent. Ce niveau d'analyse permet de tester la pertinence de la définition des compétences dans les évaluations nationales en examinant de près les corrélations entre les différents items censés mesurer une même compétence. On s'aperçoit en effet que certaines compétences dans les évaluations nationales sont composées d'items très peu corrélés ente eux. L'approche par les items suppose ainsi qu'il est possible de mettre en évidence des liaisons fines entre items et d'identifier ainsi de « nouvelles compétences » ou de recomposer les compétences initiales en effectuant de nouvelles associations d'items.

Pour ces raisons, nous avons travaillé à partir des 171 items composant les évaluations nationales de français et mathématiques de CE2. Le point de départ de l'analyse est la matrice de corrélations établie entre tous ces items. Compte tenu du

[4] On mentionnera que cette dichotomie (échec / réussite) n'est pas adaptée à une interprétation diagnostique des évaluations nationales et que les barèmes de cotation de certains items comportent à l'origine d'autres paliers qui permettent une analyse des erreurs des élèves. A notre niveau, seule la réponse effectivement attendue à l'item nous intéresse et par conséquent les « autres réponses » fournies par les élèves n'ont pas été considérées comme des «bonnes réponses ».

nombre important de corrélations existantes, nous avons décidé de ne retenir que celles supérieures à 0.20. Sans surprise, on remarque qu'une bonne partie des corrélations (57% d'entre elles) se rapporte à des items appartenant à un même exercice et 86% à la même discipline, cela laisse donc aussi la place à des corrélations faisant intervenir des items d'origines diverses. Après avoir effectué des regroupements entre items les plus corrélés, nous avons testé statistiquement la pertinence des liaisons entre les items d'un même groupe.

L'objectif a donc été d'analyser les liaisons statistiques à l'intérieur de chaque groupe d'items de façon à mettre à jour des variables latentes pouvant être interprétées comme des compétences, des aptitudes ou des capacités mobilisées par les élèves dans les évaluations nationales. Les modèles ont été estimés avec le logiciel LISREL, qui permet notamment l'estimation de modèles de mesure. Ceux-ci établissent les relations entre les variables latentes (compétences supposées) et leurs indicateurs (les items). Le logiciel LISREL permet de valider la mesure de la compétence par les indicateurs retenus et d'estimer les liens supposés entre les indicateurs et la (ou les) variable(s) latente(s).



Graphique 1 : modèle de mesure entre deux variables latentes et leurs indicateurs

Les flèches en traits pleins matérialisent l'intensité des relations qui lient chacun des indicateurs (items) à la variable latente (compétence), un coefficient de régression (et son degré de significativité) pour chaque indicateur fournit une indication sur la validité du modèle. LISREL donne également une mesure de fiabilité du modèle grâce à l'indication de l'erreur de mesure sur chacun des indicateurs (flèches en pointillés sur la figure). Il est donc possible de savoir si les indicateurs retenus sont « robustes » pour mesurer la variable latente. LISREL fournit également un certain nombre de statistiques générales rendant compte de la qualité du modèle de mesure[5] (Morlaix, 2002)

Des modèles de mesure ont été estimés sur l'ensemble des regroupements d'items. 63 variables latentes ont ainsi pu être identifiées (Morlaix, Suchaut, 2006). Par défaut ces variables latentes ont été identifiées comme des compétences[6]. L'étape suivante du travail a consisté à examiner les relations statistiques entre ces nouvelles variables qui rendent compte des compétences des élèves au début du cycle III.

III. Structure relationnelle et hiérarchique des compétences à l'entrée au cycle III

L'analyse en variables latentes conduit à une recomposition des compétences des élèves à l'entrée au CE2. Parmi les 63 variables mises à jour, 27 d'entre elles (soit 43%) correspondent, souvent de façon partielle, à des regroupements d'items déjà présents dans les évaluations nationales. La correspondance entre les compétences des épreuves et les variables latentes est néanmoins très imparfaite puisque seules 5 variables correspondent exactement à des compétences figurant dans les évaluations nationales (il s'agit uniquement d'exercices de mathématiques). Cette nouvelle configuration des compétences des élèves soulève d'autres questions. En premier lieu, certaines variables latentes semblent à première vue rendre compte de compétences composites issues d'exercices variés. L'objectif est alors de pouvoir donner un sens précis à ces compétences mesurées par les variables latentes. En second lieu, il est

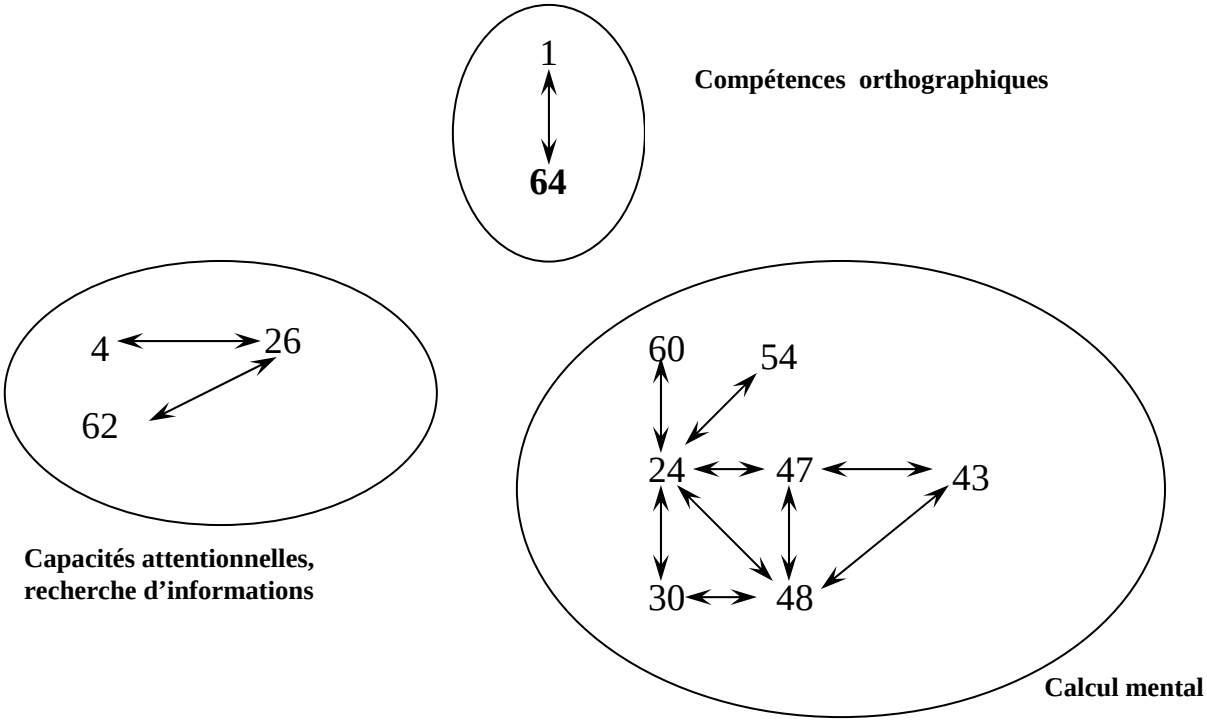
[5] Le χ^2 (khi carré) mesure la distance séparant les deux matrices de corrélations, celle théorique et celle portant sur les corrélations estimées. Cet indicateur du χ^2 doit être le plus petit possible et il doit être aussi proche que possible du nombre de degrés de liberté du modèle. La π value (P value) mesure la probabilité d'obtenir la valeur du χ^2 correcte. Le modèle est considéré comme recevable si cette valeur est supérieure à 0,5 et on cherchera en outre à ce qu'elle soit la plus élevée possible. Enfin, l'indice RMSEA (Root Mean Square Error of Approximation) se rapporte à la moyenne des résidus[5] du modèle et à leur significativité. Le modèle est acceptable si la valeur de cet indicateur est inférieure ou égale à 0,05.

[6] Nous l'avons rappelé précédemment, et notre collaboration avec des psychologues nous l'a montré : elles peuvent dans certains cas plutôt correspondre à des habiletés ou aptitudes.

indispensable d'étudier les relations statistiques entre les variables latentes pour mettre à jour la structure des acquisitions des élèves ; il se pose notamment la double question de la hiérarchisation des compétences et de leur pouvoir de prédiction de la réussite scolaire globale à ce niveau de la scolarité.

Pour répondre à ces questions, nous avons analysé la matrice de corrélations entre les différentes variables latentes. Les coefficients de corrélation présentent des valeurs allant de 0 à +0,80 (la moyenne sur l'ensemble des corrélations étant de +0,24). Nous avons sélectionné dans un premier temps les corrélations les plus importantes (supérieures à +0,70). Cela permet de dégager trois grands groupes de compétences qui sont représentées sur le graphique 2 (les chiffres correspondent aux numéros associés aux variables latentes).

Un premier ensemble est composé des variables « *comp1* » et « *comp64* ». Il s'agit ici clairement de compétences orthographiques qui sont mesurées avec ces deux variables. En effet, la presque totalité des items rendant compte de ces variables latentes concernent deux exercices de dictée. Le deuxième ensemble regroupe les variables « *comp4* », « *comp26* » et « *comp62* ». Cet ensemble d'apparence disparate prend sens quand on l'examine sous l'angle de la psychologie cognitive. Autant les items de français que ceux de mathématiques présents dans ce regroupement évaluent la capacité que peuvent avoir les élèves à rechercher de l'information plus ou moins complexe à partir de supports divers (textes, mots, calendriers, emploi du temps, plans, énoncés de problème). Ce sont donc les capacités attentionnelles des élèves qui sont mises à contribution pour la maîtrise de cette compétence globale. Le troisième ensemble établi sur les corrélations les plus fortes, regroupe 7 variables latentes. Le lien commun entre ces variables est également clair puisque les items de calcul mental interviennent systématiquement pour chacune d'entre elles.



Graphique 2 : Blocs de compétences des évaluations CE2

Les acquisitions des élèves à l'entrée au CE2 s'organisent principalement autour de ces trois blocs de compétences qui ne sont pas de même nature. Si l'acquisition des compétences orthographiques dépend principalement d'un enseignement systématique, les deux autres blocs de compétences sont davantage associées à des processus plus complexes qui interviennent de façon transversale dans de nombreuses situations d'apprentissage. On devrait s'attendre à ce que ces blocs majeurs contribuent fortement à l'explication des différences de réussite entre élèves à l'entrée au CE2. Pour vérifier cela, une régression « pas à pas » a été estimée avec comme variable dépendante le score global de CE2 et comme variables explicatives l'ensemble des compétences mises à jour précédemment. Le tableau 1 présente les résultats en ne conservant que les 12 compétences les plus prédictives. Celles-ci expliquent à elles seules plus de 92% de la variance du score global, soit presque sa totalité.

On constate que les trois compétences les plus prédictives (« *comp48* », « *comp4* », « *comp64* ») appartiennent chacune à un des blocs identifiés auparavant. Ces trois compétences expliquent à elles seules 82% de la variance du score global. Cette analyse de la prédictivité montre que certaines compétences sont au cœur des acquisitions des élèves à l'entrée au CE2. Les habiletés en calcul mental, la capacité à retrouver rapidement des informations dans des supports variés, la maîtrise de l'orthographe structurent ainsi fortement les résultats des élèves au début du cycle III.

Tableau1 : Modèle analysant la prédictivité des variables latentes sur le score global de CE2 (régression pas à pas)

Ordre d'introduction	Variables	Coefficients standardisés	t de Student
1	Comp48	+0,23	12,3 ***
2	Comp4	+0,24	14,7 ***
3	Comp64	+0,17	12,4 ***
4	Comp58	+0,13	10,2 ***
5	Comp13	+0,12	10,0 ***
6	Comp17	+0,11	8,9 ***
7	Comp60	+0,13	7,2 ***
8	Comp41	+0,08	6,6 ***
9	Comp16	+0,08	6,8 ***
10	Comp34	+0,08	7,2 ***
11	Comp15	+0,08	6,8 ***
12	Comp3	+0,08	6,5 ***
	Constante	68,76	510,02
	R^2	0,92	

*** : significatif au seuil de 1%

Nous allons examiner à présent comment ces compétences les plus prédictives des évaluations nationales se hiérarchisent[7]. Le tableau 2 expose les relations entre la réussite et l'échec aux 12 compétences les plus prédictives du score global de CE2. Les compétences sont présentées de façon hiérarchisée dans le tableau (de « *comp15* » à « *comp3* »). La réussite à la première compétence « *comp15* » est donc censée être une condition à la réussite de la seconde (« *comp58* »), qui elle-même détermine la réussite à la suivante (« *comp41* »)…, jusqu'à la dernière compétence (« *comp3* ») qui ne peut être acquise que si toutes les autres compétences sont maîtrisées. Dans chaque case, le nombre indiqué représente le pourcentage d'élèves (sur la totalité de l'échantillon) qui réussit une compétence tout en ayant échoué à une autre. Ainsi 4% des élèves réussissent « *comp58* » tout en échouant à « *comp15* » ; de même, 3,3% des élèves réussissent « *comp41* » et échouent à « *comp15* ». Si la structure hiérarchique des compétences se vérifie, on devrait donc s'attendre à ce que les chiffres du tableau présentent des valeurs faibles. Cela est effectivement le cas même si certaines valeurs s'écartent de la règle générale. Ainsi, si la maîtrise de « *comp15* » semble bien être une condition indispensable à la réussite à toutes les autres compétences (la première ligne du tableau présente des valeurs très faibles), la relation entre « *comp13* » et « *comp16* » ne semble pas vérifiée dans la mesure ou plus de 20% des élèves réussissent « *comp13* » tout en échouant à « *comp16* ». Globalement la structure de dépendance entre compétences se vérifie dans la majorité des cas ; les cases grisées dans le tableau représentent ces situations (pourcentages d'élèves inférieurs ou voisins de 10%).

[7] Une étape préalable à l'analyse est de rendre compte systématiquement de la maîtrise ou de l'échec à une compétence, ce qui suppose une transformation de l'échelle de mesure, on passe alors d'une échelle d'intervalle (qui varie selon la compétence considérée) à une échelle nobminale, représentée de façon dichotomique : échec ou réussite. Il se pose alors la question du seuil à partir duquel on va considérer que l'élève a acquis la compétence visée; il n'existe pas de réponse parfaite mais nous avons considéré qu'un score supérieur à 75% de réussite soit associé à la maîtrise de la compétence.

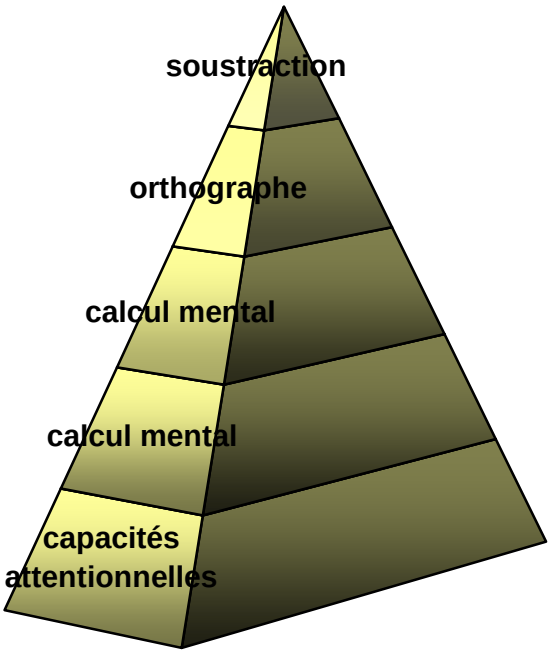
Tableau 2 : Relations entre l'échec et la réussite aux compétences de CE2
(% d'élèves)

	REUSSITE											
ECHEC	comp15	comp58	comp41	comp17	comp16	comp13	comp34	comp4	comp48	comp60	Comp64	comp3
comp15		**4,0**	3,3	2,1	3,0	2,4	2,8	0,0	0,4	0,1	0,1	0,6
comp58			**8,5**	6,7	7,9	7,6	8,0	1,2	1,3	0,7	0,7	1,2
comp41				**14,6**	14,6	14,0	13,7	2,8	3,1	1,6	1,3	1,6
comp17					**19,5**	17,7	16,1	4,5	4,9	3,6	2,5	3,1
comp16						**20,6**	16,7	6,7	6,7	4,3	4,8	3,1
comp13							**16,5**	4,5	6,1	4,2	2,2	3,3
comp34								**6,9**	7,2	4,6	5,0	4,9
comp4									**11,2**	7,9	6,7	8,9
comp48										**5,8**	7,5	6,1
comp60											**8,6**	7,5
comp64												**9,8**

Les compétences semblent donc assez bien hiérarchisées dans l'ensemble ; la compétence qui se trouve au sommet de cette hiérarchie correspond à la maîtrise de la soustraction (variable « *comp3* »). Ceci signifie que l'acquisition de la technique opératoire de la soustraction est un processus qui nécessite de la part des élèves diverses capacités et habiletés préalables. En effet, très peu d'élèves réussissent dans cette dimension des acquisitions sans maîtriser toutes les autres compétences qui figurant dans le tableau 2. La seconde compétence dont la maîtrise dépend de nombreuses autres est de nature orthographique (« comp64 ») ; viennent ensuite deux compétences dans lesquelles le calcul mental intervient majoritairement, soit directement (« comp48 »), soit indirectement par le biais d'autres compétences (« comp60 »). La compétence qui se situe au plus bas niveau de cette structure hiérarchique correspond à des items évaluant la connaissance des unités de mesure. En fait, c'est davantage la compréhension de consignes simples qui semble être la compétence véritablement visée dans cet exercice. Le tableau en annexe, fournit la liste l'ensemble des compétences du tableau 3 en indiquant les items qui s'y réfèrent. Le graphique 2 symbolise quant à lui la structure pyramidale des 5 premières compétences du tableau 2.

Cette structure pyramidale des compétences peut fournir des indications didactiques et pédagogiques pour l'enseignement au cycle II, en amont du CE2. Ces indications concernent principalement les contenus d'enseignement et leur

programmation dans le temps, ils s'appuient sur le double constat suivant établi à l'entrée du CE2 : certaines compétences sont difficilement accessibles à l'ensemble des élèves, certaines compétences sont essentielles à l'acquisition d'autres compétences



Graphique 2 : Structure pyramidale des compétences des élèves au CE2

Suite à ces deux constats, on peut déjà suggérer que la définition des programmes scolaires mette l'accent de façon explicite sur les compétences définies comme essentielles précédemment, cela nécessite aussi que des situations d'enseignement concrètes puissent être envisagées pour l'acquisition de ces compétences, ce qui n'est pas immédiat quand elles sont transversales. Il est ensuite utile de s'interroger sur la planification de l'enseignement de ces compétences en prenant en compte la structure hiérarchique d'apprentissage telle qu'elle apparaît sur le graphique précédent, c'est-à-dire en respectant le fait que certains apprentissages ne peuvent se réaliser que si

certains autres sont déjà maîtrisés. Le temps consacré à l'acquisition des différentes compétences est aussi une question centrale, sachant que le volume de temps alloué est très lié aux progrès des élèves et celui-ci s'inscrit dans des situations d'arbitrage : plus de temps consacré à une activité réduit le temps attribué à une autre (Suchaut, 1996, Morlaix, 2000). En outre, certains élèves ont besoin de plus de temps que d'autres pour acquérir certaines notions. Les recommandations pédagogiques doivent donc aussi tenir compte de la diversité des élèves et de leur rythme d'apprentissage spécifique.

Conclusion

L'objectif principal de cet article était d'étudier la structure des acquisitions des élèves à l'entrée en CE2 et d'identifier, par une approche empirique, les compétences principales. Un premier constat qui se dégage des résultats est l'aspect transversal des acquisitions scolaires. La vision institutionnelle des évaluations nationales masque cet aspect par le regroupement, parfois artificiel, de certains items au sein d'un même exercice alors que l'approche statistique permet d'identifier des compétences de nature variée. De façon complémentaire, des items provenant d'exercices variés, voire de disciplines différentes peuvent ont été identifiés comme mesurant une seule et même compétence. Ce premier constat peut avoir des implications évidentes quant à l'utilisation des évaluations nationales sur le plan pédagogique dans une perspective diagnostique. Cela permet en effet de mieux cerner les processus d'apprentissage des élèves qui sont mobilisés lors de la réalisation d'un exercice et donc de fournir des éléments pour organiser des situations pédagogiques ciblées sur l'acquisition de certaines compétences.

Le deuxième constat a trait au statut particulier de certaines compétences qui semblent structurer véritablement les apprentissages des élèves et qui concernent trois domaines : l'orthographe, l'attention et la recherche d'information, le calcul mental. Ces deux derniers éléments se rattachent d'ailleurs plus à des capacités qu'à des compétences puisqu'elles interviennent dans des situations d'apprentissage très variées qui dépassent largement l'approche disciplinaire. Ceci présente également un intérêt sur le plan pédagogique dans le sens où des activités systématiques peuvent contribuer à développer les mécanismes d'apprentissage intervenant dans la maîtrise de ces compétences essentielles.

Un troisième constat porte sur la structure hiérarchique des acquisitions des élèves. Il apparaît que les acquisitions qui rendent le mieux compte du niveau global des

élèves à l'entrée au cycle III, correspondent à des mécanismes de base qui détermineraient la maîtrise de compétences plus spécifiques en matière d'orthographe ou de technique opératoire.

Bibliographie

Demeuse M. Henry G (2004), « La théorie classique des tests », in Demeuse M. (2004), *Introduction aux théories et aux méthodes de la mesure en sciences psychologiques et en sciences de l'éducation*. Les Editions de l'université de Liège.

Laveault D., Grégoire J. (1997), *Introduction aux théories des tests en sciences humaines*. De Boeck Université.336 pages.

Morlaix S. (2000). Rechercher une meilleure répartition du temps scolaire en primaire pour favoriser la réussite au collège. *Revue française de pédagogie.* N°130, 121-131

Morlaix S. (2002). Intérêts et apports de l'analyse des variables latentes pour le chercheur en sciences sociales : exemple d'application à l'économie de l'éducation. *Orientation scolaire et professionnelle*, vol 31., N°1, 117-138.

Morlaix S., Suchaut B. (2006). Evolution et structure des compétences des élèves à l'école élémentaire et au collège. Une analyse empirique des évaluations nationales. Rapport pour l'UNSA-Education. 196 p. et annexes.

Suchaut B. (1996), « La gestion du temps à l'école maternelle et primaire : diversité des pratiques et effets sur les acquisitions des élèves ». *L'année de la recherche en sciences de l'éducation*, 123-153.

ÉVALUATION ET AUTOÉVALUATION DES COMPÉTENCES COMME MOYEN DE QUALIFICATION DE L'INSTRUCTION A DISTANCE.

Gaetano Domenici, Università Roma III

Résumé :

Dans le cadre des études sur les problèmes de l'évaluation, tant dans la formation en présence que dans la formation à distance, plusieurs auteurs (Domenici 1988,1990, 2005, Calonghi 1992, Trentin 2001, Ghislandi 2002, Rivoltella 2003, Scheerens 2003, Khan 2004, La Marca 2004, Varisco 2004, Galliani 2005, Trinchero 2006) conviennent que les pratiques d'évaluation, surtout celles qui concernent l'*assessment,* contribuent de façon significative à déterminer les conditions de connaissance sur la base desquelles tous les acteurs du processus de formation (étudiants, enseignants, proviseurs) prennent des décisions.

Dans les cours *e-learning,* toutefois, on enregistre souvent une distance entre, d'un côté, finalités et modèles d'enseignement déclarés par le projet de formation - irréprochables sur le plan théorique – et, d'autre côté, procédures, instruments et utilisation des données d'évaluation. A un extrême on remarque que : a) la seule variable d'évaluation considérée est l'apprentissage acquis, exprimé à peu près seulement en termes de connaissances, b) les instruments utilisés sont presque exclusivement des tests objectifs ne relevant que quelques aspects du plus complexe processus d'apprentissage. Dans d'autres cas, au contraire, on souligne l'importance que les dimensions relationnelles, coopératives et de construction sociale des savoirs assument dans l'apprentissage individuel et collectif (Khan 2004), en laissant fréquemment aux marges le contrôle d'importantes dimensions cognitives (exprimées et relevables en termes de compétences stratégiques de secteur) et celui des compétences affectives-motivationnelles de l'apprentissage. Il en dérive, dans les deux cas, une sous-estimation de la complexité systémique et du poids que

l'évaluation revêt, surtout dans la régulation et dans l'optimisation des processus de formation, en présence comme à distance.

Un plus systématique relevé des facteurs qui qualifient la communication pédagogique, déjà important quand on fait un processus d'hétéro-évaluation, devient décisif dans l'activité d'auto-évaluation des étudiants, surtout dans un cadre opérationnel de matrice constructiviste qui normalement caractérise la formation *on line*, où les processus d'auto-évaluation résultent irremplaçables pour promouvoir dans chacun la conscience des propres stratégies cognitives, relationnelles et affectivo-motivationnelles, afin de pouvoir les améliorer progressivement (Horton 2001, La Marca 2004).

Les épreuves traditionnelles et celles objectives représentent les deux pôles de l'ample gamme d'instruments pour vérifier l'apprentissage. Mais s'il est vrai que chacune d'entre elles possède des qualités, les deux présentent néanmoins des limites métrologiques qui en déconseillent une utilisation diffusée ou exclusive. En effet, le cadre informatif que l'on dégage de leur utilisation est pauvre; la pertinence et l'efficacité des décisions que l'on peut assumer sur cette base pour qualifier la didactique résultent basses. Les épreuves demi-structurées possèdent les qualités des épreuves traditionnelles et de celles objectives sans pour autant présenter leurs défauts. Elles sont nées des tentatives de réaliser la pleine valorisation des aspects structuraux plus profitables de l'expérience des enseignants en matière d'évaluation et de surmonter la rigidité et le schématisme des tests objectifs. Non seulement elles permettent de rendre toujours cohérentes les caractéristiques de l'instrument d'évaluation par rapport aux connaissances à relever, mais, en permettant la simulation de contextes inusuels, elles favorisent l'émergence des stratégies d'utilisation des savoirs acquis et des procédures d'évaluation et de décision adoptées au fur et à mesure. Par la promotion de ces capacités métacognitives et de méta-évaluation, les épreuves demi-structurées font assumer au moment de la vérification aussi cette valeur de formation habituellement sous-estimée dans la pratique didactique.

L'expérience de recherche, d'abord observative, ensuite empirique-expérimentale, a été conduite dans le cadre du projet pilote de formation en service des enseignants d'école maternelle et primaire de la région du Lazio (I), finalisé à l'obtention de la licence en sciences de l'éducation , en modalité *blended learnin*g, en 2004/2005. Une première analyse de la variance relative aux résultats des examens des 1048 étudiants universitaires inscrits et des corrélations entre certaines variables d'entrée et de procès avec leur carrière universitaire (nombre et notation des examens soutenus) a montré une connexion étroite entre les dimensions quali-quantitatives des résultats d'apprentissage et les modalités de participation aux activités didactiques et d'évaluation proposées.

En particulier, on a constaté que les meilleurs résultats correspondaient à une positive participation au processus et à l'intégration des exercices d'approfondissement, extension et consolidation des connaissances promues par chaque cours grâce à l'utilisation d'épreuves demi-structurées pour l'autoévaluation.

Dans la phase de travail actuelle, suite à la détermination a posteriori (sur la base des données recueillies après la première année de formation à distance) de deux échantillons statistiques et à la définition d'un groupe de contrôle et du groupe expérimental de la recherche, l'emploi des épreuves demi-structurées comme instrument d'autoévaluation (au lieu des test objectifs) a été considéré comme variable indépendante. Avant l'été, on relevera dans les deux groupes les compétences acquises et le niveau des attitudes et des motivations face à de nouvelles tâches d'apprentissage hétéro et autodirect, outre que des statistiques relatives aux données de tracement de la plateforme *e-learning*, concernant notamment la durée et la typologie des activités d'autoévaluation.

La recherche a permis de repérer quelques aspects cruciaux - qui feront objet de discussion à partir des résultats de toutes les phases - qui devraient faciliter la mise au point progressive d'un modèle systémique d'évaluation, surtout mais pas seulement pour la formation *on line* :

1. faire pratiquer aux étudiants l'autoévaluation pour les rendre conscients des caractéristiques spécifiques de leurs propres stratégies cognitives relationnelles et affectivo-motivationnelles et si nécessaire de les modifier
2. faire utiliser les résultats de la comparaison entre l'hétéro et l'autoévaluation de processus comme moyen de connaissance pour la structuration des décisions de la part des étudiants et de la part des formateurs
3. intégrer une pluralité d'instruments qui permettent le relevé et l'évaluation tant des apprentissages que de la qualité de la participation et de l'interaction dans le cadre des activités de communication de laboratoire de *coooperative learning*
4. utiliser, surtout dans les processus d'autoévaluation, les épreuves demi-structurées , en particulier la simulation de contextes dans lesquels on demande, en présence de contraintes formelles opportunément définies, l'activation de processus de métacognition évaluation et méta-évaluation, décision et méta-décison pour la solution d'un problème donné c'est-à-dire la manifestation des compétences stratégiques de secteur, acquises (Domenici-Cajola 2005).

L'ÉVALUATION DES APPRENTISSAGES DANS UN ENVIRONNEMENT SCOLAIRE MARQUÉ PAR LA DIFFÉRENCIATION PÉDAGOGIQUE ET L'ENSEIGNEMENT EN FAVEUR DU TRANSFERT : QUELQUES RÉSULTATS D'UNE RECHERCHE-ACTION MENÉE AVEC DES ENSEIGNANTS DU PRIMAIRE.

Annie Presseau, Stéphane Martineau, Université du Québec à Trois-Rivières

Résumé:

Contexte, questions de recherche et cadre de référence

La démocratisation de l'école, puis plus récemment de la réussite, n'est pas sans conséquence en terme d'attentes à l'égard des enseignants. Elle implique la prise en compte d'une certaine diversité au sein de la population scolaire (Zakhartchouk, 2001). C'est ainsi que la différenciation pédagogique apparaît comme une réponse à cette diversité et réaliser le mandat de l'école québécoise : instruire, socialiser et qualifier (MEQ, 2001). En outre, avec la réforme des curriculums, il est dorénavant exigé des enseignants qu'ils forment tous les élèves à des compétences qu'ils pourront transférer dans divers contextes. Qui dit transformation de sa pratique au regard de la différenciation pédagogique et dans une optique de transfert des apprentissages, dit aussi nécessité de revisiter ses pratiques d'évaluation des apprentissages des élèves…

C'est en tenant compte de ces divers éléments de contexte qu'a été mise en place la recherche-action à l'origine des résultats présentés dans le cadre de cette communication. Plus précisément, nous avons voulu répondre aux questions suivantes : À quelles pratiques les enseignants recourent-ils pour soutenir l'apprentissage et le transfert chez leurs élèves aux différents temps de recherche-action et comment les apprentissages des élèves sont-ils évalués? Quels dispositifs les enseignants mettent-ils en place pour favoriser la différenciation pédagogique et comment la différenciation pédagogique se manifeste-t-elle en situation d'évaluation?

Un cadre portant sur le transfert des apprentissages (Barnett et Ceci, 2002; Del'Guidice, 1996 ; Frenay, 2004), sur la différenciation pédagogique (Ducette et al., 1996; Perrenoud, 1997. Prud'homme et al., 2005; Przesmycki, 1991) et sur l'évaluation des apprentissages (Del'Guidice, 1999; Louis, 1999; Scallon, 2004) sert de base à l'analyse des données.

Méthodologie

La recherche-action est envisagée ici, à l'instar de Savoie-Zajc (2001) et de Lavoie, Marquis et Laurin (1996, p. 31), comme « une approche de recherche, à caractère social, associée à une stratégie d'intervention et qui évolue dans un contexte dynamique ». La collecte des données s'est déroulée en trois phases : 1- diagnostique (avril à juin 2003), 2- d'intervention (août 2003-février 2005) enchâssée à cinq reprises de portraits évaluatifs provisoires et 3- d'évaluation (février 2005). Divers outils de collecte de données ont été utilisés pour répertorier un maximum de pratiques déclarées et observées : les focus group (n=4), les entrevues individuelles (n=3), les observations en classe (n=2) ainsi que les rencontres collectives de travail

(n=15). Les données ont été transcrites, codées et analysées à l'aide du logiciel N'Vivo.

Principaux résultats

Les résultats mettent en relief que le recours à de multiples projets a été identifié par les enseignants comme étant un dispositif particulièrement riche et pertinent pour soutenir l'apprentissage et le transfert chez les élèves. Par ailleurs, de nombreuses prises de conscience faites pendant le déroulement de la recherche ont conduit les enseignants à arrimer aux projets un ensemble de moyens pour encadrer plus étroitement les élèves, pour leur fournir davantage de rétroaction, pour les amener à réaliser des prises de conscience sur les apprentissages faits et sur ceux qu'il leur reste à faire, pour les amener à s'auto-réguler et pour qu'ils développent des stratégies (cognitives, métacognitives, affectives et de gestion) de plus en plus efficaces. Les représentations de l'évaluation des enseignants se sont graduellement transformées pour devenir non seulement une occasion de certifier des apprentissages et de vérifier les capacités de transfert, mais aussi de soutenir les apprentissages et leurs transferts. Au regard de la différenciation, en lien avec l'évaluation, nous constatons que pour les enseignants, différencier était initialement associé à la diminution des exigences envers certains élèves (ceux jugés les plus faibles) alors qu'à l'égard des autres, des attentes similaires étaient formulées. Au terme de la recherche-action, la différenciation pédagogique implique pour ces enseignants nécessairement l'idée qu'un défi soit posé à l'ensemble des élèves. Les enseignants paraissent aussi maintenant avoir une vision plus éclatée et plus dynamique de la diversité, ce qui rejoint celle de Przesmycki (1991). Cette vision semble ouvrir la voie à une prise de conscience des différents chemins de l'apprentissage pris ici au sens large, à savoir un apprentissage qui inclut à la fois des connaissances disciplinaires, mais aussi une compréhension du monde, le développement personnel et la socialisation des élèves, pour lesquels des outils d'évaluation diversifiés sont essentiels. Si la croyance au potentiel de progression et de réussite semble présente chez les enseignants dès le départ, en revanche, leur réalité les contraints à douter, à plusieurs moments, de la faisabilité d'une telle entreprise. Car il faut du temps, de la créativité pour trouver des moyens inexplorés jusqu'ici pour faire en sorte que tous les élèves apprennent vraiment. La planification serrée des situations d'apprentissage et d'évaluation est ressortie comme une condition absolument essentielle à la mise en place d'une différenciation pédagogique réussie (Presseau et al., 2005). Elle ne règle toutefois pas tout, notamment le manque de temps pour réfléchir et ajuster son action professionnelle, ni ne permet de résoudre l'épineuse question d'équité qui rejaillit systématiquement lorsqu'il est question d'évaluation différenciée.

Quelques références

Barnett, S. et Ceci, S. (2002). When and where do we apply what we learn? A taxonomy for far transfer, Psychological Bulletin, 128(4).

Del'Guidice, J. (1996). Évaluation-régulation et apprentissage de transferts. Construction de situations d'apprentissage à l'école élémentaire, Thèse de doctorat, Université de Provence Aix-Marseille 1.

Del'Guidice, J. (1999). L'évaluation-régulation, vecteur des transferts d'apprentissages. Dans Depover, C. et Noël, B. (Dir.). L'évaluation des compétences et des processus cognitifs (99-114). De Boeck Université.

Ducette, J.P., Sewell, T.E. et Poliner Shapiro, J. (1996). Diversity in education: Problems and Possibilities. Dans Murray, F.B. (dir.), The teacher educator's handbook (323-381). San Francisco: Jossey-Bass.

Frenay, M. (2004). Du transfert des apprentissages au transfert des connaissances. Dans Presseau, A. et Frenay, M. (Dir.), Le transfert des apprentissages. Comprendre pour mieux intervenir (7-48). Sainte-Foy : Presses de l'Université Laval.

Prud'homme, L., Dolbec, A., Brodeur, M., Presseau, A. et Martineau, S. (2005). La construction d'un îlot de rationalité autour du concept de différenciation pédagogique. Journal of the Canadian Association for Curriculum, 3(1), 1-32.

Przesmycki, H. (1991). Pédagogie différenciée. Paris : Hachette éducation.

Zakhartchouk, J-M, (2001). Au risque de la pédagogie différenciée. Paris : Institut National de Recherche Pédagogique.

PARASCHOOL, EVALUATION D'UNE EXPERIENCE D'ENSEIGNEMENT ASSISTE PAR ORDINATEUR AU SECONDAIRE 1 : QUELS ENJEUX ? QUELLES IMPLICATIONS POUR L'ENSEIGNEMENT ?

Anne-Marie Broi, chargée de recherche au Bureau de l'informatique scolaire, Neuchâtel, Suisse, Anne-Marie.Broi@ne.ch

Mots-clés : suivi qualité, cours en ligne, intégration d'une innovation

Résumé :

Dans le cadre du déploiement du Réseau Pédagogique Neuchâtelois (RPN) dans les écoles, l'Office de la statistique et de l'informatique scolaires du canton de Neuchâtel (OSIS), aujourd'hui le Bureau de l'informatique scolaire, offre depuis l'année scolaire 2003-2004 aux enseignants des trois sections du secondaire 1 (préprofessionnelle, moderne, maturités) la possibilité d'expérimenter un dispositif d'enseignement assisté par ordinateur Paraschool (www.paraschool.com) dans les domaines du français (langue 1) et des mathématiques. L'expérimentation de ce programme dans les classes par des enseignants volontaires est prévue jusqu'en 2007.

En vue de mieux comprendre les apports de ce dispositif d'enseignement pour les enseignants et les élèves, les responsables du projet ont mandaté le Centre de compétence - recherche et statistique – (CCRS) pour établir un premier état de situation. Une enquête de type exploratoire a dès lors été menée auprès des utilisateurs de la plate-forme. Afin d'affiner les résultats, cette enquête a été complétée par une investigation sous forme d'interviews auprès d'enseignants et de leurs élèves (étude en cours).

Le but de la communication vise à rendre compte des premiers constats que livre cette expérience. L'accent sera mis sur l'intégration de cet «outil» dans la salle de classe et sur la perception que les enseignants en ont au regard de leur conception de l'enseignement.

Nous aborderons notamment des questions liées à la praticabilité de ce nouveau moyen d'enseignement en fonction des disciplines concernées, des filières scolaires et des caractéristiques des enseignants (enseignants porteurs d'un titre d'instituteur, enseignants porteurs d'un titre académique)[1]. Nous mettrons également en évidence quelques enjeux et implications que soulève l'utilisation de «cours en ligne» pour la pratique: comment les enseignants intègrent-ils la plate-forme dans leur enseignement

[1] Dans le canton de Neuchâtel, le corps enseignant du secondaire 1 est composé d'instituteurs (les généralistes) et d'enseignants porteurs d'un titre académique (les maîtres de branches). Les premiers enseignent aux élèves de classes conduisant à des études courtes (apprentissage), alors que les seconds enseignent à des élèves susceptibles de poursuivre des études longues (Master et Bachelor).

? Quelle est l'influence de cet outil sur leur manière d'enseigner ? Quels effets relèvent-ils sur les apprentissages de leurs élèves ? Quelles craintes disent-ils éprouver par rapport à ce dispositif ? Quels décalages existent-ils entre leur perception du logiciel et ce qu'offre véritablement celui-ci?

Introduction

Contexte de la recherche

Depuis la fin des années 90, le canton de Neuchâtel, à l'instar des autres cantons romands, a marqué sa volonté d'intégrer les technologies de l'information et de la communication (TIC) dans les écoles. Les réflexions menées par les responsables de l'éducation ont abouti à un concept pédagogique lié à l'intégration des TIC ainsi qu'à un concept de formation des enseignants (Jornod, 2004). L'ensemble de cette tâche et son suivi sont menés conjointement par le Service de l'enseignement obligatoire et l'Office de la statistique et de l'informatique scolaires (aujourd'hui le Bureau de l'Informatique Scolaire).

Dans le cadre du déploiement des TIC, un réseau pédagogique a été créé, www.rpn.ch, assorti notamment de la mise à disposition des enseignants et des élèves de logiciels éducatifs et de cours en ligne[2].

Les objectifs

Notre communication a donc pour but de rendre compte du suivi d'une expérience d'enseignement assisté par ordinateur, Paraschool (Broi, 2005). Cette étude mandatée par le responsable de l'OSIS vise à réguler et à adapter le processus d'introduction d'outils pédagogiques liés aux TIC dans les classes. Dans le cadre de notre présentation, un accent sera porté sur l'utilisation que les enseignants font de ce nouvel outil par rapport à l'enseignement-apprentissage et les représentations que cette innovation suscite chez ces derniers au regard de leur pratique habituelle. Nous chercherons à comparer les écarts entre les attentes institutionnelles (Rapport au Grand Conseil: Nouvelles technologies dans les écoles, 2000.) à propos de l'introduction des TIC à l'école et les pratiques déclarées par les enseignants à propos de Paraschool. Ces constats seront mis en perspective avec les points de vue d'un groupe d'enseignants sur l'utilisation de la plate-forme Paraschool (Broi, op. cit., étude en cours). Les questions de recherche portent notamment sur les thématiques suivantes:

---

[2] (…) "Utilisation de cours en ligne: les cours sur ordinateurs consistent par exemple à: 1) disposer en permanence d'une personne qui suit l'activité de l'élève 2) utiliser des séries d'exercices qui apportent eux-mêmes des feed-back 3) réaliser des apprentissages dynamiques et interactifs." (…). "L'enseignement différencié est préconisé (…) dans ce type d'enseignement, l'informatique devient un outil indispensable capable de jouer le rôle d'assistant. Il stimule, relance et corrige les élèves. (…)". In **Rapport au Grand Conseil**: Nouvelles technologies dans les écoles 01.004, 16 décembre 2000. p. 17.

- Comment les enseignants utilisent-ils les fonctionnalités proposées par la plate-forme Paraschool?
- Quelles craintes disent-ils éprouver par rapport à ce dispositif ?
- Comment les enseignants intègrent-ils la plate-forme dans leur enseignement ?
- Quels effets les enseignants relèvent-ils sur les apprentissages de leurs élèves ?
- Comment les enseignants perçoivent-ils la formation ?

Méthode

Recueil et traitement des données

Le recueil des informations a été réalisé, en juin 2005, sur la base d'un questionnaire composé de questions fermées. Des questions ouvertes ont pourtant été prévues afin de préciser et d'affiner certains points. L'analyse statistique du questionnaire a porté sur la fréquence des réponses. Les pourcentages exprimés sont calculés uniquement sur la base des répondants aux questions considérées. Des corrélations ont été établies entre quelques questions afin d'approfondir certains résultats. Les entretiens (en cours d'analyse) effectués auprès d'un petit échantillon d'enseignants ont fait l'objet d'une analyse qualitative dont les principaux résultats seront présentés ici[3].

Population

La population totale de notre étude est composée de 61 sujets (les enseignants utilisateurs de la plate-forme Paraschool durant l'année scolaire 2004 – 2005).

37 enseignants sur 61 ont répondu à l'enquête ce qui correspond à un retour d'environ 60% des questionnaires. Parmi ces réponses, on comptabilise 67% d'hommes et 33% de femmes. Cette représentativité entre hommes et femmes correspond à la population totale de l'enquête.

Parmi les 37 enseignants, on relève 14 enseignants de branches scientifiques et 14 généralistes alors que les maîtres de branches littéraires représentent environ un quart de la population soit 9 enseignants. L'échelle des âges va de 4 personnes de moins de 30 ans (11%) à 13 personnes de plus de 50 ans (35%), environ 54% des répondants soit 20 personnes ont entre 30 et 50 ans.

Un peu moins de 20% des enseignants ont entre 5 et 15 ans d'années d'expérience de l'enseignement, 25% ont entre 16 et 25 ans et 41% plus de 25 ans; seul 14% des enseignants ont moins de 5 ans d'années de pratique.

La plate-forme Paraschool

Avant d'entrer dans l'analyse des résultats, il nous paraît nécessaire de donner quelques précisions sur la plate-forme Paraschool.

---

[3] Les entretiens auprès des enseignants ont été suivis d'un questionnaire aux élèves. Toutefois dans les limites de cette communication, nous n'aborderons pas cet aspect.

Paraschool offre un service «en ligne» de remédiation scolaire personnalisé (www.paraschool.com). Dans notre étude, il s'agit d'exercices à l'attention d'élèves de la 7e à la 9e année scolaire de l'enseignement obligatoire (élèves de 12 à 15 ans); ce service propose des activités dans le domaine de l'enseignement du français et des mathématiques.

Les contenus, conçus et développés par des enseignants du terrain en France et à Neuchâtel, ont été validés pour la majorité des exercices par le Ministère de l'Education nationale française et par les responsables des domaines concernés de la Haute Ecole Pédagogique (BEJUNE).

Ce cours en ligne permet à l'enseignant depuis «l'espace professeur» d'administrer les données de ses élèves, de suivre leurs progrès, d'attribuer un plan de travail à chaque élève à l'aide de la bibliothèque des activités et d'envoyer des courriels à un ou plusieurs élèves pour les aider à gérer leur avancement dans les exercices. Le suivi des progrès des élèves se fait par le biais d'un historique d'utilisation pour chaque élève (les données de sessions : temps passé par jour, mois); l'édition de statistiques permet en plus de connaître le pourcentage des items validés et non validés par niveau (degré scolaire) et par matière.Sur le plan de la littérature, Paraschool pourrait être situé entre l'aspect «artisanal» et «professionnel-industriel", disons «semi-professionnel» (Pochon, 2004, p.14).

Quelques résultats du questionnaire aux enseignants

Appréciation du dispositif Paraschool

D'une manière générale, les enseignants se disent satisfaisants ou très satisfaisants de la plate-forme soit 70%; la même proportion d'enseignants recommanderait ce dispositif à leurs collègues. Les enseignants sont en revanche plus mitigés lorsqu'il s'agit de se positionner par rapport à une échelle progressive (1 étant la valeur minimale et 4 la valeur maximale) : 24% choisissent le code 3 et 22% se situent entre les codes 3 et 4. La satisfaction des autres enseignants (38%) se situe entre le code 1 et l'intervalle 2 et 3. 16% des enseignants ne répondent pas à cette question La perception que les enseignants ont de l'appréciation de leurs élèves se rapproche de leurs propres points de vue. Selon les enseignants environ 19% des élèves choisiraient entre le code 3 et 4 et 5% dans le code 4.

Si du point de vue du nombre d'enseignants, la plate-forme est majoritairement utilisée par des maîtres de mathématiques, notamment des hommes, on relève indépendamment du nombre d'enseignants représentés dans les différentes disciplines que les enseignants généralistes fréquentent plus régulièrement la plate-forme que leurs collègues spécialistes de mathématiques et de français.

Les réponses au questionnaire montrent que la fréquence d'utilisation est corrélée positivement avec le degré de satisfaction; ainsi 11 enseignants sur 26 apprécient la plate-forme et l'utilisent plus d'une fois par semaine ou une fois par semaine. Les enseignants, plutôt négatifs, voire très négatifs, à propos de Paraschool, y vont une fois par mois, rarement ou jamais.

De même, la fréquence d'utilisation de l'espace professeur est corrélée positivement avec la fréquentation de la plate-forme : 11 enseignants sur 37 fréquentent souvent Paraschool; parmi ces enseignants, 7 fréquentent aussi souvent l'espace professeur. 12 enseignants sur 21 utilisent rarement ou jamais la plate-forme et utilisent aussi rarement voire jamais l'espace professeur.

Par ailleurs, la facilité d'utilisation de l'espace professeur n'est pas une garantie de sa fréquentation; 13 enseignants sur 21 y vont rarement ou jamais bien qu'ils estiment l'utilisation de l'espace professeur facile voire très facile.

L'utilisation de Paraschool par les élèves

Les élèves utilisent essentiellement le dispositif Paraschool à l'école, selon les enseignants; ils ne sont plus que 25% à réaliser des exercices depuis la maison. Nous verrons plus loin dans ce document les raisons invoquées par les enseignants pour argumenter l'utilisation du «cours en ligne» par les élèves en dehors de l'école.

L'adéquation entre les activités proposées et les besoins des enseignants

Pour une majorité d'enseignants (75%), les activités proposées permettent d'atteindre partiellement les objectifs du programme.
Ce constat provient surtout des enseignants qui fréquentent plus d'une fois par semaine la plate-forme, soit 10 enseignants sur 28.

Indépendamment des disciplines, les activités proposées en mathématiques et en français répondraient en grande partie ou partiellement aux besoins de 30 enseignants sur 37. Les enseignants généralistes et les maîtres de français y trouveraient davantage leur compte que les enseignants de mathématiques. Cette remarque est sans doute liée au fait qu'il existe moins d'exercices à disposition dans ce domaine que dans celui du français.

L'appréciation de la formation et ses effets sur l'enseignement

D'une manière générale, les enseignants estiment à 89% la formation suffisante ou partiellement suffisante pour utiliser la plate-forme.

On relève que les enseignants qui apprécient la plate-forme sont aussi les mêmes à considérer suffisant ou partiellement suffisant le temps prévu pour la formation (24 enseignants sur 26)
Le croisement des questions entre la satisfaction face à la formation et la fréquence d'utilisation de l'espace professeur montre que les enseignants satisfaits par la formation, n'utilisent pas nécessairement, ou du moins très irrégulièrement, cet espace de travail. C'est le cas de 24 enseignants sur 32. Ces réponses seront étayées par l'analyse des questions ouvertes et des interviews.

Mise en perspective de deux prises d'information

Huit questions ouvertes ont été prévues dans le questionnaire afin de permettre aux enseignants d'exprimer de manière plus qualitative leur appréciation de l'expérience. Les questions ouvertes ont suscité des réponses chez plus de 20% des enseignants de l'enquête. Certaines questions ont rencontré davantage d'écho que d'autres, parmi lesquelles l'appréciation du dispositif, les questions relatives aux développements et perspectives d'un tel outil pour favoriser l'enseignement apprentissage. Une question portait sur la formation reçue.

Afin d'affiner cette première prise d'information, une étude exploratoire sous forme d'entretiens a été effectuée, durant l'année scolaire 2005 – 2006, auprès d'un petit échantillon d'enseignants (op. cit. étude en cours). Nous nous centrerons, dans ce sous-chapitre, sur les questions qui ont suscité le plus de réponses chez les deux populations d'enseignants de notre étude. Nous chercherons à dégager les points communs, les différences et les spécificités du discours des différents protagonistes. Nous prendrons également appui sur certaines données chiffrées de l'enquête pour argumenter nos constats.

Les points de vue sur le dispositif

Les principales qualités attribuées au dispositif concernent des dimensions pédagogiques, à savoir le travail individuel, la motivation des élèves face à ce nouveau moyen d'enseignement et l'autonomie des élèves par rapport à l'organisation du travail. Le plan de travail de l'élève organisé par le maître offre la possibilité de différencier l'enseignement. Selon les enseignants, le plan de travail facilite l'avancement dans le traitement des notions du programme.

Les enseignants des deux populations pensent que la plate-forme joue le rôle «d'exerciseur", elle permet le drill et la consolidation des notions. L'organisation des activités permet la révision systématique des notions.
Par contre, pour introduire une nouvelle notion, les enseignants utilisent d'autres moyens d'enseignement.

Les points de vue plutôt négatifs et/ou mitigés à propos du dispositif sont plus nombreux et davantage argumentés par les enseignants que la description des aspects positifs.

a) Par ordre d'importance, les enseignants souhaiteraient une meilleure adéquation entre les activités proposées par Paraschool et le programme. Ce point de vue provient surtout des enseignants de mathématiques, alors que les enseignants des branches littéraires et les généralistes trouvent adéquat le lien entre les activités et le programme. Actuellement en Suisse romande, l'enseignement-apprentissage des mathématiques est centré sur la résolution de problèmes basés sur des principes socioconstructivistes. Les activités de mathématiques proposées dans Paraschool, surtout pour des raisons techniques, selon les dires des concepteurs, sont plus «classiques"; certains enseignants de mathématiques se disent alors déçus par rapport à leurs attentes. Ils ressentent une «rupture didactique» entre les moyens d'enseignement officiels (les fiches de l'élève) et les activités de Paraschool.

b) Les remarques concernant les aspects techniques portent sur les difficultés des enseignants à intégrer les différentes fonctionnalités du dispositif. L'utilisation de l'espace professeur - établir un plan de travail, suivre la progression des élèves, etc. - poserait d'importants problèmes aux enseignants néophytes en informatique. Le degré d'aisance de l'enseignant par rapport à la plate-forme informatique a dès lors des retombées sur la gestion de la classe et l'individualisation du travail. Le point de vue de cette enseignante reflète les propos de ses collègues *«Je suis plus attentive en classe avec eux que lorsque je suis dans une salle informatique, je ne suis pas certaine de pouvoir leur répondre s'ils rencontrent un pépin technique.* »

Le point de vue des enseignants sur l'enseignement à distance

Dans le questionnaire, les trois-quarts des enseignants estiment que moins de 25% des élèves utilisent le dispositif à la maison. En même temps, ils sont unanimes à déclarer dans les questions ouvertes (20 enseignants sur 32) que la connexion à la plate-forme

en dehors de l'école représente une plus-value pour les élèves; l'élève peut ainsi avancer à son rythme dans les exercices fixés du plan de travail.

Corollairement à ces déclarations, les interviews autorisent quelques hypothèses sur le décalage entre des constats chiffrés et un discours sur les pratiques. Les enseignants interrogés inciteraient peu leurs élèves à utiliser Paraschool depuis la maison. Ils utilisent aussi rarement la communication par courriel pour dialoguer avec leurs élèves. Certains enseignants invoquent des raisons d'égalité entre les élèves, - tous n'auraient pas à disposition un ordinateur à la maison[4] - et d'autres considèrent que donner des activités à faire à la maison en lien avec Paraschool ne fait pas partie (…) *«de ce qui a été convenu avec l'élève»* (...). Tout se passerait un peu comme si la présence d'ordinateurs à l'école influait sur «le contrat didactique» maître-élèves (Brousseau, 1996). Ces constats montreraient une utilisation hybride de Paraschool par les enseignants. L'option travail à distance à la maison serait sous-utilisée par rapport au poids donné à l'option de travail en salle informatique à l'école. On relève ainsi dans le discours des enseignants une tendance à détourner la fonction d'apprentissage à distance au profit d'une pédagogie plus traditionnelle. La plupart des enseignants préfèrent guider l'enseignement en privilégiant les contacts *«directs»* avec les élèves.

Des interrogations sur la dimension didactique de Paraschool

Lors des interviews, plusieurs enseignants relèvent l'aspect ludique de la plate-forme. Certains réagissent fortement par rapport à l'introduction de l'ordinateur à l'école: l'ordinateur ferait partie des loisirs et n'aurait pas sa place en classe. D'autres enseignants, dans les interviews et dans le questionnaire, déplorent l'assimilation faite par les élèves entre les jeux pratiqués à la maison et les exercices scolaires. Selon ces enseignants, les élèves prendraient moins au sérieux leur travail en utilisant Paraschool que lorsqu'ils réalisent des exercices de type papier crayon en classe. Ce point de vue apparaît dans les mêmes proportions chez les enseignants acquis aux nouvelles technologies et chez ceux qui émettent des doutes sur l'efficacité pédagogique de Paraschool. Environ un enseignant sur cinq partage cet avis. Quelques enseignants relèvent (…) *«des effets pervers»* (…) à propos de la dimension ludique de Paraschool. Certains élèves chercheraient davantage à gagner des points[5] qu'à réellement comprendre les notions traitées dans les exercices. L'assimilation de la plate-forme à un jeu inciterait ces élèves, notamment les plus faibles, à répondre au hasard jusqu'à l'obtention de la réponse «juste» alors que dans leur pratique habituelle, ces enseignants disent mettre l'accent sur des démarches réflexives et accorder de l'importance à la méta-réflexion dans la construction des apprentissages. La part du hasard dans la construction du logiciel freinerait en quelque sorte la confiance que les enseignants mettent dans cet instrument : *«La part du hasard, c'est propre à la*

[4] Les centres scolaires disposent d'ordinateurs dans l'école pour ce type d'élèves.

[5] Un certain nombre de points est attribué à chaque exercice. Ces points sont ensuite comptabilisés et traduits en pourcentage au moment du bilan des activités. Des couleurs (vert, orange et rouge) indiquent à l'élève le degré de réussite de chaque exercice.

construction des logiciels, mais la fiabilité d'un tel instrument c'est celle qui laisse le moins de place au hasard» relève une enseignante.

L'intégration du «cours en ligne» dans l'enseignement

Pour l'ensemble des enseignants de notre étude, la plate-forme représente un «outil» complémentaire. On relève toutefois, chez les enseignants, deux tendances par rapport à l'intégration du logiciel dans les pratiques.

1e tendance : certains enseignants sont acquis au principe des cours en ligne: il s'agit notamment des enseignants de mathématiques et des enseignants généralistes qui utilisent Paraschool pour les mathématiques. 10 enseignants sur 14 déclarent pouvoir tirer parti des activités dans la conduite de leur enseignement. Cette citation résume ce point de vue: *«Je fais des «va-et vient» entre les moments de recadrage avec toute la classe et les moments de travail avec Paraschool»*.

2e tendance : d'autres plus réservés, considèrent la plate-forme comme un «outil» supplémentaire par rapport aux moyens d'enseignement officiels. Cette population est représentée par les enseignants qui fréquentent peu souvent la plate-forme ou qui éprouvent des difficultés à utiliser l'espace professeur pour gérer le travail de l'élève. Les raisons invoquées portent sur la nécessité d'avoir une vision constante sur le travail des élèves. Pour ces enseignants, tout se passe un peu comme si le fait de travailler avec Paraschool masquait une partie du travail de l'élève. Les enseignants perdraient ainsi ,en quelque sorte, le regard sur l'avancement du travail de l'élève: *«je ne sais pas ce qu'il fait lorsqu'il est devant l'ordinateur" (…) «je ne sais pas s'il a des difficultés»* relève un de ces enseignants. Ce point de vue corrobore les réponses aux questions fermées du questionnaire au sujet du degré d'influence de Paraschool : 24 personnes sur 32 estiment que Paraschool a peu d'influence sur leur manière d'enseigner.

Un autre fait marquant est à souligner à propos des retombées de l'utilisation du logiciel. L'enseignement à partir d'une plate-forme informatique remettrait en question le rôle des enseignants concernant la conduite de l'enseignement-apprentissage. Avec l'introduction de l'informatique, certains enseignants ont en effet le sentiment d'être pris en porte-à-faux entre leurs convictions en matière de rapport au savoir et la place que prend l'ordinateur dans le processus de l'enseignement-apprentissage, par exemple les propositions d'exercices à l'élève par le logiciel. La remarque suivante résume la crainte de cette mise en concurrence pour les enseignants : *«Les personnes qui ont réalisé cet outil (Paraschool) n'ont pas pensé combien cela change notre pratique, je dois mettre en place une nouvelle pédagogie»* (…) *«Dans ma pratique, j'ai besoin que mes élèves m'appellent et me posent des questions. »*

Les effets de Paraschool sur les apprentissages

Les enseignants, toute population confondue, sont timorés lorsqu'il s'agit de se prononcer sur les effets de Paraschool auprès des élèves. Certains ne se prononcent pas car ils estiment être des utilisateurs occasionnels et les autres demandent

davantage de recul avant de se prononcer. Ces résultats corroborent les réponses au questionnaire : 10 enseignants sur 32 relèvent des effets sur les apprentissages alors que 14 se disent moyennement convaincus et 8 ne se prononcent pas. Toutefois, les enseignants, quel que soit leur degré d'utilisation de Paraschool, sont unanimes sur la nécessité d'une régularité et d'une systématicité dans l'utilisation de Paraschool afin d'espérer des bénéfices pour l'élève.

En marge des déclarations, les enseignants mentionnent dans les interviews leur besoin d'un suivi personnalisé en classe pour pallier leurs lacunes sur le plan technique. Ces mêmes personnes souhaiteraient des points de situations réguliers avec des collègues et les responsables de Paraschool : *«C'est en échangeant de vive voix que l'on apprend des collègues»*.

Principales difficultés de l'introduction du «cours en ligne»

L'introduction de la plate-forme Paraschool dans les classes n'a pas été sans problème, les enseignants ont dû surmonter un certain nombre de difficultés. Les raisons qui ont freiné le processus d'introduction de Paraschool sont regroupées ci-dessous selon cinq catégories [6]:

- *des raisons à caractère organisationnel*: les enseignants, notamment les spécialistes de branches, mentionnent souvent la difficulté d'intégrer la plate-forme dans leur horaire ou encore de prévoir et gérer le déplacement en salle informatique. Malgré la mise à disposition de salles informatiques jugée suffisante dans les établissements, le nombre d'élèves joue un rôle déterminant dans le suivi de leur travail, en particulier pour les élèves de classes préprofessionnelles.

- *des raisons à caractère pédagogique* : Nous sommes d'avis avec Blanchet & al. (1996) que les difficultés les plus importantes concernent les conditions pratiques et pédagogiques liées au contexte, notamment la gestion de l'enseignement, qui n'est pas sans relation avec les problèmes de la didactique d'ailleurs. Les enseignants de notre expérience, aussi bien les enseignants de français que de mathématiques, soulignent en effet que l'articulation entre les exercices de Paraschool et les moyens d'enseignement officiels nécessitent des recadrages systématiques en classe afin que les activités réalisées dans des contextes différents fassent sens pour les élèves.

- *des raisons à caractère didactique :* Dans le prolongement des constats ci-dessus, les enseignants de mathématiques relèvent que les exercices de Paraschool seraient davantage représentatifs d'une approche de l'enseignement centrée sur une méthode guidée et répétitive que sur une approche basée sur la compréhension et la capacité d'utiliser des connaissances procédurales pour résoudre des problèmes (telle que développée dans les moyens d'enseignement officiels). Toutefois, en même temps et au-delà de leurs déclarations, ces enseignants reconnaissent les apports du logiciel dans sa fonction de «répétitoire» pour les élèves. Ils feraient alors des «va-et-vient» entre les moyens d'enseignement officiels et les activités du logiciel. Dans le domaine du français, en revanche, le contenu des activités proposées par Paraschool semble

[6] Cette catégorisation est une adaptation issue du questionnaire de Blanchet, A. & al. (1996)

correspondre davantage aux exercices réalisés habituellement en classe. Pour ces enseignants, les freins mis par rapport à l'utilisation de Paraschool seraient alors principalement liés à des questions techniques.

- *des raisons liées au caractère de l'activité elle-même*: Les remarques concernent moins les activités elles-mêmes que les fonctionnalités du logiciel : l'élaboration d'un plan de travail, le suivi personnalisé du parcours de l'élève, la communication par courriel avec les élèves pour gérer l'avancement de leur travail. Les enseignants de français et les généralistes exploitant les activités de français sont plus nombreux à mentionner ce type de difficulté. Ces enseignants préfèrent de loin les contacts directs avec les élèves.

- *des raisons à caractère motivationnel et métacognitif*: Certains enseignants trouvent trop ludique «l'habillage» des activités. Les élèves auraient tendance à assimiler les exercices aux logiciels récréatifs utilisés à la maison. Ces enseignants, plus particulièrement les personnes qui utilisent Paraschool dans le domaine du français, sont dubitatives sur ses effets auprès des élèves.

Corollairement à ces doutes, la fonction auto-corrective intégrée dans le logiciel - le feed-back à l'élève - suscite des interrogations. Selon les enseignants, et toutes disciplines confondues, les élèves chercheraient à parvenir rapidement à la réponse juste au lieu de chercher à comprendre leur erreur. Cette stratégie est surtout utilisée par les élèves en difficulté.
D'autres constats touchent les représentations des enseignants à propos de leur rôle. Ainsi, la place que prend la plate-forme interactive par rapport à l'enseignement habituel les interpelle. Certains enseignants ressentent une concurrence entre le dispositif informatique et leur mission d'enseignant. Les cours en ligne priveraient en quelque sorte l'enseignant d'être en prise directe avec le processus d'apprentissage de l'élève.

Discussion

L'introduction de l'informatique scolaire … des questions récurrentes au fil du temps
Afin de prendre la mesure des apports de l'informatique dans le temps, nous nous proposons, dans les limites de la comparaison, de dépasser les aspects locaux de notre étude en profilant les questions restées en suspens dans notre étude à la lumière des résultats de l'enquête conduite par Blanchet & al. (op. cit.). Cette étude de nature exploratoire portant sur les effets de l'introduction de l'informatique à l'école obligatoire a été menée dans le milieu des années 1990, auprès de 317 enseignants des cantons romands et du Tessin, dont Neuchâtel[7]. La mise en perspective des deux études sur des domaines tels que l'informatique et les apprentissages, l'influence de

[7] En 1995, une période/semaine (45') est prévue à l'horaire en 9e année et une période en 8e dans certains centres scolaires (sensibilisation et techniques de base, traitement de texte, dessin, tableaux, fichiers, etc). Actuellement, une période est à l'horaire en 8e année. Par ailleurs, dans le plan d'études du secondaire 1 (PENSE) et en cohérence avec le nouveau plan d'étude-cadre romand (PECARO), l'informatique est devenue une discipline transversale.

l'informatique sur les représentations des enseignants, les difficultés rencontrées par ces derniers montrerait qu'il n'y aurait pas de décalage sensible entre ce que disent les enseignants interrogés dans notre enquête et ce que révèlent les résultats de Blanchet & al. (op. cit.). Ces concordances de points de vue concernent surtout des domaines tels que la motivation, l'autonomie des élèves, la différenciation de l'enseignement. Dans les deux populations, le degré d'accord n'est pas toujours le même selon l'expérience de l'enseignant avec l'ordinateur. Les enseignants qui ont eu une expérience informatique avec leurs élèves sont plus positifs à son égard. De même, le discours sur les difficultés à surmonter dans la conduite des activités de type multimédia est proche: le manque de temps, le nombre d'élèves, la gestion de la classe. Les deux populations ont conscience du changement à amorcer dans leurs pratiques pour intégrer les TIC. L'effort à fournir pour relever ce défi est un élément présent dans les discours des enseignants. Cette proximité des constats à dix ans d'intervalle nous interpelle à plus d'un titre. En effet, que penser de l'aptitude des enseignants «à faire leur» une innovation? Que dire aussi des réels apports pédagogiques des TIC pour l'enseignement ? Et sur le plan institutionnel, qu'ont fait les autorités scolaires pour donner suite aux constats de la recherche ? Ces questions sont d'autant plus brûlantes aujourd'hui dans le contexte politique actuel que des mesures d'économie sont discutées à tous les niveaux de l'école pour assainir les finances des cantons.

Conclusion

Pour conclure, nous essayerons tout d'abord de tirer quelques lignes de forces de nos résultats; nous avancerons quelques éléments qui caractérisent l'utilisation du dispositif «Paraschool» en classe par rapport aux visées éducatives promulguées par les responsables scolaires (op. cit. 2000). Nous esquisserons ensuite quelques pistes de développement et des perspectives.

Il faut toutefois encore relever que ces résultats sont à considérer avec prudence étant donné qu'il s'agit d'une première approche d'une part, et que, d'autre part l'enseignement à distance n'est qu'un aspect parmi d'autres du phénomène complexe qu'est la mise à disposition d'outils informatiques au service de l'enseignement-apprentissage.

Des options politiques sur les TIC vers l'implantation d'un outil pédagogique dans les classes

Que retenir dès lors de cette étude au regard des visées pédagogiques exprimées dans les documents officiels qui ont précédé l'introduction de l'informatique scolaire dans le canton de Neuchâtel (2000, op. cit.) ? Force est de relever qu'au-delà des résultats globalement positifs de notre étude, nos investigations auprès des enseignants montrent que le processus d'intégration des TIC est en route mais que la distance à parcourir pour atteindre les objectifs stratégiques des politiques est encore longue. Les enseignants interrogés reconnaissent que l'utilisation de logiciels éducatifs, plus particulièrement la plate-forme Paraschool, entraîne une transformation de la démarche pédagogique. Leur demande à être mieux accompagnés dans ce processus d'innovation nous conduit à penser que les enseignants seraient plutôt d'avis que le

challenge de l'introduction des TIC est nécessaire pour préparer l'élève à entrer dans la vie active et qu'ils sont d'accord de s'y investir. Pourtant, dans leur pratique, tout porterait à penser que les enseignants privilégieraient des pratiques anciennes. La plate-forme Paraschool reste un complément, une aide utile pour effectuer le travail scolaire habituel. Ainsi, l'informatique prend sa place dans l'école, mais une place secondaire.

Quelques éléments prospectifs et de développement

Les questions abordées, dans les limites de cette étude, montrent modestement qu'une étape a été franchie entre le monde de l'école et celui de l'informatique, mais que le processus d'intégration de ces «nouveaux produits» demande encore des recherches approfondies tant sur le plan de la compréhension des processus d'apprentissages des élèves que sur celui des effets de l'informatique en terme de plus - value pour l'enseignement. Une approche systémique serait alors à privilégier.

- Sur le plan pédagogique:

a) Une attention particulière est à porter sur la façon dont les connaissances sont organisées et transmises ensuite au travers des activités. On peut citer plus particulièrement la cohérence à établir entre les objectifs visés dans des plans d'études et la table des matières des activités. Dans la version actuelle de Paraschool, la table des matières est centrée sur des contenus alors que pour aller dans le sens des nouveaux plans d'études, elle devrait être formulée en termes d'objectifs d'apprentissage modulés par degré de complexité.

b) La transcription d'un contenu pédagogique sur support multimédia peut générer de nouveaux «gestes» pédagogiques. L'analyse didactique des tâches et des activités est dès lors nécessaire pour mieux instrumenter les enseignants dans la maîtrise de ces gestes.

c) Dans le domaine de la gestion des activités par l'ordinateur, une exploitation efficace des exercices impliquerait d'inclure une phase de «debriefing». Le «debriefing» engage l'élève au transfert des connaissances. Une investigation serait à faire pour faire évoluer cette dimension dans Paraschool. Si le logiciel doit permettre «d'apprendre à apprendre», une proposition d'aide à la suite d'une réponse inexacte ne suffit pas, d'autres outils d'auto-évaluation seraient à prévoir.

d) Toutefois, corollairement au développement de l'aide à l'élève par l'ordinateur, il s'agirait de repenser le rôle et la place de l'enseignant dans le triangle «maître-élève-démarche pédagogique». De nouvelles postures sont à formaliser pour engager l'enseignant à passer de la relation «maître-élèves-moyen d'enseignement de type papier-crayon» à une interaction «maître-élèves-TIC».

- Sur le plan institutionnel :

a) D'une manière générale, et d'autant plus dans le domaine des nouvelles technologies, il ne faut pas sous-estimer le temps que demande l'intégration d'une innovation dans l'école. Des prises d'information systématiques et régulières, selon une méthodologie du projet adaptée au contexte, seraient à mettre en œuvre dans un processus de suivi des expérimentations pour effectuer les ajustements souhaités et prévoir les adaptations de la formation des enseignants. Il s'agit à la fois de former mais aussi d'outiller mieux les enseignants en les rendant attentifs à la relation à l'outil et à la nature de cet outil.

- Sur le plan politique :

a) Au vu des enjeux de l'informatique dans le monde scolaire et des attentes de la société, le pôle des décideurs ne peut être occulté dans le processus d'intégration des TIC à l'école. Nous pensons à l'instar de Pochon (2004, p. 175) que des canaux de communication, en particulier «des boucles» de retour seraient à favoriser, éclairant notamment les usages de l'informatique pour orienter les options à prendre et les investissements financiers à prévoir.

Références

Baron, G.-L; & al. (eds) (2002). *Les technologies en éducation: Perspectives de recherche et questions vives*. INRP/IUFM de Basse-Normandie.

Blanchet, A. & al. (1996). *Informatique scolaire 10 ans d'expériences. Etude exploratoire sur les effets de l'introduction de l'ordinateur à l'école obligatoire.* Neuchâtel : IRDP.

Broi, A.-M. (2005). *Projet Paraschool, Résultats du questionnaire aux enseignants,* Neuchâtel: *OSIS-CCRS*.

Brousseau, G.(1996). Fondements et méthodes de la didactique des mathématiques. In J. Brun, *Didactique des mathématiques*, Delachaux et Niestlé, Coll. Textes de base en pédagogie, 45 à 143.

Jornod, V. (2004). Formation décentralisée du corps enseignant. In *Entre techniques et pédagogie: la création de contenus multimédia pour l'enseignement et la formation*. (pp.96 – 102) IRDP. Neuchâtel.

Perriault, J. (2002). *Education et nouvelles technologies, théories et pratiques.* Nathan-Université.

Pochon, L.-O. & al. (eds). (1997). *L'ordinateur à l'école de l'introduction à l'intégration*. Neuchâtel: IRDP; Lausanne: LEP.

Pochon, L.-O. & al. (eds). (2004). *Entre techniques et pédagogie: la création de contenus multimédias pour l'enseignement et la formation*. IRDP. Neuchâtel.

Nouvelles technologies dans les écoles. *Rapport au Grand Conseil No. 01.004.* Neuchâtel 16 décembre 2000.

UN OUTIL D'EVALUATION DU PROFIL ET DES COMPETENCES DES DEMANDEURS D'EMPLOI

Claude Houssemand, Romain Martin , Raymond Meyers, Université du Luxembourg

Mots-clés : chômeurs, profilage, évaluation informatique

Résumé :
Le profilage des chômeurs est devenu une technique de suivi utilisée par les services publics de l'emploi (SPE) dans de nombreux pays. Il vise à adapter l'intervention de ces services en fonction des profils des demandeurs d'emploi. Mais les outils utilisés se sont jusqu'à maintenant centrés uniquement sur les variables socio-professionnelles et socio-économiques traditionnelles (âge, sexe, nationalité, diplôme, expérience professionnelle, etc.). Un outil de profilage multidimensionnel évaluant l'ensemble des variables pouvant prédire la durée du chômage au Luxembourg, y compris les variables psychologiques et sociales, a été mis au point. Il s'agit d'un instrument qui a été développé sous une forme informatique conviviale. Cet outil ne nécessite aucune compétence informatique de la part des utilisateurs et provoque moins d'anxiété des ordinateurs que les PC clavier-souris. Ces points sont importants pour une utilisation avec des publics peu ou pas qualifiés, ce qui est le cas de la majorité des demandeurs d'emploi inscrits dans les SPE luxembourgeois. Cet outil permet aussi des évaluations plus objectives, dégagées des partis pris subjectifs sur les personnes. Il a été testé sur une cohorte de 385 demandeurs d'emploi au Luxembourg. Certains résultats de cette passation sont présentés afin de montrer que ce genre d'instrument permet d'obtenir avec ce public des données qui montrent une validité de construit satisfaisante.

Le profilage dans les services publics de l'emploi

Le profilage des chômeurs est devenu une technique de suivi utilisée par les services publics de l'emploi (SPE) dans de nombreux pays (Australie, USA, Pays-Bas, Allemagne, Suisse, etc.). Il vise à ajuster l'intervention de ces services en fonction des profils différentiels des demandeurs d'emploi. L'idée de base de cette démarche est que le cas de tous les chômeurs n'est pas semblable et qu'il faut dès lors établir des profils spécifiques. Toutes les catégories de demandeurs d'emploi ne courent pas le même risque de devenir chômeur de longue durée. Pour prendre un exemple simple, on sait que les personnes sans qualification et âgées de plus de 40 ans ont plus de difficultés à retrouver de l'emploi que les personnes plus jeunes et qualifiées. Ceci dit, la majorité des demandeurs d'emploi inscrits à l'Administration de l'Emploi (ADEM) luxembourgeoise trouve du travail endéans quelques mois et seule une minorité dépasse la durée critique d'une année d'inscription.

Le profilage permettrait dès lors des prises en charge différentes selon le profil et donc selon le degré de vulnérabilité des chômeurs. Au lieu de prévoir des formules

relativement identiques pour tous, des mesures plus ajustées pourraient être ciblées en direction des différentes catégories de personnes.

Lors d'une récente conférence internationale qui s'est tenue à Nuremberg des 12 au 14 janvier 2005, sur invitation de la Commission européenne, des expériences réalisées notamment en Australie, au Canada, aux Etats-Unis, aux Pays Bas, en Suède, en Allemagne, en Suisse et en Belgique ont été présentées. Par exemple, aux Pays-Bas, chaque chômeur remplit dès son inscription dans les services publics de l'emploi un questionnaire-type, le *Kansmeter*, qui permet de mesurer son niveau d'employabilité sur une échelle allant de 1 (directement plaçable) à 4 (reclassement peu envisageable avant un long trajet), et de mettre ainsi en œuvre des parcours d'intégration adaptés. Une étude de profilage danoise (*Danish Employability Profiling System*) a quant à elle permis de prédire correctement dans 66% des cas le risque de chômage de plus de 26 semaines (Rosholm, Svarer & Hammer, 2004).

Limites du profilage traditionnel

Mais les outils utilisés se sont jusqu'à maintenant centrés uniquement sur les variables socio-professionnelles et socio-économiques traditionnelles (âge, sexe, nationalité, diplôme, expérience professionnelle, etc.). Or, de nombreuses recherches universitaires ont également montré l'importance de variables psychologiques et comportementales dans la façon de faire face (*coping*) des chômeurs et dans leur rapidité de sortie du chômage (pour des revues, voir Cortese & Meyers, 2004; Kanfer, Wanberg & Kantrowitz, 2001). Des recherches réalisées précédemment au Luxembourg avaient d'ailleurs montré l'incidence corrélationnelle des ces facteurs dans le chômage de longue durée (Houssemand et al., 1997). Par la méthode des interviews en face à face, 777 demandeurs d'emploi inscrits depuis plus d'un an à l'ADEM ont été soumis à un questionnaire systématique. Une au moins des quatre caractéristiques socio-psychologiques suivantes se retrouvait chez trois quarts environ de ces personnes :

- 29,5% des demandeurs d'emploi de longue durée ne recherchaient pas réellement un emploi,
- 38,2% de ces personnes se jugeaient inaptes pour un emploi régulier,
- 47% de ces personnes ne réalisaient pas ou peu de démarches personnelles pour une insertion professionnelle,
- 36,5% de ces demandeurs d'emploi avaient une lacune sur au moins deux facteurs psycho-sociaux sur les trois qui avaient été mesurés (facteurs relationnels, présentation physique et projection dans l'avenir).

Les mêmes auteurs ont constaté que 19,4% de ces personnes cumulaient deux caractéristiques, 16,5% en cumulaient trois et 7,7% cumulaient les quatre (Houssemand et al., 1997).

Cette étude est intéressante car elle tente d'expliquer le chômage de longue durée, non seulement par des caractéristiques démographiques, professionnelles et économiques, mais aussi par l'introduction de dimensions souvent non considérées dans ce type de recherches, à savoir des variables psychologiques et sociales. Il semblerait donc que certaines caractéristiques personnelles des demandeurs d'emploi aient une influence sur la durée de leur interruption de travail.

Les difficultés de l'étude des facteurs affectant le chômage

Mais certaines difficultés apparaissent dans l'étude des facteurs affectant la durée du chômage, difficultés s'étant d'ailleurs manifestées dans l'étude qui vient d'être citée. L'absence d'échantillon comparatif de chômeurs de courte durée interdit de pouvoir conclure quant à l'ampleur différentielle des phénomènes constatés. Par exemple, quelle est la différence entre chômeurs de courte durée et chômeurs de longue durée quant à leur capacité ou incapacité à se projeter dans l'avenir ?
Et, en l'absence d'un suivi longitudinal, il est également difficile de juger du sens des relations mises en évidence dans cette étude. Les analyses statistiques présentées ne mettent en évidence que l'existence de covariations entre la durée du chômage de plus d'un an et les caractéristiques citées des demandeurs d'emploi. Par exemple, l'auto-évaluation négative de l'aptitude au travail est-elle à l'origine du chômage ou une conséquence de celui-ci ? Il était impossible dans cette étude de confirmer l'une ou l'autre de ces hypothèses, d'où l'intérêt d'une nouvelle étude longitudinale que nous avons mise en chantier.

Les variables en rapport avec la durée du chômage

Cette étude partait donc de l'idée d'évaluer l'ensemble des variables qui pourraient être pertinentes pour être en rapport avec la durée de chômage des personnes inscrites dans les SPE luxembourgeois. Une revue de la littérature internationale (Cortese & Meyers, 2004) a permis de dégager celles qui méritaient une investigation plus poussée au Luxembourg (voir tableau 1).

Tableau 1 - Variables potentiellement en rapport avec la durée du chômage

Variables	Variables influençant la durée du chômage	Variables pouvant être affectées par la durée du chômage
Variables démographiques	Age, sexe, nationalité, etc.	Statut marital, etc.
Variables socio-économiques	Formation, compétences, revenus, etc.	Formation d'insertion, mesures pour l'emploi, etc.
Variables psychologiques	Intelligence, sentiment de compétence, estime de soi, etc.	Extraversion, centralité du rôle professionnel, etc.
Variables de comportement	Compétences sociales, emploi du temps, stratégies de recherche d'emploi, etc.	Compétences sociales, stratégies de recherche d'emploi, etc.

On peut citer à titre d'exemple les variables démographiques (comme l'âge et la nationalité), les variables socio-économiques (comme la formation initiale et le métier), les variables psychologiques (comme l'estime de soi et l'extraversion) et les variables comportementales (comme l'emploi du temps et les stratégies de recherche d'un emploi). Nous avons distingué deux cas de figure : soit la variable influence la durée de chômage (deuxième colonne du tableau), soit la variable est affectée par la durée de chômage (troisième colonne).

Une étude de profilage multidimensionnel

L'Université du Luxembourg est en train de réaliser cette étude de profilage en collaboration avec l'ADEM et financée dans le cadre du programme VIVRE du Fonds National de la Recherche (projet FNR 02/05/07). Elle vise à tester ces différentes hypothèses. Une cohorte de chômeurs inscrits pour la toute première fois sera suivie sur une période de deux ans avec prise de données tous les six mois. Lors de la première évaluation, qui vient de se terminer, un profil très complet de chaque personne a été établi. Ce dernier reprend d'abord des données socio-économiques et socioprofessionnelles comme l'âge, le statut marital, le niveau de formation, etc. Il comprend également un profil de personnalité établi par questionnaire et incluant des informations pouvant jouer un rôle dans le maintien au chômage. A titre d'exemple, on peut citer l'estime de soi, l'ouverture aux expériences nouvelles, la projection dans l'avenir, le contrôle perçu par rapport au chômage, l'auto-efficacité, l'intelligence générale, etc. Et finalement, on a mesuré certaines démarches comportementales du chômeur, comme sa façon de chercher un emploi, sa mobilité, sa capacité à se faire aider, etc.

A partir d'une revue de la littérature, nous avons intégré dans notre outil des échelles psychométriques qui ont fait leur preuve et qui ne sont pas trop longues. A titre d'exemple, on peut citer le *General Health Questionnaire* en 12 items (Goldberg, 1992), la *Rosenberg Self-Esteem Scale* en 10 items (Rosenberg, 1965), les échelles «Ouverture aux actions » et « Ouverture aux idées » du NEO-PI-R comprenant deux fois huit items (Costa &McCrae, 1998), la *Self-Efficacy Scale* de Sherer et al. (1982) de 21 items. Certaines échelles ont été créées pour les besoins spécifiques de l'étude, comme l'échelle de contrôle perçu par rapport au chômage en 18 items. Cette dernière échelle, inspirée des travaux de Levenson (1973), mesure si le demandeur d'emploi croit que sa situation de chômage est influencée par ses propres actions (contrôle I pour interne), par la chance (contrôle C) ou par l'action de personnages tout-puissants (contrôle P).

Bien entendu, au fur et à mesure des prises de données, le nombre de personnes faisant initialement partie de l'échantillon diminuera et au bout d'un an ne resteront que les chômeurs de longue durée. A la fin du suivi, et certainement déjà au bout d'un an, on déterminera quelle part d'explication statistique chacune des variables testées a eu pour prédire le chômage de longue durée. A partir de ces résultats, on pourra alors créer un outil qui ne reprendrait que les variables qui se sont révélées pertinentes, en les pondérant en fonction du poids qu'elles auront eu dans les résultats trouvés. Cet outil pourra ensuite être utilisé dans le repérage précoce des chômeurs à risque.

Un outil de profilage adapté

Les profilages traditionnels se basent sur les données qui se trouvent dans les dossiers individuels des chômeurs, celles-ci sont donc facilement traitables. Pour recueillir des informations de type personnalité ou comportements, une méthode et un outil spécifiques de recueil de ces données doivent être employés.

Certaines caractéristiques indispensables de l'outil peuvent être citées. Il devait être utilisable à grande échelle (plusieurs centaines de personnes), pouvoir comporter un grand nombre de questions, permettre la standardisation des questions (afin d'éviter les biais d'interviews bien connus) et des réponses (afin de permettre un traitement statistique). Par ailleurs, il était important de pouvoir disposer d'un outil de prise de données qui pouvait être utilisé avec un public souvent peu à l'aise avec l'informatique et avec l'écrit. En effet, environ 50% des personnes inscrites à l'ADEM n'ont terminé que leur scolarité obligatoire (ADEM, 2006). L'outil devait également, vu le multilinguisme de la population active luxembourgeoise, être disponible en français, allemand et portugais. La passation devait être répétable dans le temps avec une grande flexibilité au niveau des questions posées (possibilité de retrancher ou d'ajouter des items). Différents formats d'items devaient être possible (texte, image). Et finalement, la passation devait être anonyme, avec néanmoins la possibilité de retrouver des informations pertinentes dans les dossiers de l'ADEM.

On sait que les procédures d'évaluation comme les entretiens psychologiques ou les bilans de compétences en face à face sont souvent trop lourds à employer avec ce genre de population. Notre type d'outil permet aussi des évaluations plus objectives, dégagées des partis-pris subjectifs sur les personnes. Ces biais introduits par la subjectivité de l'évaluateur ont d'ailleurs été bien étudiés, il s'agit surtout des stéréotypes, de l'effet de similitude (sympathie plus grande pour des personnes qui nous ressemblent), de l'effet de halo (les indices particulièrement saillants ou subjectivement importants pour l'évaluateur prennent un poids démesuré dans l'évaluation), le souci de cohérence dans le jugement (qui amène à gommer des facteurs non congruents avec l'impression d'ensemble), etc. (pour une revue, voir Dawes, 1994). Par ailleurs, les entretiens et les bilans de compétences traditionnels ne présentent pas toujours des caractéristiques psychométriques assez valides.

Un outil a été mis au point sous une forme informatique conviviale. Les variables ont été opérationnalisées sous forme de questionnaires. Les questions ont été traduites afin d'avoir une version française, anglaise et portugaise. Ensuite, elles ont été informatisées à l'aide du logiciel Quest. Le programme résultant de ce travail permet des passations dans les locaux de l'ADEM. Nous avons eu recours à des ordinateurs portables de type "tablette PC", où les personnes n'ont besoin que de cliquer avec un stylet sur l'écran tactile pour fournir leurs réponses, pour la majorité des passations. Cet instrument ne nécessite aucune compétence informatique de la part des utilisateurs et on a montré dans diverses recherches que ce genre d'outil provoque moins d'anxiété informatique que les ordinateurs clavier-souris (Tseng, Tiplady, Macleod & Wright, 1998). Néanmoins, certains chômeurs plus à l'aise avec les outils informatiques ont répondu aux questionnaires sur des portables classiques (avec une souris au lieu du stylet). Cette précaution a été prise afin de pas "infantiliser" des personnes se sentant parfaitement à l'aise avec cet outil et rendant, de ce fait, les observations de ceux-ci plus fiables. Avec les deux types de dispositifs, les données recueillies ont été sauvegardées sur disques durs contrairement aux prévisions

initiales du projet qui voulait tenter un recueil on-line via Internet des informations. L'avantage de cette procédure est qu'elle nous met à l'abri des problèmes de connexion et de transfert des données, malheureusement encore fréquents avec Internet.

Premiers résultats et validité de construit de l'outil

L'instrument a été testé sur une cohorte de 385 demandeurs d'emploi nouvellement inscrits et n'ayant donc aucun passé de chômeurs. De fait, on a pu constater que peu de demandeurs d'emploi ont dû être éliminés de la passation (surtout des personnes illettrées et/ou incapables de comprendre les consignes). On a également constaté que très peu de personnes ont refusé, alors que la passation n'était pas obligatoire.

Le suivi longitudinal de ces personnes continuera jusqu'à deux ans d'inscription dans les SPE. Les premiers résultats au temps initial sont actuellement disponibles (Meyers & Houssemand, 2006) et permettent notamment de vérifier la validité de construit de l'outil. Celle-ci est particulièrement importante, car elle permet d'évaluer la correspondance entre la théorie et les données empiriques telles que recueillies par l'instrument (Cronbach & Meehl, 1955; Hogan & Nicholson, 1988). Elle peut se faire individuellement pour les différentes échelles utilisées, mais peut aussi se faire globalement par rapport à l'ensemble des données.

Pour les échelles individuelles, on peut citer l'exemple du *General Health Questionnaire* en 12 items que nous avons utilisé, (il existe une version en 28 et une en 30 items), en abrégé GHQ-12 (Goldberg, 1992). Cette échelle mesure l'état général de mal-être perçu, est largement utilisée dans les études épidémiologiques et a déjà été utilisée de façon répétée avec des chômeurs (Avery et al., 1998; Banks et al., 1980; Banks & Jackson, 1982; Cassidy, 2001; Claussen, 1999; Creed & Machin, 2003; Viinamaki, Niskanen & Koskela, 1995). On peut citer comme exemples d'items : "Ces derniers temps, avez-vous été capable de vous concentrer sur tout ce que vous faites?" ou "Ces derniers temps, avez-vous manqué de sommeil à cause de vos soucis?".

Différentes méthodes de cotation des réponses existent. Dans la version originale, Goldberg (1992) propose une cotation bi-modale (0-0-1-1), mais une cotation en échelle de Likert (0-1-2-3) est également souvent utilisée (Campbell, Walker, & Farrell, 2003). C'est cette dernière que nous avons adoptée, car elle permet des discriminations plus fines. L'échelle totale est donc notée de 0 à 36, des scores plus élevés indiquant un degré plus élevé de mal-être perçu.

Nos données ont montré une consistance interne mesurée par l'alpha de Cronbach de 0,86, ce qui est conforme à ce qui est généralement cité dans la littérature (0,82 à 0,91 selon les études et les échantillons). La moyenne était de 13,0 pour les femmes et 13,9 pour les hommes. Dans les études publiées, on cite des chiffres similaires (10,5 à 12,0 pour les femmes; 9,9 à 10,7 pour les hommes, selon les études, voir Pevalin, 2000; Shevlin & Adamson, 2005).

L'écart-type était de 6,5 pour les femmes et de 6,6 pour les hommes (3,9 à 5,9 pour les femmes; 4,3 à 5,4 pour les hommes, selon les études, voir également Pevalin, 2000; Shevlin & Adamson, 2005).

L'analyse factorielle permet d'approfondir la validité de construit. La littérature a montré jusqu'à présent soit l'existence d'un facteur (Goldberg, 1992), de deux facteurs (Politi, Piccinelli & Wilkinson, 1994) ou de trois facteurs (Shevlin & Adamson,

2005). Nos données rentrent le mieux dans une structure à trois facteurs similaire à celle mise en évidence par Politi et al. (1994) : à savoir "Dysphorie", "Perte de confiance" et "Dysfonctionnement".
Concernant la validité de construit plus globale de l'outil, on peut examiner les corrélations des différentes échelles entre elles et vérifier si celles-ci correspondent bien à la théorie définitoire de ces échelles. C'est ce que nous avons fait pour un certain nombre de variables données à titre d'exemples dans le tableau 2. Seules les corrélations très significatives, c'est-à-dire à p<0,001, ont été reprises.

Tableau 2 - Corrélations significatives (p<0,001) entre échelles

	0GHQ12	**Estime de soi**	**Ouverture aux idées**	**Contrôle I**	**Auto-efficacité**	**Extraversion**	**Centralité rôle professionnel**	**Anxiété sociale**	**Stress perçu**	**Dépression**	**Conscience**	**Compétences sociales**	**Soutien social**
1GHQ12		-0,47			-0,26	-0,23		0,20	0,56	0,59			
Estime de soi			0,18		0,58	0,36		-0,32	-0,62	-0,66	0,47	0,25	0,19
Ouverture aux idées					0,21	0,22		-0,30				0,36	0,26
Contrôle I						0,18							
Auto-efficacité						0,49	0,39	-0,36	-0,57	-0,55	0,64	0,36	0,23
Extraversion							0,23	-0,23	-0,40	-0,40	0,41	0,42	0,31
Centralité rôle professionnel											0,45	0,22	
Anxiété sociale									0,36	0,42	-0,23	-0,25	-0,22
Stress perçu										0,73	-0,46	-0,30	-0,31
Dépression											-0,39	-0,22	-0,30
Conscience												0,39	
Compétences sociales													0,26
Soutien social													

Un examen attentif de ce tableau montre que les variables sont effectivement corrélées en accord avec les attentes théoriques et les résultats déjà publiés dans la littérature. Ainsi, pour revenir au GHQ12, on voit que les scores sont corrélés

positivement avec l'anxiété sociale, le stress perçu et la dépression, tandis qu'on a une corrélation négative avec l'estime de soi, l'auto-efficacité et l'extraversion.

Ainsi, notre analyse structurale a permis de distinguer deux groupes de variables (voir tableau ci-après). On a d'un côté celles dont on peut faire l'hypothèse qu'elles sont adaptatives et pro-actives pour l'employabilité : estime de soi, ouverture aux idées, contrôle interne, auto-efficacité, évaluation centrale de soi, projection dans l'avenir, extraversion, centralité du rôle professionnel, conscience, compétences sociales, soutien social, satisfaction avec le soutien social, intelligence. Et on a d'autre part celles qui seraient plutôt inadaptatives et invalidantes pour l'employabilité : mal-être perçu, contrôle chance, contrôle personnages tout-puissants, anxiété sociale, stress perçu, dépression.

Tableau 3 - Variables pro-actives (1e colonne) et variables invalidantes (2e) chez les demandeurs d'emploi

estime de soi	
ouverture aux idées	
contrôle interne	
auto-efficacité	
évaluation centrale de soi	état de mal-être perçu
projection dans l'avenir	contrôle chance
extraversion	contrôle personnages tout-puissants
centralité du rôle professionnel	anxiété sociale
conscience	stress perçu
compétences sociales	dépression
soutien social	
satisfaction avec le soutien social	
intelligence	

Les corrélations sont positives à l'intérieur de chaque groupe et négatives entre les deux groupes, vérifiant la validité convergente et divergente des concepts mesurés dans l'étude. Evidemment, les mesures ultérieures permettront de vérifier dans quelle mesure ces variables sont effectivement prédictives de la durée du chômage et quel est le degré exact de ce pouvoir prédictif.

Conclusions et perspectives

Une évaluation détaillée du profil et des compétences d'un public peu qualifié, comme c'est le cas pour une majorité de demandeurs d'emploi, pose des défis et des difficultés. Afin d'avoir des données extensives et standardisées, d'éviter les différents biais d'interviews et d'évaluation subjective, nous avons mis au point un outil informatique convivial qui a été testé sur 385 demandeurs d'emploi nouvellement inscrits à l'ADEM. L'outil a montré une facilité d'utilisation et une bonne acceptation de la part des demandeurs d'emploi. Plus fondamentalement, la validité de construit de l'outil s'est avérée bonne, montrant une bonne adéquation avec la théorie définitoire des items et des échelles.

Des passations abrégées auront lieu au bout de 6, 12, 18 et 24 mois d'inscription, avec les personnes qui seront alors encore inscrites comme demandeurs d'emploi. Les deux dernières passations ne concerneront plus qu'un nombre plus limité de personnes. On repassera un certain nombre de questions qui seront choisies pour

pouvoir vérifier si l'état psychologique et les démarches de recherche d'emploi sont affectés en retour par le chômage.

Nous pourrons alors reprendre l'analyse avec la principale variable dépendante, qui est la durée de chômage pour chacune des personnes interrogées. Des conclusions plus approfondies pourront alors être tirées du matériel recueilli. A ce moment, il sera possible de déterminer dans quelle mesure les variables que nous avons mesurées sont prédictives par rapport à cette durée. On pourra alors créer des typologies de chômeurs selon leur risque de rester plus ou moins longtemps au chômage. Par ailleurs, il sera possible de créer un outil de profilage nettement plus court que le questionnaire actuel et permettant une prédiction fiable de la durée du chômage pour les personnes au moment de leur inscription.

Nous pourrons alors également déterminer dans quelle mesure le comportement et l'état psychique des personnes sont influencés par 6 mois, respectivement 12, 18 et 24 mois de chômage.

Références

ADEM (2006). *Rapport d'activité 2005*. Luxembourg : Administration de l'Emploi.

Avery, A. J., Betts, D. S., Whittington, A., Heron, T. B., Wilson, S. H., & Reeves, J. P. (1998). The mental and physical health of miners following the 1992 national pit closure programme: a cross sectional survey using General Health Questionnaire GHQ-12 and Short Form SF-36. *Public Health*, *112*, 169-173.

Banks, M. H., Clegg, C. W., Jackson., P. R., Kemp, N. J., Stafford, E. M., & Wall, T. D. (1980). The use of the general health questionnaire as an indicator of mental health in occupational studies. *Journal of Occupational Psychology*, *53*, 187-194.

Banks, M. H., & Jackson, P. R. (1982). Unemployment and the risk of minor psychiatric disorders in young people: cross-sectional and longitudinal evidence. *Psychological Medecine*, *12*, 789-798.

Campbell, A., Walker, J., & Farrell, G. (2003). Confirmatory factor analysis of the GHQ-12: can I see that again? *Australian and New Zealand Journal of Psychiatry*, *37*, 475–483.

Cassidy, T. (2001). Self-categorization, coping and psychological health among unemployed mid-career executives. *Counselling Psychology Quarterly*, *14*, 303-315.

Claussen, B. (1999). Health and re-employment in a five-year follow-up of long-term unemployed. *Scandinavian Journal of Public Health*, *2*, 94-100.

Creed, A. P., & Machin, M. A. (2003). Multidimensional properties of the Access to Categories of Experience scale. *European Journal of Psychological Assessment*, *19*, 85-91.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281-302.

Cortese, L., & Meyers, R. (2004). *Discussion des variables pouvant intervenir dans le chômage de longue durée au Luxembourg*. Working paper n° 1. Luxembourg : Université du Luxembourg.

Costa, P. T., & McCrae, R. R. (1998). *NEO PI-R Inventaire de personnalité-Révisé - Manuel*. Adaptation française de J.-P. Rolland. Paris : Les Editions du Centre de Psychologie Appliquée.

Dawes, R. M. (1994). *House of cards - Psychology and psychotherapy built on myth*. New York : Free Press.

Goldberg, D. (1992). *General Health Questionnaire (GHQ-12)*. Windsor, UK : NFER-Nelson.

Hogan, R., & Nicholson, R. A. (1988). The meaning of personality test scores. *American Psychologist*, *43*, 621-626.

Houssemand, C., Hawiltzky, S., Kieffer, G., Lanners, Y., Leijten, T., & Reiter, R. (1997). *Enquête sur le chômage de longue durée – Situation actuelle au Luxembourg*. Luxembourg : ADEM-CEPS.

Kanfer, R., Wanberg, C. R., & Kantrowitz, T. M. (2001). Job search and employment: A personality-motivational analysis and meta-analytic review. *Journal of Applied Psychology*, *86*, 837-855.

Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology*, *41*, 397-404.

Meyers, R., & Houssemand, C. (2006). *Résultats de la première passation du questionnaire informatisé auprès d'une cohorte de demandeurs d'emploi au Luxembourg*. Working paper n° 4. Luxembourg : Université du Luxembourg.

Pevalin, D. J. (2000). *Investigating long-term retest effects in the GHQ-12*. Colchester, UK: Institute for Social and Economic Research.

Politi, P. L., Piccinelli, M., & Wilkinson, G. (1994). Reliability, validity and factor structure of the 12-item General Health Questionnaire among young males in Italy. *Acta Psychiatrica Scandinavica*, *90*, 432–437.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, New Jersey : Princeton University Press.

Rosholm, M., Svarer, M., & Hammer, B. (2004). *A Danish Profiling System*. IZA Discussion paper N° 1418. Bonn : IZA.

Sherer, M., Maddux, J., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. (1982). The Self-Efficacy Scale: Construction and validation. *Psychological Reports*, *51*, 663-671.

Shevlin, M., & Adamson, G. (2005). Alternative factor models and factorial invariance of the GHQ-12: A large sample analysis using confirmatory factor analysis. *Psychological Assessment*, *17*, 231–236.

Tseng, H. M., Tiplady, B., Macleod, H. A., & Wright, P. (1998). Computer anxiety: A comparison of pen-based personal digital assistants, conventional computers, and paper assessment of mood and performance. *British Journal of Psychology*, *89*, 599-610.

Viinamaki, H., Niskanen, L., & Koskela, K. (1995). General health questionnaire and Beck depression scale as screening methods for psychiatric morbidity among the unemployed. *European Journal of Psychiatry*, *9*, 209-215.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

EVALUATION D'UN LOGICIEL D'(AUTO)APPRENTISSAGE DES LANGUES

Audrey Thonard, Université de Liège

Mots-clés : logiciel, (auto) apprentissage, langues

Résumé :

Soucieuse du développement de son internationalisation, l'Université de Liège a émis le souhait de mettre en place un nouveau programme visant l'amélioration du niveau de connaissances en langues de l'ensemble de ses membres. Ce programme est centré sur l'enseignement à distance, avec l'encadrement d'un tuteur qui guiderait ses apprenants et qui leur proposerait une série d'activités adaptées à leurs besoins propres.

En 2006, l'Université de Liège a invité deux cents volontaires à utiliser le logiciel « Tell me more » (TMM dans la suite de l'article) pendant quatre mois et a ensuite recueilli leurs impressions. Deux formules d'encadrement ont été testées sur des apprenants étudiant diverses langues et relevant de niveaux différents (du niveau A1 au niveau B2). Des experts de l'enseignement des langues et de l'enseignement à distance ont étudié l'outil et l'ont analysé.

Si les volontaires ont globalement bien accueilli le logiciel et lui ont trouvé un grand nombre de qualités, ils ne l'ont pas utilisé aussi assidûment qu'ils l'avaient prévu. Les experts se sont penchés sur les raisons, internes et externes au logiciel, qui ont provoqué la démotivation des participants. Il ressort de leurs analyses que les principaux défauts du logiciel sont la faible variété de paradigmes proposés, la faible diversité des thèmes des séquences, le manque de gradation de la difficulté, le manque d'intérêt de certaines activités et l'absence d'activités visant le développement de l'expression. La démotivation des participants s'explique également par le type d'encadrement qui leur a été proposé et par la durée de l'apprentissage.

L'expérience montre que le logiciel constitue un outil agréable et intéressant mais qu'il ne peut permettre un apprentissage complet. La présence d'un professeur, l'ajout d'activités adaptées et l'établissement d'un échéancier s'avèrent être des éléments indispensables à la réussite d'un système de formation centrée sur ce type de logiciel.

Introduction

Partant du fait que la majorité des membres de l'Université de Liège ont besoin d'une ou de plusieurs langues étrangères dans leur vie professionnelle ou dans leurs études et que la plupart d'entre eux rencontrent des difficultés à suivre assidûment les cours traditionnels, l'ULg souhaite mettre à la disposition de tous une solution alternative, à savoir le recours à un logiciel d'(auto)apprentissage en ligne, permettant à l'apprenant une meilleure flexibilité de son emploi du temps. Son choix s'est porté sur le logiciel « Tell me more on line », qui lui paraissait être le plus complet.

1. Déroulement de l'expérience

En novembre 2005, les membres de l'ULg ont été invités à compléter une enquête en ligne, dont l'intérêt était double puisqu'elle visait non seulement à sonder la communauté universitaire à propos d'un tel outil mais aussi à recruter des participants pour l'évaluation du logiciel choisi. En janvier 2006, deux cents volontaires ont été sélectionnés sur base de leur langue maternelle, de la langue qu'ils souhaitaient apprendre, de leur niveau et, enfin, de leur statut au sein de l'ULg. Huit groupes ont été constitués autour de quatre langues (le français, l'anglais, le néerlandais et l'espagnol), de trois types d'encadrement (*e-coaching*, *e-tutoring* et *e-learning*) et de quatre niveaux (de A1 à B2). Ils étaient composés d'étudiants et/ou de membres du personnel. Un groupe de contrôle a été constitué pour le français. Chacun de ces groupes comportait, *grosso modo*, un tiers d'apprenants de niveau A1-A2, un tiers d'apprenants de niveau B1, un tiers d'apprenants de niveau B2.

Tous les participants se sont engagés à utiliser cet outil *on line* deux heures par semaine, pendant quatre mois (de février à mai). Pour les accompagner, sept professeurs de langue ont joué le rôle de tuteur à distance et deux informaticiens ont assuré la maintenance du serveur et ont répondu aux questions des différents acteurs de l'expérience. A deux reprises, tous les participants de l'expérience ont été invités à communiquer leurs impressions générales et leurs commentaires sur les aspects pédagogique, informatique et pratique du logiciel.

Il est important de signaler que, en raison de sérieux problèmes d'ordre logistique, un tiers des étudiants francophones et la majorité des étudiants non-francophones n'ont pu accéder au logiciel et se sont vus retirés de l'expérience. Par conséquent, ni l'enseignement du français, ni la formule d'*e-coaching* n'ont pu être testés au cours des quatre mois de l'expérience. Par ailleurs, aucun groupe de contrôle n'a pu valider les qualités et les lacunes du logiciel. Mais si les difficultés rencontrées au cours de cette phase expérimentale n'ont pas permis d'étude complète et précise de l'impact du logiciel sur l'apprentissage des langues, elles n'ont pas empêché les participants ainsi que des experts de l'enseignement des langues et de l'enseignement à distance de se plonger dans l'outil et d'analyser sa forme et son contenu.

2. Accueil du logiciel par les volontaires

Impressions et commentaires

Après un mois, les impressions générales des participants sont bonnes. Ils qualifient ce logiciel de convivial, d'amusant, de stimulant et de motivant. Ils apprécient la liberté d'action qu'il autorise et la facilité de manipulation. La cinquantaine de participants ayant complété la seconde enquête de satisfaction affirment que leur motivation initiale est restée intacte, voire plus grande, à de rares exceptions près. Cette motivation s'explique surtout par l'aspect ludique du logiciel et par sa malléabilité aux exigences de l'utilisateur et à ses compétences et, bien sûr, par le besoin de maîtriser la langue étudiée dans la vie professionnelle.

Cependant, la majorité des participants regrettent la faible variété de paradigmes proposés, la faible diversité des thèmes des séquences, le peu d'activités visant au développement de l'expression orale et écrite, le manque de précision des explications grammaticales, le manque de clarté de certaines consignes, le manque de gradation de la difficulté, la rigidité de la correction de l'expression écrite, la longueur de certains exercices et, enfin, le manque d'intérêt de certaines activités.

Evaluation du taux de participation

Bien que chaque volontaire se soit engagé, par écrit, à consacrer en moyenne deux heures par semaine à l'apprentissage de la langue étrangère choisie pendant toute la durée de l'expérience, peu ont réellement honoré le contrat. Au sein des six groupes restants :

- 16 % des volontaires n'ont pas du tout utilisé le logiciel ;
- 38 % d'entre eux ont utilisé le logiciel moins de deux heures par semaine ;
- 47 % d'entre eux ont utilisé le logiciel au moins deux heures par semaine.

Utilisation du logiciel	Nulle	Moins de deux heures par semaine	Au moins deux heures par semaine
Académique	10%	20%	70%
Scientifique	8%	31%	61%
PATO	20%	37%	43%
Etudiants francophones	18%	47%	35%
Moyenne des groupes	**16%**	**38%**	**47%**

Tableau 1 : Statistiques des données issues de l' « Outil de l'Administrateur » de TMM

Le tableau 1 le montre, plus de la moitié des participants ont rapidement été découragés par l'apprentissage. Pourtant, les volontaires qui ont interrompu leur formation n'ont jamais remis en cause le logiciel lui-même.

Cette faible participation semble s'expliquer par le manque de temps dû, soit à une charge plus élevée de travail, soit à des problèmes d'ordre personnel : les volontaires ont rapidement pris conscience que le temps consacré à l'étude de la langue dépassait largement celui voué à l'utilisation du logiciel. Plusieurs participants ont estimé que la durée de l'expérience et le moment où elle s'est déroulée ont été deux facteurs plutôt négatifs.

L'absence d'un encadrement adéquat semble avoir participé, lui aussi, à la démotivation des apprenants. C'est parmi les participants en auto-apprentissage que l'on dénombre le plus grand taux d'abandon, principalement au sein des débutants puisque deux tiers d'entre eux ont travaillé moins d'une heure par semaine alors que seule la moitié des débutants accompagnés dans leur apprentissage ont travaillé à ce rythme. Les participants accompagnés, eux, sont presque unanimes quant à la qualité, à l'intérêt et à la nécessité de ce contact. Ils ont cependant regretté l'absence de rencontres réelles avec leur professeur et de contacts avec leurs condisciples, indispensables à la création d'une dynamique de groupe.

Evaluation du travail accompli

En accédant à son compte, l'apprenant peut visualiser le parcours d'apprentissage de manière synoptique et être informé sur les compétences visées par chaque activité au moyen d'un tableau de bord. Ce tableau participe à la dynamique motivationnelle puisqu'il permet à l'apprenant d'avoir une vue d'ensemble du travail accompli et de celui qu'il doit encore effectuer. Il permet également au tuteur de vérifier la masse de travail effectuée par ses apprenants et son assiduité. Dans le cadre de l'expérience, ce contrôle n'a pu être effectué que partiellement car le temps imparti au travail des tuteurs s'est révélé insuffisant. Les messages d'encouragement et d'alerte envoyés n'ont eu que peu d'effet puisque de nombreux participants, pour qui le logiciel était facilement accessible, ne l'ont pas utilisé une seule fois.

Evaluation des progrès accomplis

Avant et après l'expérience, les compétences linguistiques des deux cents volontaires ont été appréciées, d'une part, par les membres de l'Institut Supérieur des Langues Vivantes de l'Université de Liège (pour l'expression écrite et orale) et, d'autre part, par le logiciel lui-même (pour les autres compétences). Les participants non-francophones avaient présenté, en outre, le *Test d'Evaluation du Français* de la Chambre de Commerce et d'Industrie de Paris.

Si tous les volontaires avaient participé au pré-test, seuls 93 se sont présentés au post-test organisé par les professeurs de l'ISLV, malgré l'engagement écrit et les invitations répétées. La post-évaluation effectuée par les tuteurs montre que ces apprenants ont, en moyenne, amélioré leur score d'1,7 point sur 20, en expression orale et en expression écrite. Seule une petite trentaine d'entre eux ont effectué le test de progression proposé par le logiciel. Ce test de progression indique que les

participants ont évolué d'un point sur 20, en moyenne. Pour rappel, aucun post-test n'a été réalisé avec le public non-francophone puisque ce dernier n'a pu accéder au logiciel.

En raison du nombre réduit de volontaires effectifs, ces informations sont à prendre avec précaution. La courte durée de l'apprentissage et le faible nombre de participants ayant respecté leur engagement jusqu'au terme de l'expérience rendent donc peu fiables les résultats en matière d'efficacité du logiciel.

3. Analyses des experts

Analyse pédagogique

Les professeurs de langue ont émis les critiques suivantes :

- Si l'on se réfère aux travaux de Leclercq et Poumay (2005) de l'ULg qui ont identifié huit événements d'apprentissage, décrits dans une structure identifiant la posture de l'apprenant et celle de l'accompagnateur, l'on peut dire que, malgré leur variété, les activités proposées par TMM se situent essentiellement dans un paradigme d' « Exercication / Guidage » ou dans celui de l' « Imitation / Modélisation ». Les paradigmes tels que « Débat / Modération » ou « Création / Encouragement » en sont complètement absents. TMM néglige donc la production et la communication, qui constituent, de nos jours, le stade ultime de l'apprentissage des langues.
- Les explications grammaticales sont correctes mais ne sont pas toujours claires et efficaces. Elles sont simples et suffisantes pour des non-spécialistes mais sont peu orientées vers la production. Leur ordre d'apparition manque parfois de cohérence et la notion de grammaire contrastive est tout à fait étrangère au logiciel.
- Le logiciel aide peu l'apprenant à mémoriser le vocabulaire qui n'est ni introduit, ni exploité.
- Les fiches culturelles, constituées d'une image et d'un court texte, sont intéressantes et bien conçues, même si elles recouvrent parfois certains clichés. Il est cependant regrettable que leur contenu n'ait peu ou aucun lien avec les autres activités et que rien n'ait été prévu pour l'exploitation du contenu ou la vérification de la compréhension.
- Les tests proposés se composent d'une série de questions à choix multiples, alternant compréhension orale et écrite. L'apprenant est informé explicitement sur quoi porte l'évaluation, quand elle est prévue et par qui elle est opérée. Cependant, si la formation proposée et l'évaluation mise en place concordent au niveau de la méthodologie, elles ne s'accordent pas toujours au niveau des contenus, ce qui explique parfois une régression dans les résultats obtenus au test de *progression*.

Analyse interpersonnelle

La seule possibilité d'échanges se résume à l'outil de communication interne entre l'apprenant et son éventuel tuteur, ce qui prive l'apprenant des nombreuses possibilités offertes en matière de communication synchrone et asynchrone, qui sont généralement des fonctions de base de toute plate-forme *e-learning*. Tous les utilisateurs de TMM *on line* ont accès à un forum, « Le Club Auralog », qui n'a pas été exploité dans le cadre de l'expérience. La création d'un forum propre à chaque groupe permettrait de communiquer avec des apprenants partageant un profil semblable, et d'effectuer d'autres types d'activités que celles proposées par le logiciel, ce qui contribuerait à établir un réel sentiment d'appartenance à un groupe d'apprentissage.

L'expérience montre que la présence d'un professeur s'avère indispensable pour traiter les erreurs, suggérer des pistes d'analyse et de remédiation mais aussi pour encourager et stimuler les apprenants et leur proposer des tâches qu'ils ne peuvent effectuer avec TMM.

Analyse motivationnelle

Comme tout système de formation, l'enseignement à distance doit veiller à maintenir l'envie d'apprendre et provoquer l'engagement des apprenants dans une démarche d'apprentissage. Il doit, lui aussi, leur permettre de prendre conscience de leurs incompétences dans le domaine étudié et leur proposer les outils pour parvenir à résoudre leurs problèmes. Selon Viau (2006), pour qu'une activité pédagogique suscite et maintienne la motivation des apprenants, il faut notamment :

- qu'elle soit **signifiante** à leurs yeux c'est-à-dire qu'elle « correspond[e] à ses intérêts, s'harmonise avec ses projets personnels et répond[e] à ses préoccupations » (Viau, 2006). Dans le cadre de TMM, les objectifs généraux visés sont disponibles dans l'interface, au travers d'un document générique qui les définit par niveau mais qui n'est pas adapté à chaque langue. De ce fait, l'apprenant peut difficilement avoir conscience de l'objectif général qu'il est censé atteindre au long de son cursus et se voit donc engagé dans un travail dont il ne connaît pas explicitement les finalités. D'autre part, les thématiques abordées dans le logiciel ne sont pas adaptées à toutes les disciplines étudiées ou à toutes les professions exercées à l'Université. Cette faiblesse peut être palliée par le recours à divers outils (internes ou non au logiciel) permettant au tuteur d'ajouter ses propres leçons, construites autour d'une thématique particulière, et de mieux répondre de la sorte aux besoins de ses apprenants.
- qu'elle soit **diversifiée** et qu'elle s'intègre aux autres activités : si les activités proposées par TMM sont assez variées, elles se révèlent assez répétitives et lassantes au terme de plusieurs semaines d'apprentissage. Beaucoup de tâches à accomplir se ressemblent puisque la plupart du temps, elles mènent l'apprenant à reconnaître ou à choisir la « bonne » réponse.
- qu'elle représente un **défi** pour l'apprenant : si la majorité des apprenants s'accordent pour affirmer que le niveau de difficulté est adéquat, les experts, quant

à eux, estiment que le niveau est trop faible par rapport au niveau annoncé. Quoi qu'il en soit, à tout moment de sa formation, l'apprenant a la possibilité de descendre ou de monter de niveau, ou bien de sélectionner les compétences qu'il souhaite surtout développer.

- qu'elle soit **authentique**, c'est-à-dire qu'elle mène à « une réalisation, c'est-à-dire à un produit qui ressemble à ceux que l'étudiant trouvera dans sa […] vie professionnelle » (Viau, 2006). Dans son état actuel, le logiciel ne peut satisfaire à cette exigence puisqu'il n'aborde qu'un nombre limité de sujets et n'invite pas réellement son utilisateur à « produire ». La fabrication d'activités et de parcours adaptés à chaque public paraît donc, une nouvelle fois, indispensable.
- qu'elle exige un **engagement cognitif** de la part de l'apprenant : les participants de la phase expérimentale ont pris conscience très rapidement que, pour approfondir la connaissance de la langue cible, ils allaient devoir dépasser le logiciel et avoir recours à d'autres supports de travail et à d'autres moyens de mémorisation. Le logiciel mène l'apprenant à établir des liens entre les nouvelles notions et celles déjà apprises mais uniquement au sein d'une même séquence.
- qu'elle **responsabilise** l'étudiant en lui permettant de faire des choix : cette condition est tout à fait respectée par TMM. La liberté d'action et la facilité d'accès sont désignées, par les participants et par la plupart des experts, comme deux des plus grandes qualités du logiciel. Selon Viau, « [i]l revient toutefois à l'enseignant de décider quels éléments de l'enseignement et de l'apprentissage demeureront sous sa responsabilité et quels seront ceux dont il pourra déléguer la responsabilité à l'étudiant. » (Viau, 2006). Un excès de liberté peut en effet nuire à l'apprenant qui peut accéder directement à la solution de l'exercice ou à l'activité suivante.
- qu'elle permette à l'apprenant **d'interagir et de collaborer avec les autres** : l'apprentissage coopératif est négligé par le logiciel. Certes, il propose l'accès à un forum mais ce dernier est ouvert à l'ensemble des utilisateurs du logiciel et ne peut donc créer une réelle atmosphère de coopération.
- qu'elle comporte des **consignes claires** : en général, les consignes des activités proposées par TMM le sont. Certaines consignes techniques pour la réalisation d'exercices d'appariement, par exemple, pourraient l'être davantage mais, dans l'ensemble, l'utilisateur manipule le logiciel de façon intuitive. Pour anticiper plus efficacement les difficultés auxquelles les futurs utilisateurs pourraient être confrontés, la prise de contact entre tuteur et apprenants devrait se dérouler dans une salle multimédia où les futurs utilisateurs pourraient accéder au logiciel et poser des questions. Le tuteur pourrait également remettre, en début d'expérience, un document reprenant les réponses aux questions les plus fréquemment posées, les ressources disponibles et les démarches techniques à accomplir. Ce guide devrait également exposer les objectifs généraux et spécifiques visés, présenter la méthodologie et les fondements didactiques sous-tendus par TMM et par le système de formation mis en place. Il pourrait également comporter un calendrier indiquant les échéances à respecter.
- qu'elle se déroule pendant une **période de temps suffisante** : dans le cadre de cette expérience, les participants n'étaient pas tenus de réaliser une tâche précise, ni d'atteindre des objectifs précis. Les notions de temps et d'échéances étaient donc absentes du projet. Beaucoup de participants ont le sentiment que l'utilisation

du logiciel sur une période plus longue aurait permis de rattraper les retards accumulés pendant les semaines de surcharge de travail, ce qui aurait été tout bénéfice.

Analyse informatique et multimédia

Selon le LabSET (*Laboratoire de Soutien à l'Enseignement Télématique*) de l'ULg, « l'utilisation du multimédia est très variée et elle est au service de l'apprentissage. Sur ce point, les développeurs de TMM ont réalisé un travail très soigné. Tous les tests et les exercices proposent des feed-back automatiques, la vidéo, le son sont très utilisés[…]» (LabSET, 2006) Cependant, « [l]e principe de reconnaissance vocale pourrait être plus performant car il est parfois nécessaire de répéter une phrase plusieurs fois avant de la voir acceptée (sans que la manière de prononcer ait été changée). Ce défaut est très irritant et pourrait s'avérer démotivant pour les utilisateurs. » (LabSET, 2006). Divers tests ont été réalisés par les tuteurs de l'expérience à ce propos et permettent d'affirmer que la reconnaissance vocale, proposée par TMM :

- ne sanctionne pas toujours la prononciation de phonèmes inadéquats (voire de mots entiers inadéquats), ni l'ajout d'un phonème ou d'une syllabe,
- attribue parfois un bon score à l'apprenant qui s'exprime avec un fort accent étranger à la langue cible (accent qui peut rendre le mot incompréhensible pour un natif),
- ne sanctionne ni un phonème allongé inadéquatement, ni un accent tonique mal placé,
- sanctionne de diverses façons un mot prononcé de façons identiques.

L'intérêt des exercices de prononciation réside plutôt dans l'amélioration de la compréhension de la langue cible et dans la possibilité qu'a l'apprenant d'entendre indéfiniment les mots cibles.

Conclusion

Au bout de quatre mois d'utilisation, les participants sont satisfaits de ce logiciel et la plupart d'entre eux souhaitent continuer à l'utiliser. Ils apprécient principalement la facilité d'accès et de manipulation, qui leur permet de suivre aisément une formation en langue.

Cependant, de nombreuses activités proposées par TMM négligent des conditions *sine qua non* au maintien de la motivation des apprenants. Il conviendrait donc de créer une motivation extrinsèque au logiciel, telle que l'attribution de « crédits-temps », la remise d'un diplôme, d'un « passeport langues » en fin de cycle ou encore la perspective d'effectuer un stage ou de participer à un colloque à l'étranger.

Les diverses analyses des experts montrent clairement que le logiciel constitue un bon outil d'exercisation, qu'il développe surtout la compréhension, la justesse grammaticale et phonétique mais qu'il ne peut, à lui seul, préparer l'apprenant à « analyser, traiter et solutionner une situation inédite ou réaliser un projet » (Gérard, 2001). C'est pourquoi l'encadrement des utilisateurs et la mise en place d'activités visant au développement des compétences peu travaillées avec TMM s'avèrent indispensables.

Références

Gérard, F.-M. (2001), L'évaluation de la qualité des systèmes de formation, *Mesure et Évaluation en Éducation*, 24, n°2-3.

LabSET (2006), Evaluation du multimédia pédagogique d'apprentissage des langues *Tell me more*.

Leclercq, D. & Poumay, M. (2005), The 8 Learning Events Model and its principles, en ligne sur http://www.labset.net/media/prod/8LEM.pdf (consulté le 23 novembre 2005).

Viau, R. (2006), La motivation des étudiants à l'université : mieux comprendre pour mieux agir, en ligne sur http://www.ifres.ulg.ac.be/texte_motivation_viau.doc (consulté le 14 juin 2006).

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

EVALUER L'EFFICACITE PEDAGOGIQUE D'UNE FORMATION OU D'UN COURS A L'AIDE D'UN OUTIL D'AUTOEVALUATION

François-Marie GERARD, Directeur adjoint du BIEF
(Bureau d'Ingénierie en Education et en Formation, Louvain-la-Neuve, Belgique)

Jean-Marc BRAIBANT, Thérèse BOUVY, Conseiller en pédagogie universitaire à l'IPM (Institut de Pédagogie Universitaire et des Multimédia, Louvain-la-Neuve, Belgique)

Résumé :
Lorsqu'elle a lieu, l'évaluation d'une action de formation, que celle-ci se déroule dans le monde de l'enseignement ou dans les organisations professionnelles, est la plupart du temps une évaluation du processus. Le plus souvent, il s'agit d'un questionnaire de satisfaction qui aborde des questions telles que l'adaptation du contenu de la formation au projet personnel, la qualité de l'équilibre entre les activités théoriques et pratiques, l'ouverture sur la vie active et/ou professionnelle, les conditions de déroulement de l'action de formation (équipement, tenue des locaux, accessibilité), etc. Les réponses à ces questions sont bien sûr intéressantes, car elles permettent de réguler le processus, ce qui devrait permettre à une nouvelle formation du même type d'améliorer l'atteinte des résultats.

Néanmoins, une telle évaluation ne permet pas d'attester que l'action de formation a atteint ou non ses objectifs. L'évaluation en termes de résultats est difficile, pour plusieurs raisons. D'abord, les objectifs liés à une action de formation se situent la plupart du temps à plusieurs niveaux. Ainsi, une formation professionnelle cherche la plupart du temps :

- à développer et/ou à renforcer auprès des participants une série de compétences ;
- à inciter ceux-ci à mettre en œuvre les acquis de la formation, c'est-à-dire ces compétences, une fois qu'ils sont sur leur poste de travail ;
- à améliorer de cette sorte la qualité du travail accompli et par là les performances de l'organisation.

Une évaluation des résultats atteints par une action de formation devrait donc idéalement se situer à ces 3 niveaux : les participants ont-ils acquis les compétences visées (*évaluation des acquis*) ? Les ont-ils mises en œuvre sur leur poste de travail (*évaluation du transfert*) ? Cela a-t-il permis d'améliorer les performances de l'organisation (*évaluation de l'impact*) ?

Introduction

Le processus de formation et d'évaluation peut se formaliser de la manière suivante (GERARD, 2003) :

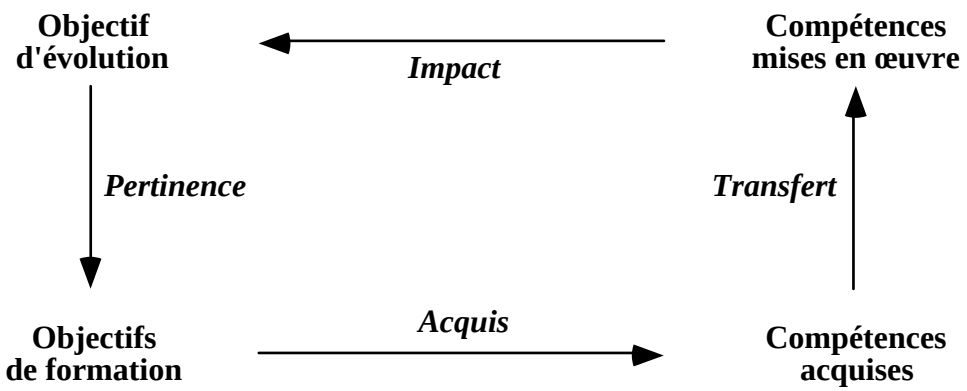


Figure 1 - Processus de formation et d'évaluation

Une action de formation est mise en place parce que l'organisation s'est définie, en fonction de ses besoins, un **objectif d'évolution** (HAUSER et al., 1985) : diminuer le nombre de pannes, accroître le chiffre d'affaires, augmenter la qualité du service à la clientèle, etc.

L'action de formation mise en place doit permettre a priori d'atteindre cet objectif d'évolution, en déterminer les **objectifs de formation** les plus adéquats ou les compétences que l'on va essayer de développer ou de faire acquérir et qui devraient permettre, en étant mises en œuvre, d'atteindre l'objectif d'évolution. La ***pertinence*** des objectifs de formation est donc primordiale : sont-ils ceux qui, a priori, permettront le mieux d'atteindre l'objectif d'évolution ?

L'action de formation consistera à mettre en œuvre un processus permettant d'atteindre les objectifs de formation et que les **compétences** soient **acquises** par les participants. L'évaluation des ***acquis*** permettra d'attester qu'il en est bien ainsi. Elle pourra se réaliser durant ou à la fin de la formation.

Lorsque l'action de formation est terminée (ou même parfois lorsqu'elle est encore en cours dans le cas d'actions étalées sur plusieurs semaines, voire plusieurs mois), une évaluation du ***transfert*** permettra de vérifier que les compétences acquises sont effectivement **mises en œuvre** sur le poste de travail.

L'**impact** d'une action de formation est dès lors directement lié à la pertinence des objectifs de formation, à l'efficacité pédagogique de la formation et à la qualité du transfert des compétences acquises.

L'outil présenté dans cette communication se contente d'aborder l'évaluation des acquis ou de l'efficacité pédagogique. Cette évaluation devrait aller de soi : quoi de plus naturel que de vérifier qu'on a bien atteint son objectif ! Dans le monde de l'enseignement, une telle évaluation est la plupart du temps présente, mais elle se réalise, surtout au niveau universitaire, sur un mode très normatif. Dans l'univers professionnel, rares sont les actions de formation qui font l'objet d'une telle évaluation. Tout se passe comme si on considérait que le simple fait de participer à la formation devait suffire à acquérir les compétences visées, en se basant sur le postulat qu'un adulte est maître de son apprentissage et suffisamment responsable pour tirer pleinement profit des stratégies didactiques qui lui sont proposées. Ce postulat, certes généreux, est malheureusement faux, au moins partiellement.

La plupart des gestionnaires de formation sont pleinement conscients de cette réalité, mais l'évaluation des acquis reste peu présente, essentiellement en raison de deux types de problème :

- d'une part, un problème méthodologique. Comment faire pour évaluer les acquis d'une formation qui vise la mise en place ou le renforcement de compétences ? Une restitution d'informations sans réellement les utiliser dans une situation proche de la situation professionnelle ne servirait pas à grand chose. Il conviendrait donc de mettre les participants face à une situation complexe, proche de la réalité. Mais de telles situations ne sont pas faciles à construire ni à mettre en œuvre, et cela prend du temps. De plus, la dynamique de formation fortement liée au travail de groupe risque de conduire à évaluer une compétence collective alors qu'elle devra être mise en œuvre individuellement sur le terrain ;
- d'autre part, un problème psychologique. Il n'est pas évident d'imposer à des adultes ce regard extérieur, qui ne rappelle que trop les sinistres évaluations scolaires.

Face à ces difficultés, un simple questionnaire peut apporter une ébauche de solution, qui – certes – ne remplace pas une véritable évaluation de la maîtrise des compétences par les participants, mais qui peut constituer un indicateur précieux pour évaluer un certain degré d'efficacité pédagogique.

L'outil que nous proposons — utilisé avec beaucoup de bénéfices dans de nombreuses actions de formation ou d'enseignement — se fonde sur l'idée que l'évaluation des compétences passe peut-être nécessairement par une autoévaluation, permettant à l'apprenant comme au travailleur, de circonscrire dans quelle mesure il est à même d'affronter de manière pertinente les situations auxquelles il peut — dans le cadre de la compétence visée — apporter une réponse satisfaisante (GERARD & VAN LINT-MUGUERZA, 2000). En effet, autant un observateur extérieur est à même d'observer et d'évaluer si tel ou tel individu maîtrise ou non tel ou tel comportement, autant il est difficile, voire impossible, à cet observateur extérieur de dire si oui ou non une compétence est réellement maîtrisée. En effet, l'exercice de la compétence

fait appel à la mobilisation intégrée d'un ensemble approprié de savoirs, de savoir-faire et de savoir-être dans une situation propre à la compétence mais qui peut être en soi inédite, imprévisible. La personne elle-même n'est-elle pas la mieux placée pour estimer si elle est à même de réaliser cette mobilisation dans toutes les situations ? Le principe de l'outil proposé est dès lors de demander au participant d'exprimer dans quelle mesure il s'estime lui-même plus ou moins compétent au début et à la fin de la formation.

Présentation de l'outil et des traitements quantitatifs

Pour construire l'outil, il convient de reprendre les différents objectifs de la formation, formulés à partir des compétences visées en formation professionnelle, et de demander au participant d'estimer pour chaque objectif son niveau de compétence avant et au terme de la formation. La validation que nous avons faite de l'outil nous a montré que ce double questionnement peut se faire en même temps, à la fin de la formation.

Nous proposons ici un exemple de questionnaire, tel qu'il a pu être élaboré dans le cadre d'une formation en gestion de projet.

Pour chaque objectif, estimez, selon vous, votre niveau de compétence.

0 = pas du tout capable
10 = tout à fait capable

	Avant la formation	Au terme de la formation
a) Planifier et gérer des projets par des objectifs	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
b) Développer des programmes d'activités et des plans de travail	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
c) Développer des plans budgétaires	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
d) Évaluer le projet	0 1 2 3 4 5 6 7 8 9 10	0 1 2 3 4 5 6 7 8 9 10
e) Construire et manager des équipes	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9

Figure 2 - Exemple de questionnaire

Cet outil permet un certain nombre de traitements quantitatifs, relativement simples et pouvant être standardisés à l'aide d'un tableur, qui permettent d'analyser de manière intéressante les résultats de la formation.

Par exemple, prenons le cas où 10 personnes ont participé à la session sur la gestion de projets. Pour l'objectif « a) Planifier et gérer des projets par des objectifs », ils ont donné les évaluations suivantes :

	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**
Avant la formation	5	5	4	6	2	5	2	3	5	4

Au terme de la formation	8	7	7	8	7	8	5	7	8	6

À partir de ces résultats, on peut dresser un bilan global des compétences, pour chaque objectif, en calculant un certain nombre d'indices :

- calcul de la moyenne (μ) avant la formation et au terme de la formation ;
- calcul de l'écart-type (σ) avant la formation et au terme de la formation ;
- calcul du taux d'hétérogénéité (ou coefficient de variation) (η). Ce taux correspond au rapport entre l'écart-type et le score moyen ($\eta = \sigma / \mu \times 100$). Son avantage par rapport à l'écart-type est qu'il donne un pourcentage qui permet de supprimer la référence à la moyenne. L'intérêt de cet indice est qu'il permet d'avoir une idée du degré d'accord entre les personnes interrogées. On considère qu'en dessous de 15%, l'accord (ou l'homogénéité) est important(e), alors qu'au-dessus de 30%, il existe un désaccord (ou une hétérogénéité) important(e) (D'HAINAUT, 1975 ; OUELLET, 1985).

Sur la base des données ci-dessus, le calcul de ces trois indices peut déboucher sur les tableaux suivants :

AVANT LA FORMATION	μ_1	σ_1	η_1
a) Planifier et gérer des projets par des objectifs	4,1	1,3	31,7%

AU TERME DE LA FORMATION	μ_2	σ_2	η_2
a) Planifier et gérer des projets par des objectifs	7,1	0,94	13,3%

Ces premiers indices apportent déjà des informations intéressantes :

- puisque la moyenne au terme de la formation est plus élevée qu'avant la formation, il y a eu une amélioration dans la maîtrise de la compétence — nous y reviendrons —, mais le niveau de maîtrise final est encore relativement faible. Un score moyen de 8 sur 10 est un meilleur indicateur de maîtrise, mais il convient bien sûr d'être prudent à cet égard (il est important de comparer les niveaux atteints pour les différentes compétences afin de mieux analyser les résultats) ;
- la comparaison des **taux d'hétérogénéité** montre que la formation a permis de réduire la disparité qui existait au départ, ce qui signifie que l'apprentissage a eu un effet d'« équité », dans la mesure où les écarts de compétences entre les participants se sont réduits et que la formation a donc contribué à un plus grand « partage » des compétences.

L'analyse de **l'effet d'« apprentissage** » peut être affinée en calculant des indices de gains entre les deux moyennes :

- le gain brut moyen correspond à ce qui a été effectivement gagné et se calcule simplement par la différence entre le score moyen « après » et le score moyen « avant » la formation ($\mu_2 - \mu_1$) ;

- le gain relatif moyen est le rapport entre ce qui a été gagné et ce qui pouvait être gagné et se calcule par la formule suivante (Score APRÈS - Score AVANT) / (Score MAXIMUM - Score AVANT) × 100, soit — selon notre exemple — par le calcul $\frac{(7{,}1-4{,}1)}{(10-4{,}1)} \times 100 = 50{,}8\%$. On peut considérer qu'il y a un effet positif d'apprentissage lorsque ce gain relatif est supérieur à 30 ou 40%[1].

Ces deux indices peuvent être synthétisés dans un nouveau tableau :

	Gain brut moyen	Gain relatif moyen
a) Planifier et gérer des projets par des objectifs	3	50,8%
b) …		

Ces indices de gain permettent de mieux cerner l'effet d'« apprentissage ». Notamment, on voit que le gain relatif de 50,8% est important et indique que les participants estiment avoir réellement progressé dans leur maîtrise de l'objectif, même si nous avons vu que le score moyen de 7,1 n'indique pas une maîtrise absolue.

Le traitement de ces indices par rapport à chaque objectif permet d'avoir une appréciation assez fine de l'efficacité pédagogique de la formation, notamment en mettant en évidence les objectifs qui ne seraient pas suffisamment maîtrisés et pour lesquels un complément de formation serait éventuellement nécessaire.

Cet outil ne peut en aucun cas se substituer à une véritable évaluation de la maîtrise des compétences développées par les étudiants ou les participants, mais il est utilisé actuellement avec beaucoup de bénéfices dans de nombreuses formations que ce soit dans des centres de recherche pédagogique, des IUFM ou des entreprises… ainsi que par les conseillers en pédagogie universitaires de l'IPM (Institut de pédagogie universitaire et des multimédias) dans le cadre de la politique d'évaluation formative des cours universitaires. Récemment, il a été utilisé à une échelle relativement importante pour évaluer la politique de formation à l'apprentissage d'une langue étrangère (anglais, néerlandais, ou autre langue) auprès des étudiants de l'UCL (BRAIBANT, 2004 ; BOUILLON, 2005).

Utilisation de l'outil pour évaluer la politique de formation d'une langue étrangère auprès des étudiants universitaires de l'UCL

[1] Il est difficile de dégager un seuil précis, car cela dépend aussi du type d'apprentissage. Par exemple, pour des apprentissages techniques, la différence est souvent nette : « je ne savais rien, et maintenant je sais ». Le gain relatif pourrait dans ce cas être supérieur à 50% pour considérer qu'il y a un réel accroissement de compétences. Par contre, pour des apprentissages plus sociaux ou comportementaux, on ne part pas de rien et les progrès sont moins visibles. Dans certains cas, un gain relatif de 25% pourrait alors être considéré comme significatif. Il faut donc analyser les chiffres avec prudence et circonspection, et en fonction du contexte.

Le contexte

Lors de l'année académique 2002-2003, l'IPM a été sollicité par l'Institut des Langues Vivantes (ILV) de l'UCL, opérateur de formation en matière d'apprentissage des langues étrangères, pour l'aider à évaluer l'efficacité pédagogique du Plan langues de l'UCL. Cette collaboration a débouché sur la mise au point d'un questionnaire d'évaluation qui a été administré à la fin du 2e quadrimestre 2003 à plus de 1400 étudiants appartenant à 8 facultés différentes.

Le Plan Langues de l'UCL s'articule en 2 étapes :

- les candidatures imposent dans leurs programmes au moins 60 h de cours de langues de type essentiellement **réceptif** axés sur la compréhension à la lecture ou à l'audition ;
- au cours de leur 2e cycle, tous les étudiants de l'UCL ont l'occasion de suivre au moins 30 heures de cours de langues de type **interactif.** Ces cours ont pour principal objectif de rendre l'étudiant opérationnel dans le contexte socioprofessionnel où il est appelé à s'insérer. Les cahiers de charges stipulent notamment que « les activités simuleront des situations réelles faisant appel à l'expression orale (participation à une réunion, entretien d'embauche…) et dans une moindre mesure à l'expression écrite (rédaction d'un CV, d'un courriel…) ».

L'évaluation dont il est question ici concerne uniquement **les cours interactifs du 2e cycle** pour les **4 langues majeures** du plan langues (anglais, néerlandais, allemand, espagnol).

Le questionnaire

Le questionnaire comprend deux parties distinctes. La **première partie** se présente sous la forme très répandue d'un questionnaire de satisfaction. Il est demandé aux étudiants d'exprimer leur avis à propos d'une série d'items portant sur le processus. Cette façon classique de procéder permet de recueillir de manière rapide et économique toute une série d'informations permettant au responsable d'enseignement d'avoir une idée relativement précise en ce qui concerne les qualités ou les lacunes d'un cours donné (méthodes, organisation, évaluation, enseignant, etc.), mais elle ne nous apporte que très peu de renseignements quant à l'efficacité pédagogique du cours en question. Les étudiants estiment-ils avoir acquis les compétences qui étaient visées par les objectifs de formation ? Ont-ils le sentiment d'avoir progressé depuis le début du cours ? Et si oui, dans quelle mesure ? La **deuxième partie** du questionnaire vise à répondre à ces questions particulièrement importantes. Elle se présente sous la forme de l'outil d'évaluation présenté plus haut (GERARD, 2003).

Le principe est de demander à chaque étudiant d'exprimer dans quelle mesure il s'estime lui-même plus ou moins compétent au début et à la fin de sa formation. Pour construire le questionnaire, nous avons repris les 19 objectifs de formation formulés préalablement par l'ILV et il a été demandé aux étudiants d'estimer pour chaque objectif leur niveau de compétence avant et après la formation (cf. exemples ci-dessous).

Pour chacun des objectifs suivants, estimez votre niveau de compétence atteint dans la langue cible avant et après ce cours en utilisant l'échelle suivante : 0 = pas du tout capable → 10 = tout à fait capable

	Avant le cours ILV	*A l'issue du cours ILV*
a) Dialoguer avec un interlocuteur	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9
b) Lire et comprendre un article relevant de mon domaine d'études	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9
c) Me présenter à un entretien d'embauche	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9
d) Rédiger mon curriculum vitae	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9
e) Tenir une conversation téléphonique	0 1 2 3 4 5 6 7 8 9	0 1 2 3 4 5 6 7 8 9

Figure 3 - Exemples des questions posées aux étudiants du Plan Langues

La Figure 4 montre que pour chacune des 19 compétences, la moyenne au terme de la formation est plus élevée qu'avant la formation. En première analyse, on peut donc dire **que les étudiants estiment avoir amélioré leurs compétences linguistiques**. Cela étant, le niveau estimé de maîtrise final est encore relativement faible pour toute une série de compétences qui n'atteignent pas un score moyen de 7 ou de 7,5. Ce seuil est par contre atteint dans 5 types de situations différentes, celles pour lesquelles les étudiants s'estiment les plus compétents au terme de leur formation dans la langue-cible : « *Me présenter et parler de moi* » (7,5), « *Lire et comprendre un article relevant de mon domaine d'études* » (7,3), « *Rédiger mon CV* » (7,1), « *Dialoguer avec un interlocuteur* » (7,1) et « *Rédiger un courrier électronique* » (7,0).

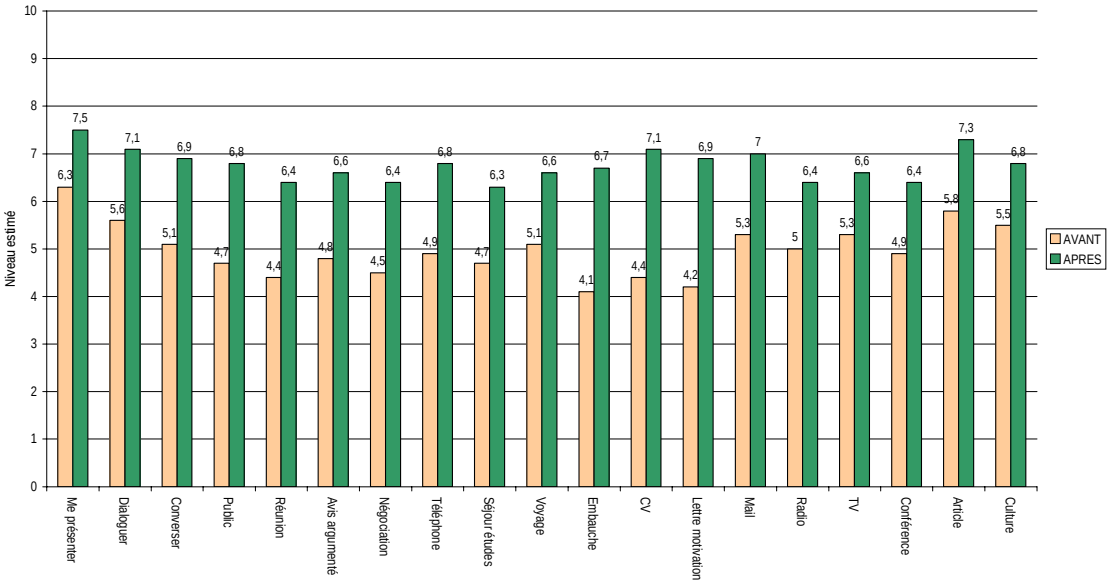


Figure 4 - Comparaison des scores moyens avant et après la formation linguistique pour chacune des 19 compétences mesurées

Comparaison des coefficients de variation

La comparaison des coefficients de variation (cf. Figure 5 ci-dessous) montre que la formation a permis de réduire sensiblement la disparité qui existait au départ, ce qui signifie que **l'apprentissage a eu un effet d' « équité »,** dans la mesure où les écarts de compétences entre les étudiants se sont réduits et que la formation a donc contribué à un plus grand « partage » des compétences. Certes, pour la majorité des compétences linguistiques, les coefficients de variation au terme de la formation sont encore relativement élevés (taux proche ou supérieur à 30 %) mais on peut se demander si on peut espérer de meilleurs résultats compte tenu de l'importante hétérogénéité caractérisant le public étudiant au début de leur formation de 2e cycle et qui se manifeste par des taux élevés généralement supérieurs à 45 %, voire à 50 %.

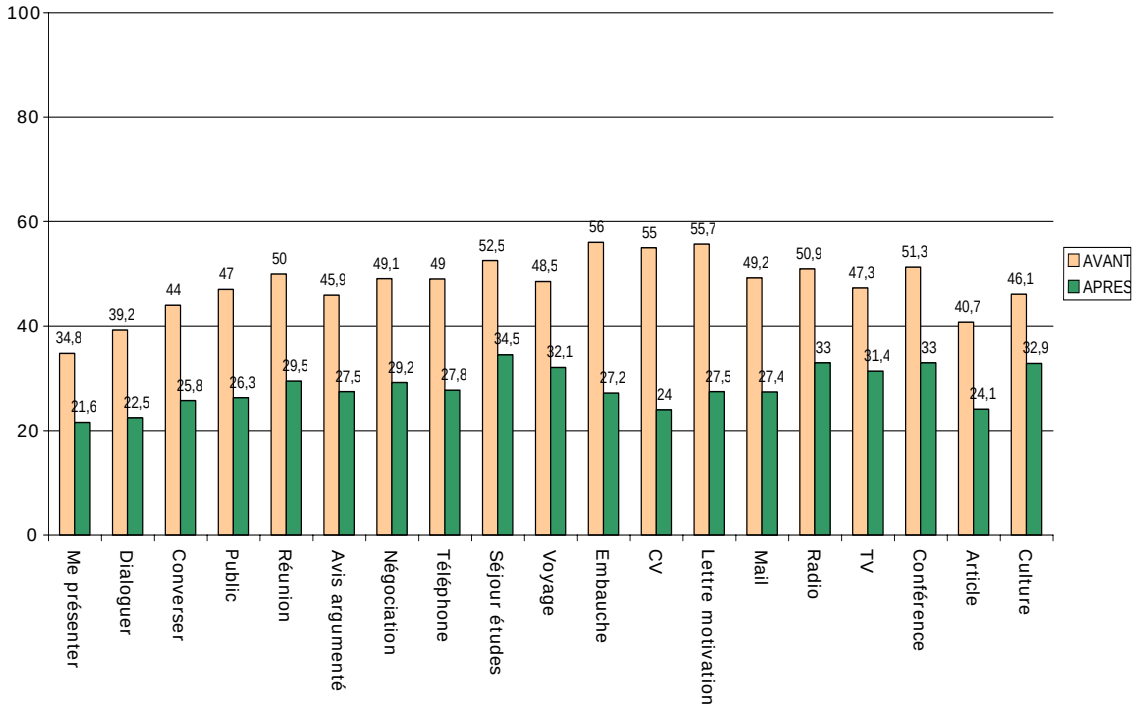


Figure 5 - Comparaison des coefficients de variation avant et après la formation linguistique pour chacune des 19 compétences visées

Analyse des effets d'apprentissage : calcul des gains relatifs moyens

En référence aux analyses proposées ci-dessus, l'analyse de l'**effet d'apprentissage** a pu être affinée en calculant des indices de **gains** entre les deux moyennes (avant /après). Ces deux indices ont été synthétisés dans le Tableau 1 et dans la Figure 6.

	Gain brut moyen	**Score MAX-Score AVANT**	**Gain relatif moyen**
Me présenter et parler de moi (famille, études, centres d'intérêt)	1,25	3,71	**33,7%**
Dialoguer avec un interlocuteur	1,58	4,44	**35,6 %**
Participer à une conversation avec plusieurs interlocuteurs	1,72	4,86	**35,4 %**
Présenter en public des informations relatives à mon domaine d'études	2,03	5,28	**38,4 %**
Participer activement à une réunion /mise en situation liées à mon domaine d'études	2,02	5,58	**36,2 %**
Présenter et développer un avis argumenté	1,81	5,21	**34,7 %**
Jouer un rôle dans une négociation	1,96	5,50	**35,6 %**
Tenir une conversation téléphonique	1,92	5,13	**37,4 %**
Préparer et effectuer un séjour d'études	1,57	5,26	29,8 %
Obtenir et utiliser des informations dans le cadre d'un voyage	1,51	4,89	**30,9 %**
Me présenter à un entretien d'embauche	2,63	5,89	**44,6 %**
Rédiger mon CV	2,77	5,62	**49,3 %**
Rédiger une lettre de motivation	2,64	5,78	**45,7 %**
Rédiger un courrier électronique	1,81	4,74	**38,2 %**
Comprendre une émission à la radio	1,40	4,95	28,3 %
Comprendre une émission télévisée	1,36	4,74	28,7 %
Comprendre une conférence	1,47	5,11	28,8 %
Lire et comprendre un article relevant de mon domaine d'études	1,52	4,18	**36,4 %**
M'ouvrir à des aspects culturels des pays de la langue cible	1,28	4,47	28,6 %

Tableau 1 - Analyse des gains pour chacune des 19 compétences

L'analyse des gains permet de mieux cerner l'effet d'apprentissage pour chacune des compétences visées. Pour 14 d'entre elles (soulignées en gras dans le tableau), on observe un gain relatif supérieur à 30 % ce qui semble indiquer que les étudiants estiment avoir réellement progressé dans leur maîtrise de l'objectif en question. Il est important de signaler que les trois compétences pour lesquelles on enregistre les gains relatifs les plus élevés (supérieurs à 40 %) c'est-à-dire « *Rédiger mon CV* » (49,3 %), « *Rédiger une lettre de motivation* » (45,7 %) et « *Me présenter à un entretien d'embauche* » (44,6 %) correspondent justement aux situations explicitement

travaillées dans le cadre de cours à visée socioprofessionnelle, tels que ceux du plan langues au 2e cycle.

À l'inverse, on peut être rassuré par le fait que les compétences pour lesquelles les gains relatifs sont inférieurs à 30 % ne constituent certainement pas des priorités pour ce niveau d'études. Trois de ces compétences, « *Comprendre une émission à la radio* », « *Comprendre une émission télévisée* » et « *Comprendre une conférence* », correspondent davantage aux objectifs poursuivis dans le cadre des cours réceptifs du 1er cycle. Les deux dernières compétences, « *Préparer et effectuer un séjour d'études* » et « *M'ouvrir à des aspects culturels des pays de la langue cible* », ne faisaient pas non plus partie des objectifs prioritaires de formation si on se fie au nombre très élevé de réponses « sans objet » cochées par les étudiants au regard de ces deux compétences.

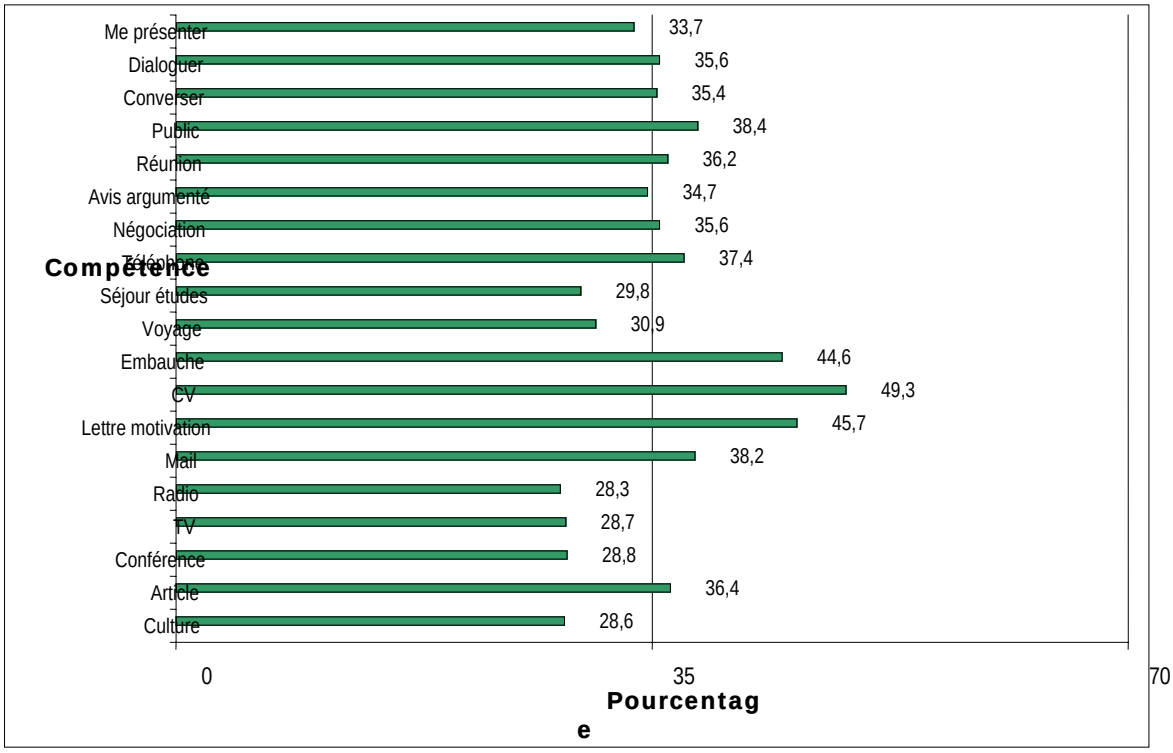


Figure 6 - Comparaison des gains relatifs d'apprentissage pour les 19 compétences de la formation linguistique

Résultats présentés en fonction des cours

Au-delà de ces constats très positifs mais aussi très globaux, il est intéressant de choisir un autre angle d'analyse et de s'intéresser aux effets d'apprentissage cours par cours et non plus pour chacune des compétences du Plan Langues. Les analyses qui

suivent ne concernent évidemment que les cours pour lesquels nous avons suffisamment d' étudiants pour tirer des conclusions statistiques (cours regroupant plusieurs séries et qui font donc appel à plusieurs enseignants). Huit cours d’anglais et un cours de néerlandais répondent à cette condition. Chacun de ces cours regroupe des étudiants appartenant à une faculté ou à un programme d’études bien déterminé. Les résultats sont résumés dans le

ci-dessous.

Cours	Nombre étudiants	Moy AVANT	Moy APREs	Coeff Var AVANT	Coeff Var APRES	Nombre d'objectifs dont le gain brut relatif est > à 30 %	Nombre d'objectifs dont le gain brut relatif est > à 40 %
ANGL 1532 (IAG/ING E)	192	5,70	7,00	37,2 %	26,1 %	9/19	3/19
ANGL 2420 (DRT)	142	4,72	6,41	46,3 %	31,5 %	12/19	2/19
ANGL 2431 (SOC)	29	5,16	7,24	44,1 %	29,8 %	18/19	10/19
ANGL 2433 (COMU)	118	5,46	6,82	43,1 %	29,6 %	11/19	3/19
ANGL 2440 (PSP)	115	4,22	6,58	60,8 %	30,7 %	18/19	9/19
ANGL 2451 (IEPR)	56	4,09	6,12	60, 0 %	31,5 %	15/19	4/19
ANGL 2454 (SBIM)	20	4,51	6,74	59,9 %	29,8 %	19/19	8/19
ANGL 2460 (SC)	53	4,13	6,63	51,0 %	25,4 %	18/19	7/19
ANGL 2470 (FSA)	103	5,35	7,11	35,4 %	20,8 %	15/19	7/19
ANGL 2480 (AGRO)	66	5,34	7,10	42,2 %	21,8 %	16/19	10/19
ANGL 2613 (GERM)	42	6,81	8,07	24,8 %	15,7 %	18/19	8/19
NEER 2420 (DRT)	182	4,86	6,52	48,0 %	28,3 %	14/19	1/19

Tableau 2 - Comparaison des résultats en fonction des cours

Les résultats confirment que les étudiants estiment avoir réellement progressé dans leur maîtrise des objectifs de la formation linguistique :

- dans chacun des cours la moyenne calculée sur l'ensemble des 19 objectifs de formation est sensiblement plus élevée au terme de la formation qu'avant la

formation (cette différence est significative au sens statistique du terme à un seuil de 0,05) ;

- si on considère qu'il y a un effet positif d'apprentissage lorsque le gain relatif est supérieur à 30 %, la lecture de l'avant-dernière colonne montre que cet effet d'apprentissage concerne un nombre important d'objectifs et ce dans tous les cours. Ceci est d'autant plus marquant que les enseignants visent nécessairement le développement et la maîtrise d'un nombre limité de compétences dans leurs cours respectifs (pas plus de 5 ou 6 compétences cibles). Certains cours enregistrent par ailleurs un nombre élevé d'objectifs pour lesquels le gain relatif est égal ou supérieur à 40 % (cf. dernière colonne) Ces résultats tendent à démontrer que les étudiants estiment avoir réellement progressé non seulement dans la maîtrise des « compétences cibles » mais aussi dans celles qui ne font pas partie des objectifs prioritaires de leur cours ;
- la comparaison des coefficients de variation (colonnes en milieu de tableau) montre que les différents cours ont réduit sensiblement les écarts de compétences qui existaient au départ (en début de formation) entre les étudiants, ce qui signifie que l'apprentissage a eu un effet d'équité et ce dans tous les cours. Dans certains cas (ANGLAIS 2440, 2451, 2454, 2460, 2480) la réduction du taux d'hétérogénéité est très importante (réduction de 50 %).

Le tableau permet également de dégager des **profils** de cours présentant des résultats quelque peu contrastés.

Un premier niveau d'analyse concerne le sentiment de compétence au départ de la formation. Par exemple, les étudiants des cours ANGL 2451 (IEPR), ANGL 2460 (Sciences), ANGL 2440 (Psychologie) et ANGL 2454 (Ssciences bio-médicales) s'estiment peu compétents au début du cours (moyenne inférieure à 4,5/10 sur l'ensemble des 19 compétences) si on les compare avec les étudiants du cours ANGL 2613 (Germanique), voire avec ceux des cours ANGL 1532 (Ingénieur commercial) ou ANGL 2433 (Communication).

Si on envisage à présent le sentiment qui anime les étudiants à la fin de la formation, on note que le niveau de maîtrise final estimé est relativement faible pour une série de cours : les cours de néerlandais (NEER 2420) et d'anglais (ANGL 2420) en faculté de droit ainsi que les cours ANGL 2451 en Education physique et ANGL 2440 en Psychologie. À l'inverse, plusieurs cours obtiennent un sentiment de compétence final supérieur ou égal à un score de 7 points sur 10. La palme revient très logiquement aux étudiants de « germanique » puisqu'ils sont les seuls à atteindre le score moyen de 8/10, mais il faut rappeler qu'ils estimaient déjà avoir un niveau de compétence élevé en début de formation.

Il est intéressant de noter que le taux d'hétérogénéité en début de formation varie considérablement d'un cours à l'autre (de 24,8 % à 60,8 %). Comme il a été dit plus haut, ce taux se réduit considérablement en sortie de formation même si certains cours enregistrent encore des taux supérieurs à 30 % (qui témoignent d'une hétérogénéité

encore importante entre les étudiants). En d'autres termes, il semble bien que certains cours soient « plus équitables » que d'autres.

La prise en compte détaillée des gains bruts relatifs est également très intéressante dans la mesure où elle met bien en évidence les cours dans lesquels certains objectifs obtiennent des gains bruts relatifs supérieurs à 60 ou 70 % (données non présentées ici). Ces scores élevés indiquent que les étudiants estiment avoir progressé de manière très nette dans leur maîtrise de l'objectif. Dans d'autres cours, les gains bruts relatifs maximum ne dépassent pas le seuil de 40 %.

Résultats du questionnaire de satisfaction

Rappelons que les étudiants ont eu l'occasion d'exprimer leur avis sur leur formation linguistique par le biais de questions plus classiques où on leur demandait d'exprimer leur accord / désaccord par rapport à toute une série d'affirmations telles que par exemple « Les thèmes abordés sont pertinents pour ma formation », « Les supports didactiques facilitent l'apprentissage », « L'enseignant(e) suscite l'intérêt des étudiant(e)s », etc.

Les résultats présentés ici ne concernent qu'un seul item, celui qui vise à résumer le sentiment général des étudiants : « *Le cours interactif m'a permis d'améliorer mes aptitudes de communication dans la langue cible* ». Les résultats présentés dans le Tableau 3 sont très positifs et confirment une nouvelle fois le « sentiment d'efficacité » dont nous avons longuement parlé ici. En effet, lorsqu'on prend en compte les 69 séries (ou cours) de notre échantillon, on constate que dans plus de la moitié des cas, 80 % au moins des étudiants expriment leur accord avec cet item. Seules 5 séries (ou cours) posent de sérieux problèmes puisque dans ce cas, moins de 50 % des étudiants estiment qu'ils ont amélioré leurs aptitudes de communication au terme de la formation linguistique.

Pourcentage moyen d'étudiants exprimant leur accord avec l'item	Nombre (et pourcentage) de séries
≥ 80 %	43 (59,4 %)
70-79 %	11 (21,7 %)
51-69 %	08 (13,0 %)
≤ 50 %	05 (05,8 %)
TOTAL	69 (100 %)

Tableau 3 - Questionnaire de satisfaction : résultats pour l' item « Le cours interactif m'a permis d'améliorer mes aptitudes de communication dans la langue cible »

Conclusions et recommandations

Sans sous-estimer l'importance de ce travail, il est important de souligner qu'il n'avait pas pour objectif de mesurer l'état réel des compétences linguistiques des étudiants, mais bien de déterminer dans quelle mesure ces derniers estimaient avoir réellement progressé entre le début et la fin de leur formation donnée dans le cadre du Plan langues. Il en ressort trois constats très positifs :

- les étudiants de 2e cycle estiment avoir améliorer sensiblement leurs compétences linguistiques ;
- globalement, on peut dire l'apprentissage a eu un effet d'équité dans la mesure où les écarts importants de compétences observés entre les étudiants en début de formation se sont fortement réduits au terme de la formation ;
- il est heureux de constater que les effets d'apprentissage se manifestent essentiellement au niveau des objectifs explicitement visés par les cours du Plan langues, c'est-à-dire les objectifs de préparation et d'insertion à la vie professionnelle (« Rédiger mon CV », « Rédiger une lettre de motivation », « Me présenter à un entretien d'embauche »).

Les responsables d'enseignement disposent à présent des résultats et des analyses nécessaires pour assurer le suivi des cours (séries) qui posent problème. De manière plus générale, il a été conseillé aux enseignants titulaires de lire attentivement les documents qui leur ont été remis par l'IPM afin d'identifier les points susceptibles d'être encore améliorés, en vérifiant par exemple si les objectifs qu'ils estiment prioritaires enregistrent bel et bien des gains bruts et relatifs élevés. Ce travail de régulation devrait être utile pour augmenter à terme l'efficacité pédagogique de ces cours (par exemple se donner l'objectif d'atteindre un sentiment d'efficacité pédagogique de 7,5 au terme de la formation), mais aussi et surtout de réduire la disparité importante qui subsiste entre les étudiants en fin de formation (se donner pour objectif d'obtenir un coefficient de variation inférieur à 30 %).

L'ILV dispose à présent des instruments nécessaires pour reproduire régulièrement ce type d'enquête, notamment au niveau des cours du plan langues du 1er cycle, pour lesquels nous ne disposons pas de données aussi précises que celles qui sont disponibles actuellement pour les cours du 2e cycle.

Atouts et limites de l'outil

Cette expérience d'évaluation du Plan Langues ainsi que bien d'autres exemples d'application (par exemple, l'évaluation formative de cours universitaires….) montre que l'utilisation de cet outil est particulièrement intéressante, mais il convient bien entendu de ne pas en oublier les principales limites :

- il ne s'agit bien sûr que d'une autoévaluation à laquelle les participants ne sont peut-être pas suffisamment préparés. Ce sont des représentations qui s'expriment et – même si nous pensons, comme l'ont montré plusieurs validations sur le terrain, que le risque d'erreur est relativement faible – on ne peut bien sûr pas être certain que la maîtrise des compétences est réellement celle déclarée ;

- tant pour des raisons de commodité que pour tenir compte de la prise de « conscience de son incompétence » (COURAU, 1993) propre à toute formation, l'outil est proposé en fin de formation, mais il est évident qu'il peut être difficile pour les participants d'estimer à ce moment quel était leur niveau de compétence avant la formation. Il est vraisemblable que les participants cherchent plus à indiquer quel est leur niveau de gain, mais cela ne représente pas une grosse difficulté dans la mesure où ce sont les calculs de gains qui permettront d'apprécier l'efficacité pédagogique de la formation ;
- surtout en formation professionnelle continue, l'expression des représentations en fin de session peut être influencée par une certaine euphorie caractéristique de ce moment, et cela d'autant plus que les questionnaires de satisfaction sont souvent remplis à la demande et en présence du formateur. Celui-ci n'aura pas manqué de créer une relation de confiance durant la formation, et il est vraisemblable que les participants ne souhaitent pas remettre en question son efficacité de manière directe. Le cas inverse peut bien sûr aussi se présenter – un formateur en tension avec les participants – et avoir également une influence sur les réponses.

En lien avec ces limites, il importe de signaler que l'utilisation des indices ne peut, selon nous, qu'être collective et non individuelle. Cela signifie que l'outil ne peut en aucun cas permettre de certifier que tel ou tel individu maîtrise ou non les acquis de la formation. Son ambition est d'apporter, grâce à un dispositif aussi léger qu'un « questionnaire de satisfaction », une information en ce qui concerne l'efficacité pédagogique globale de l'action de formation. L'outil nous semble pouvoir répondre à cette attente, pour autant que les conditions minimales d'utilisation soient remplies, notamment que les participants sachent qu'en aucun cas leurs réponses ne pourront se retourner contre eux.

Références

BOUILLON, H. (2005). Conception, réalisation et évaluation du « Plan Langues » de l'UCL. Réalisation d'objectifs en matière d'enseignement de langues étrangères à l'échelle d'une université. In COURTOIS, L. & PIROTTE, J. (Eds), L'enseignement des langues en Wallonie. Enjeux citoyens et chances pour l'avenir. Publications de la fondation Wallonne P.-M. et J.-F. Humblet. Série Recherches, tome IV. Louvain-la-Neuve.

BRAIBANT, J.-M. (2004). *Évaluation de l'efficacité du plan langues de l'UCL*. Conseil de la formation et de l'enseignement (CEFO). Université catholique de Louvain. Louvain-la-Neuve. Document interne.

COURAU, S. (1993), *Les outils d'excellence du formateur — Pédagogie et animation,* Paris : ESF éditeur.

D'HAINAUT, L. (1975). *Concepts et méthodes de la statistique* (Vol. 1). Bruxelles : Labor.

GERARD, F.M. (2003). L'évaluation de l'efficacité d'une formation, *Gestion 2000*, Vol. 20, n°3, 13-33.

GERARD, F.-M., VAN LINT-MUGUERZA, S. (2000). Quel équilibre entre une appréciation globale de la compétence et le recours aux critères ? In BOSMAN, C., GERARD, F.M., ROEGIERS, X. (Éds). *Quel avenir pour les compétences* ? Bruxelles : De Boeck Université, 135-140.

HAUSER G., MASSINGUE B., MAITRE F. et VIDAL F. (1985). *L'investissement Formation*, Paris, Les Éditions d'Organisation.

OUELLET, G. (1985). *Statistiques. Théorie, exemples, problèmes.* Sainte-Foy (Québec) : Le Griffon d'argile.

EVALUER LA QUALITE DE L'ENSEIGNEMENT SUPERIEUR : UNE CULTURE PROFESSIONNELLE EMERGENTE

HERTAY Alain, STEGEN Pierre, Haute Ecole de la Ville de Liège, Belgique[1]

Mots-clés : évaluation de la qualité - enseignement supérieur - facteurs de réussite et d'échec

Résumé :
Une Agence pour l'évaluation de la Qualité de l'Enseignement supérieur organisé ou subventionné par la Communauté française de Belgique a été créée par le décret du 14 novembre 2002. L'objectif : demander aux établissement concernés de mettre en évidence les « bonnes pratiques », les insuffisances et les problèmes à résoudre. La Haute Ecole de la Ville de Liège (établissement d'enseignement supérieur de type court, à visée professionnalisante) s'est approprié le prescrit légal en choisissant d'évaluer les facteurs de réussite et d'échec des étudiants de première année. Il s'est agi d'élaborer des outils d'évaluations et de développer des procédures relatives au pilotage du système éducatif. L'initiation de ces pratiques a doté les formateurs de nouveaux rôles en les impliquant dans une évaluation systémique de leur contexte d'enseignement.

Un questionnaire à choix multiples a été diffusé auprès des 1215 étudiants des quatre Départements entre février et mars 2005. Nous avons obtenu un retour de 605 questionnaires, soit 53% du public concerné. L'enquête a mis au jour de précieuses données concernant les différentes dimensions investiguées (facteurs motivationnels, cursus scolaire, maîtrise de la langue…). Les résultats permettent d'esquisser des pistes pour une meilleure prise en charge de l'échec : aménagements des horaires, importance accrue accordée aux facteurs identitaires, proposition d'entretiens ciblés avec les étudiants… Par ailleurs, l'enquête sera réitérée en 2006-2007.

Nous avons ensuite considéré que les conditions d'émergence d'une culture d'évaluation partagée dépendaient de l'acquisition de nouvelles compétences professionnelles, ainsi que du renouvellement de l'identité professionnelle des formateurs. Nous avons interrogé ces derniers. Nous constatons que si des modifications des comportements professionnels semblent apparaître chez certains acteurs impliqués, une identité professionnelle nouvelle n'est pas construite. Dès lors, deux questions se posent : quelles modalités d'accompagnement imaginer pour permettre aux formateurs d'acquérir une culture d'évaluation partagée ? Quelle reconnaissance accorder au sein de leurs institutions à ce nouveau rôle rempli par certains membres des Hautes Ecoles ? Leur identité professionnelle ne devrait-elle pas évoluer pour que leurs actions ne restent pas marginales et sans suite ?

---

[1] Le travail présenté ici est le fruit des activités d'un groupe de formateurs de la Haute Ecole de la Ville de Liège : Mmes A. Boileau, V. Dethier, I. Huvelle, G. Leclercq, ainsi que MM. M. Capitaine, H. Hedia, M. Herman, P. Vandenborght et P. Varetto.

Evaluer la qualité de l'enseignement supérieur : une culture professionnelle émergente

Introduction

La notion de qualité est apparue formellement dans les Hautes Ecoles (établissements d'enseignement supérieur de type court, à visée professionnalisante) de la Communauté française de Belgique, le 5 août 1995 dans le décret fixant l'organisation de l'enseignement dans les Hautes Ecoles (Chapitre 5, section 4). Plus récemment, une Agence pour l'évaluation de la Qualité de l'Enseignement supérieur organisé ou subventionné par la Communauté française a été créée par le décret du 14 novembre 2002. Cette Agence a pour missions de répondre aux recommandations européennes déjà arrêtées en cette matière et, surtout, de doter la Communauté française d'une structure organisant cette évaluation au sein des établissements. Laquelle évaluation vise à améliorer la qualité de l'enseignement supérieur dans les différents cursus tels que spécifiquement organisés par les Universités, les Hautes Ecoles ou les autres formes d'enseignement supérieur. Il ne s'agit pas d'une évaluation contrôle débouchant sur une pénalisation des établissements en termes de moyens ou d'image. Il ne s'agit pas non plus d'accréditer des institutions ou des cursus comme c'est le cas dans certains pays. En ce qui concerne les cursus ou modalités particulières à évaluer, ils sont déterminés chaque année par l'Agence de son initiative propre ou à la demande du Gouvernement, l'évaluation se référant à une série d'indicateurs[2] recouvrant l'ensemble des démarches de formation à considérer. Il s'agit donc d'améliorer la qualité de l'enseignement au sein de chaque institution en mettant en évidence les « bonnes pratiques », les insuffisances et les problèmes à résoudre.

1. Problématique

Comment la Haute Ecole de la Ville de Liège (désormais HEL) s'est-elle approprié le prescrit légal ? En initiant, dès septembre 2004, un service Qualité composé de trois coordonnateurs désignés à temps partiel (MM. M. Capitaine, A. Hertay et P. Varetto) qui a eu pour mission d'organiser plusieurs groupes de travail associant toutes les catégories d'acteurs de la HEL. L'un de ces groupes réunissait plusieurs formateurs des quatre Départements de la HEL (Mme Boileau pour le Département économique, Mmes Dethier, Huvelle et Leclercq pour le Département paramédical, MM. Herman et Stegen pour le Département pédagogique et MM. Hedia et

[2] Définis par le décret du 14 novembre 2002, ces indicateurs portent sur le cadre institutionnel, les objectifs généraux et spécifiques de l'enseignement dispensés, les programmes d'études, les étudiants, l'information et le suivi pédagogique, la recherche liée à l'enseignement, les services à la collectivité, le personnel et la gestion des ressources humaines, les ressources et les équipements, les relations extérieures, la gestion de la qualité, le fonctionnement et la politique générale de gestion et d'organisation, l'analyse des forces et faiblesses, des opportunités et des risques.

Vandenborght pour le Département technique)[3]. Ceux-ci ont choisi de faire converger leur réflexion sur une problématique transversale commune, jugée décisive : évaluer les facteurs de réussite et d'échec des étudiants de première année inscrits dans les treize baccalauréats organisés par la Haute Ecole.

Que sait-on des préacquis des étudiants ? Construisent-ils un projet personnel d'apprentissage ? Quelles sont leurs méthodes de travail ? Comment abordent-ils les évaluations et les examens ? … Telles sont les questions auxquelles nous avons voulu essayer d'apporter des éléments de réponse, bien sûr dans les limites imposées par nos moyens de recherche. En effet, un projet inter-Départements de ce type constitue une innovation au sein de la HEL et répond à l'injonction ministérielle qui lui a été faite. Il s'est agi d'élaborer des outils d'évaluations appropriés à notre propre contexte d'enseignement et de développer, en interne, des procédures relatives au pilotage de notre système éducatif. Le bilan de notre enquête revêt par conséquent un caractère exploratoire évident : le recueil d'informations qui a été réalisé représente un coup de sonde qui mériterait d'être élargi, affiné et davantage documenté par des enquêtes ultérieures, dans le cadre de recherches à visée confirmative.

Par ailleurs, l'initiation de ces pratiques a incité les formateurs à endosser de nouveaux rôles et de nouvelles responsabilités en les impliquant dans une évaluation systémique de leur contexte d'enseignement. C'est une culture professionnelle rénovée qui est convoquée, requérant des compétences et des composantes identitaires professionnelles inédites. Cette rénovation de la culture professionnelle des formateurs suppose une culture d'évaluation partagée. Nous avons donc cherché à réaliser une évaluation globale du projet visant à faire émerger les changements survenus du côté des formateurs impliqués, les difficultés qu'ils ont rencontrées, les stratégies qu'ils ont mises en place pour les contourner…

Dès lors, deux niveaux de questionnements tissent notre réflexion. Primo, quels résultats notre enquête a-t-elle permis d'obtenir ? Secundo, comment les pratiques émergentes d'évaluation tendent-elles à reconfigurer la culture professionnelle des formateurs ? Dans un premier temps, nous reviendrons sur la phase d'élaboration du questionnaire adressé aux étudiants de première année de la HEL. Nous donnerons ensuite un aperçu de notre corpus de données, ainsi que du traitement qui lui a été appliqué. Nous procéderons de même pour le questionnaire diffusé auprès des formateurs participant au groupe de travail.

2. L'enquête sur les facteurs de réussite et d'échec des étudiants de première année à la HEL

2.1. La construction du questionnaire destiné aux étudiants

[3] Signalons que ces formateurs se sont impliqués sur la base d'une participation volontaire et bénévole. Nous reviendrons sur cette modalité de constitution du groupe de travail lorsque nous aborderons notre seconde problématique portant sur la culture professionnelle des formateurs.

Nous nous sommes tout d'abord penchés sur les nombreuses études qui envisagent la question de la réussite et de l'échec dans l'enseignement supérieur. Remarquons que ces études, parfois anciennes et parfois françaises (Bourdieu et Passeron, 1964, 1970 ; Coulon, 1997 ; Bodson, 1999 ; Romainville, 1999 ; Erlich et al., 2000 ; Leclercq et al., 1998, 2000) s'attachent à la promotion de la réussite à l'Université, les Hautes Ecoles étant rarement traitées[4]. Nous avons plus particulièrement sollicité les recherches de Romainville (1999), pour l'éclairage pédagogique, et de Bodson (1999), pour l'éclairage sociologique. La taille de notre population (1215 étudiants) nous a amenés à privilégier l'emploi d'un questionnaire à choix multiples.

Les facteurs de réussite et d'échec scolaires constituent un phénomène complexe et multiforme : nous avons dès lors distingué quatre problématiques distinctes qui structurent notre questionnaire :

- Le projet personnel,
- Les préacquis,
- La vie étudiante,
- Les évaluations de janvier

Cette subdivision a pour but de nous donner plusieurs pistes d'analyses qui renvoient à autant d'hypothèses préalablement négociées par le groupe de travail. Ainsi, le volet « projet personnel » pose la question du « sens » des études entreprises. En effet, à la suite de Bodson (1999), nous formulons l'hypothèse selon laquelle le projet personnel de l'étudiant peut avoir une incidence sur sa réussite académique. Nous entendons par là qu'un choix assumé et raisonné du cursus entrepris constituerait un possible facteur de réussite. Par ailleurs, les préacquis désignent ici les filières d'études suivies dans le secondaire. Plusieurs études ont montré la valeur prédictive du passé scolaire sur la réussite ou l'échec dans l'enseignement supérieur (Dal et al., 1994 ; Romainville, 1999). Notre hypothèse est donc la suivante : les étudiants « à l'heure » et qui s'estiment avoir bien été préparés dans le secondaire seraient les plus susceptibles de réussir. Nous essayerons de vérifier si cette hypothèse se vérifie pour les étudiants de la HEL. En outre, les modalités de travail dans l'enseignement supérieur sont envisagées dans la partie que nous avons appelée « Vie étudiante ». Cette partie interroge également les étudiants sur l'information qu'ils reçoivent concernant les cours, l'administration, les examens, ainsi que sur les ressources qui leur sont proposées par leur Département. Ces données permettent d'envisager les liens éventuels entre les pratiques d'apprentissage des étudiants et leur réussite (Bodson, 1999). Il s'agit là de notre troisième hypothèse : il est légitime de penser qu'une corrélation existe entre la réussite des étudiants et certains modes de travail (étude régulière, présence aux cours, prises de notes…). Enfin, sont abordés dans la dernière partie du questionnaire les évaluations de janvier, les facteurs d'attribution des causes de réussite et/ou d'échec et les besoins de remédiations souhaitées. Ainsi, Bautier (1995) a montré que plus le sujet s'inscrit dans une logique

[4] A cet égard, il serait intéressant de disposer de données de comparaison afin de vérifier si cette problématique se décline de la même manière à l'Université et dans l'enseignement supérieur de type court.

de causalité externe (la réussite ou l'échec dépendent de facteurs qui lui sont extérieurs), moins il sera à même de se mobiliser sur un changement.

Le questionnaire, de type exploratoire, a été diffusé auprès des 1215 étudiants des quatre Départements de la HEL entre le 14 février et le 8 mars 2005. Bien que le nombre d'évaluations effectuées d'un Département à l'autre soit variable, les étudiants disposent à ce moment d'un certain nombre de résultats relatifs aux évaluations de janvier. La plupart des étudiants sont alors en mesure de se percevoir comme en situation de « réussite » ou d'« échec ». En fonction des ressources qui nous étaient allouées, nous nous sommes dirigés vers une collecte des données durant les cours, éliminant des méthodes sans doutes plus fiables mais aussi plus onéreuses comme l'enquête par voie postale à partir d'un échantillon raisonné. Le risque était, dès lors, de ne toucher que les étudiants fréquentant régulièrement les cours. Aussi, avons-nous mis en œuvre, afin d'obtenir un taux de participation maximale, plusieurs mesures qui devaient permettre de limiter ce biais. Ainsi, c'est le groupe de formateurs impliqués dans le projet – rappelons-le, groupe constitué d'enseignants provenant des quatre Départements – qui a assuré l'administration du questionnaire dans les classes. Ils ont pris soin de choisir les cours les plus fréquentés par les étudiants. Par ailleurs, dans la mesure du possible, plusieurs passages dans les classes ont eu lieu.

2.2. Notre échantillon

Pour 1215 étudiants inscrits à la HEL en première année, nous avons obtenu un retour de 605 questionnaires à la date du 8 mars 2005, soit 53% du public concerné. Il est clair qu'à ce moment de l'année de nombreux étudiants ont déjà décroché. Il nous semblait toutefois important d'attendre cette période afin de disposer d'informations sur les taux de réussites des examens du mois de janvier.

Remarquons que le dépouillement des questionnaires a démontré le sérieux avec lequel les étudiants l'ont complété. En effet, aucun commentaire déplacé n'est apparu. Seules quelques notations concernant des difficultés de compréhension de certaines questions à choix multiples ont été constatées. Par ailleurs, signalons que nous avons choisi de mettre en confiance les étudiants en leur proposant l'anonymat, ce qui ne nous a pas permis par la suite de recontacter personnellement les étudiants dont nous aurions aimé approfondir les réponses. Compte tenu des conditions d'enquête, les réponses sont donc restées tributaires de la participation volontaire des étudiants. Il faut par conséquent considérer que les effectifs qui nous ont répondu ne sont qu'imparfaitement représentatifs des populations de référence, leur représentativité pouvant être affectée par certains biais. Il sera intéressant de réitérer cette enquête ultérieurement afin de compléter ces premiers résultats. Cette démarche est envisagée pour l'année 2006-2007.

2. 3. Résultats du questionnaire destiné aux étudiants

L'analyse statistique des données (logiciel SAS) a été entreprise et une première partie du travail a été soumise à l'ensemble du groupe de formateurs. Les remarques émises nous ont permis d'approfondir le traitement des données entre septembre et

décembre 2005. Ces remarques nous ont conduits à distinguer, dans un rapport d'activité présenté en Collège de Direction en janvier 2006, des résultats relatifs à chacun des quatre Départements de la HEL. Bien évidemment, nous avons travaillé avec prudence et n'avons retenu que les résultats les plus significatifs : la nature des données recueillies n'a pas toujours permis des traitements statistiques très élaborés. Dans les limites imparties à cette communication, nous ne pouvons rendre compte de façon exhaustive des multiples enseignements apportés par le traitement des données. Par conséquent, nous présenterons ici, à titre d'exemples représentatifs des divers traitements effectués, quelques résultats obtenus auprès des étudiants du Département économique de la HEL, soit 134 étudiants (22,15% de l'échantillon global) inscrits en première année d'un des quatre baccalauréats du Département (comptabilité, gestion hôtelière, relations publiques, secrétariat de direction).

L'enquête a mis au jour de précieuses données concernant les différentes dimensions investiguées, qu'il s'agisse du choix des études, des facteurs motivationnels, du cursus scolaire, de la charge horaire, de la maîtrise de la langue…

Ainsi, nous avons cherché à mesurer la construction des projets personnels des étudiants en examinant les critères qui ont présidé au choix du baccalauréat entrepris. La question 9 de l'enquête nous informe sur le type de motivation qui a conduit l'étudiant à s'inscrire dans sa filière d'étude :

Tableau 1 – La construction du projet personnel (Département économique).

9. Quelles sont <u>les deux raisons principales</u> du choix de vos études actuelles ? (faites figurer 1 devant le plus important et 2 devant le second)		
	Choix 1	**Choix 2**
1. Par goût intellectuel	29,85	12,69
2. Pour les débouchés professionnels	34,33	26,87
3. Pour la facilité des études	0	1,49
4. Pour réaliser un projet professionnel	11,19	17,16
5. Pour le « style de vie » que cette profession génère (vacances, études, aspects financiers, etc.)	2,99	11,94
6. Par hasard, parce que je ne savais pas quoi faire	4,48	2,24
7. Parce que certain(e)s de mes ami(e)s y étaient inscrit(e)s	0	0,75
8. Parce que mon entourage me l'a conseillé	1,49	11,94
9. Parce que j'avais échoué ailleurs et/ou je n'avais pas d'autres choix	2,24	1,49

Les deux critères principaux qui incitent les étudiants à choisir leur filière d'étude sont donc les débouchés professionnels et l'engagement par goût intellectuel. Plus d'un étudiant sur trois privilégie cette dernière raison en premier choix et plus d'un quart en second choix. Autrement dit, plus de 60% des étudiants cochent l'item « Par goût intellectuel » quant au choix de leurs études actuelles. Nous avons alors cherché à savoir si ces choix ont une incidence sur la réussite académique de l'étudiant. Le test de Chi carré n'a pas permis de vérifier cette hypothèse pour les items 2 à 9 mais bien pour le premier item : l'engagement par goût intellectuel.

Tableau 2 – Effet du choix de l'item 1 sur la réussite/échec (Département économique).

Item 1 : Par goût intellectuel	**Echec (0 à 25%)**	**Echec (26 à 50%)**	**Réussite (51 à 75%)**	**Réussite (76 à 100%)**
Non choisi	30,26%	25%	19,74%	25%
Choisi	7,02%	22,81%	24,56%	45,61%

On constate que 45,61% des étudiants ayant choisi l'item « Par goût intellectuel » ont obtenu un résultat total pour leurs évaluation de janvier variant entre 76 et 100%. A l'inverse, 30,26% des étudiants n'ayant pas choisi cet item sont en échec aggravé (résultat de moins de 25%). Les étudiants qui choisissent cet item semblent donc avoir tendance à mieux réussir leurs évaluations.

L'âge, et à travers lui la réussite scolaire antérieure, nous paraissait constituer un des critères les plus déterminants du taux de réussite des étudiants en première année. Il s'agit là de notre seconde hypothèse.

Tableau 3 – Effet du passé scolaire sur la réussite/échec (Département économique).

	Echec (0 à 25%)	**Echec (26 à 50%)**	**Réussite (51 à 75%)**	**Réussite (76 à 100%)**
Etudiants à l'heure à 18 ans	12,82%	21,79%	24,36%	41,03%
Etudiants ayant un an de retard	23,08%	19,23%	23,08%	34,62%
Etudiants ayant plus d'an de retard	40,74%	33,33%	11,11%	14,81%

Nos données ne démentent pas la tendance des études antérieures sur la réussite à l'Université (Dal et al., 1994 ; Romainville, 1999) : seuls 34,61% des étudiants à l'heure sont en situation d'échec en janvier, alors que les étudiants accusant plus d'une année de retard sont 74,07% à ne pas réussir en janvier. Signalons que l'étude de Dal et al. (1994, p.136) sur les étudiants de la Communauté française (UCL/ULB) souligne le fait que les étudiants à l'heure (âgés de 18 ans) ont 50% de chances de passer dans l'année supérieure alors que la probabilité n'est que de 30% pour les étudiants âgés de 19 ans et de 20% pour les étudiants accusant deux années de retard.

On s'en souvient, notre troisième hypothèse envisageait les corrélations possibles entre les pratiques d'apprentissage des étudiants et leur réussite. Les résultats statistiques pour cette question se sont révélés peu exploitables. Nous proposerons néanmoins des résultats relatifs au temps consacré aux études en dehors des heures de cours. Cette question se subdivisait en deux volets : la fréquence moyenne de temps consacré aux études en dehors des heures de cours durant la semaine et durant le week-end. Si, à nouveau, les résultats pour le temps de travail en semaine sont non

significatifs, ceux portant sur le week-end sont quant à eux intéressants. 41,04% des étudiants déclarent travailler plus d'1h30 le week-end, 18,66% travaillent entre 1h et 1h30, 21,64% entre 30 minutes et 1h et 17,16% travaillent moins de 30 minutes (1,49% de non-réponse). Comparons ces données avec les taux de réussite et d'échec de notre échantillon.

Tableau 4 – Effet du temps de travail le week-end sur la réussite/échec (Département économique).

	Echec (0 à 25%)	**Echec (26 à 50%)**	**Réussite (51 à 75%)**	**Réussite (76 à 100%)**
0 à 30 min.	26,09%	34,78%	17,39%	21,74%
30 min. à 1h.	35,71%	32,14%	7,14%	25%
1h. à 1h. 30 min.	16%	28%	24%	32%
Plus d'1h.30 min.	10,91%	12,73%	30,91%	45,45%

On s'aperçoit que les étudiants qui assurent travailler le week-end plus de 1h30 minutes réussissent à 76,36% alors que les étudiants qui disent travailler moins de 30 minutes échouent à 60,87%. Nuançons cependant ces résultats en observant que 23,64% des étudiants affirmant travailler plus de 1h30 minutes sont en situation d'échec en janvier. Cette donnée devrait être approfondie dans la perspective des études qui ont montré que le nombre d'heures de travail hebdomadaire importait moins que la manière dont le travail était organisé (Leclercq et al., 1997). Lors de l'administration d'un prochain questionnaire, il serait sans doute utile d'essayer d'appréhender plus finement l'organisation et la qualité du travail effectué par les étudiants.

Pour clore cet aperçu de nos résultats, nous considérerons le type de logique auquel recourt l'étudiant pour expliquer son échec ou sa réussite : s'agit-il en effet d'une causalité externe ou interne ? La question 27 de notre enquête abordait ce point :

Tableau 5 – Attribution de la réussite ou de l'échec (Département économique).

27. Selon votre expérience personnelle, quelles sont les causes générales de réussite aux évaluations et/ou examens ?

Dans la colonne « Tout à fait » classez les deux items avec lesquels vous êtes tout à fait d'accord en faisant figurer 1 devant le plus important et 2 devant le second.

Dans la colonne « Pas du tout » classez les deux items avec lesquels vous n'êtes pas du tout d'accord en faisant figurer 1 devant le plus important et 2 devant le second

a.Tout à fait			**b. Pas du tout**	
Choix 1	**Choix 2**		**Choix 1**	**Choix 2**
0	4,48	1. La chance	41,04	8,21
38,06	17,16	2. L'organisation du travail	1,49	4,48
	3,73	3. Les recherches personnelles en dehors du cours	2,24	8,21
17,91	16,42	4. Les aptitudes intellectuelles	2,99	3,73
0,75	6,72	5. Les relations avec les condisciples	5,97	12,69
2,24	0,75	6. L'absence de problèmes d'ordre affectif ou personnel	3,73	3,73
1,49	1,49	7. Des facteurs d'ordre culturel (culture ou langue d'origine)	2,24	3,73
0,75	5,22	8. La bonne maîtrise des compétences linguistiques	2,99	6,72
1,49	2,99	9. Les relations avec les professeurs ou certains professeurs	6,72	17,16
13,43	14,18	10. L'intérêt pour les études entreprises	0,75	0,75
	2,24	11. Des facteurs d'ordre socio-économique (origine sociale, moyens financiers)	10,45	12,69
11,94	12,69	12. La formation préalable dans l'enseignement secondaire	6,72	5,22

Nous avons considéré les différents items en fonction de leurs caractère plutôt interne (2, 3, 4, 5, 6, 8, 10) ou plutôt externe (1, 7, 9, 11, 12) et réalisé un nouveau test de Chi carré. Il s'avère significatif pour les items 4 et 10.

Tableau 6 – Choix de l'item 4 comme facteur de réussite (Département économique).

Item 4 : Les aptitudes intellectuelles	**Echec (0 à 25%)**	**Echec (26 à 50%)**	**Réussite (51 à 75%)**	**Réussite (76 à 100%)**
Non choisi	27,27%	23,86%	20,45%	28,41%
Choisi	6,67%	24,44%	24,44%	44,44%

Tableau 7 – Choix de l'item 10 comme facteur de réussite (Département économique).

Item 1 : L'intérêt pour les études	**Echec (0 à 25%)**	**Echec (26 à 50%)**	**Réussite (51 à 75%)**	**Réussite (76 à 100%)**

entreprises				
Non choisi	22,68%	28,87%	23,71%	24,74%
Choisi	13,89%	11,11%	16,67%	58,33%

Remarquons qu'une majorité d'étudiants ayant obtenu de très bons résultats en janvier attribuent leur réussite à des facteurs internes : leur intérêt pour leurs études et leurs aptitudes intellectuelles. Les données ne nous permettent cependant pas de dégager précisément des causes d'attribution internes ou externes pour les étudiants en situation d'échec.

3. Les pratiques émergentes d'évaluation reconfigurent-elles la culture professionnelle des formateurs ?

3.1. Le profil des formateurs impliqués dans le projet

Comme nous l'avons signalé dans notre introduction, la notion de gestion de la qualité est relativement récente dans l'enseignement supérieur organisé par la Communauté française. Un bref historique nous semble utile pour expliquer comment une procédure d'évaluation de la qualité a progressivement vu le jour au sein de la HEL. Entre janvier et juin 2004, plusieurs formateurs des quatre Départements de la HEL ont été contactés par leurs directions respectives. Choisis pour leur intérêt pour l'évaluation de la qualité ou leurs expériences professionnelles de gestionnaires de la qualité dans le monde industriel, ces formateurs ont constitué un groupe de réflexion dont la principale mission était de doter la HEL d'un dispositif de gestion de la qualité en adéquation avec les exigences du législateur. Ces échanges ont conduit à la création, en septembre 2004, d'un Service Qualité chargé de définir des objectifs d'évaluation de la qualité et de piloter divers groupes de travail pour atteindre ces objectifs. Trois formateurs ont été détachés partiellement de leurs charges d'enseignants pour assurer la gestion du Service. Il s'agissait de deux maîtres-assistants en communication et d'un maître-assistant en sciences économiques[5].

Parallèlement à un groupe de formateurs développant des actions de mobilité des étudiants dans le cadre des programmes européens Leonardo et Erasmus, un second groupe de travail a choisi d'axer son action sur la réalisation de l'enquête évoquée dans les pages qui précèdent. La plupart des membres constituant ce groupe sont issus du groupe de réflexion impliqué dès janvier 2004 par leurs directions. Leur investissement dans le projet est bénévole. Ces formateurs continuent à prester leurs charges d'enseignant à temps complet et ne bénéficient pas de détachement de spécifique pour cette action.

Tableau 8 – Profil des formateurs.

Département économique	Une Maître-assistante en langues germaniques

[5] « Maître-assistant » et « maître de formation pratique » sont les principaux titres qui désignent les formateurs dans l'enseignement supérieur de type court de la Communauté française de Belgique.

Département pédagogique	Un Maître-assistant en mathématiques Un Maître-assistant en pédagogie
Département paramédical	Une Maître-assistante en logopédie Une Maître de formation pratique en logopédie Une Maître de formation pratique en logopédie
Département technique	Un Maître-assistant en chimie Un Maître-assistant en électronique appliquée

Remarquons que tous ces formateurs possèdent une certaine ancienneté à la HEL. Par ailleurs, la plupart sont déjà impliqués dans leurs Départements à divers titres : développement de projets, gestion des stages, membres d'instances internes de la HEL (Conseil pédagogique, Conseil social…), etc. En outre, plusieurs d'entre eux disposent d'expériences professionnelles dans le secteur privé ou universitaire (assistanat, activité de chercheurs, etc.).

3.2. Le questionnaire adressé aux formateurs

Comme le spécifie le décret Qualité du 14 novembre 2002, l'évaluation en cours doit impliquer des membres issus des différentes composantes de l'entité évaluée : personnel académique, scientifique, administratif, technique et étudiant. Il s'agit là d'un véritable processus de développement des compétences professionnelles des individus puisque ceux-ci se voient contraints de faire face à une situation professionnelle inédite les amenant à évaluer la qualité de l'enseignement au sein de leur institution. Dans le cadre de l'enquête sur les facteurs de réussite et d'échec des étudiants de la HEL, il était donc légitime de s'interroger sur les effets occasionnés par le projet sur les formateurs qui l'ont conduit. En effet, si, à la suite de Barbier (1996, pp. 40-41), nous considérons l'identité professionnelle comme l'ensemble des composantes représentationnelles, opératoires et affectives produites par l'histoire particulière d'un individu et qu'il est susceptible de mobiliser à un moment donné pour une pratique, quelles modifications des composantes identitaires des formateurs la réalisation de l'enquête aura-t-elle produites ? Nous avons interrogé les formateurs afin de recueillir leurs représentations. Six questionnaires complétés nous ont été retournés, ce qui nous amènera à nouveau à rester très mesurés dans nos conclusions. Le questionnaire utilisé[6] s'articulait sur plusieurs hypothèses que nous présenterons en les faisant suivre par une synthèse des réponses fournies par les acteurs impliqués. Nous tirerons ensuite des conclusions de ces données.

Dans un premier temps, nous avons formulé l'hypothèse selon laquelle le décret Qualité avait amené les formateurs à définir un positionnement critique personnel vis-à-vis de celui-ci. La lecture des réponses nous permet de constater que ce positionnement a été compliqué par un manque de précisions préalables du législateur. Ainsi, une confusion a, un temps, existé entre l'évaluation telle que proposée par le décret et les processus visant à une accréditation qualité pratiqués par le monde industriel. Dès que des éclaircissements ont été obtenus pour cette question, certains formateurs ont émis des doutes sur les possibilités d'être à la fois évaluateurs

[6] cf. annexe 1.

et évalués dans le projet. D'autres ont estimé que la démarche pouvait entraîner des changements profitables pour la HEL.

Les ressources mobilisées (savoirs, savoir-faire, etc.) par le dispositif Qualité laissaient supposer le développement de nouvelles compétences professionnelles des formateurs. Il s'agit là de notre seconde hypothèse. Ces compétences portent sur l'évaluation des étudiants, la réflexion sur une méthode de travail commune, le partage de pratiques, l'adoption d'une posture de chercheur, etc. Pour les formateurs, il s'agissait d'analyser une situation-problème complexe et de développer des solutions adaptées, construites et partagées avec des collègues (Beckers, Paquay et al., 2003, p. 6). Les attitudes des formateurs sont ici plus contrastées. Un formateur estime n'avoir développé aucune compétence nouvelle, voire ne pas avoir pu faire preuve de compétences déjà acquises. Les autres considèrent que le projet les a amenés à modifier leur perception de l'institution grâce aux contacts avec des collègues de disciplines et de Départements différents. Cette modification affecte aussi leur façon d'envisager les comportements étudiants et, pour certains, a modifié leurs pratiques (évaluation des étudiants, charge de travail, etc.). Relevons que la fierté d'avoir conçu et mis en pratique un outil d'évaluation est manifeste, ainsi que le désir de rendre cet outil récurrent.

Nous avons ensuite envisagé les possibles transformations identitaires liées au processus, élargissant de la sorte notre problématique. La construction de compétences professionnelles conduit-elle à une nouvelle identité professionnelle ? L'enjeu est ici de savoir si cette double construction a eu lieu, amorçant idéalement une construction professionnelle identitaire positive et forte (Beckers, 2002). Ce processus repose sur deux facettes distinctes : la reconnaissance ou l'attribution d'une identité par autrui d'une part et l'intériorisation des modèles identitaires dans l'identité pour soi d'autre part (Dubar, 2000). Les réponses au premier aspect de la construction identitaire, la reconnaissance par autrui, sont hélas unanimes : l'institution ne dispose pas d'un dispositif de reconnaissance suffisamment fort pour que l'identité professionnelle des formateurs en soit changée. Ceux-ci le regrettent et souhaiteraient des renforcements de leur action auprès de leurs collègues. L'un des formateurs reconnaît que son image personnelle du métier a été modifiée, l'expérience lui ayant apporté une vision pédagogique d'ensemble qui lui paraît enrichir sa pratique et pouvoir amorcer une nouvelle identité professionnelle personnelle.

Notre dernière hypothèse porte sur l'émergence possible d'une nouvelle culture professionnelle conduisant à une culture d'évaluation partagée au sein de l'institution. Les résultats qui précèdent limitent pour l'instant le partage des pratiques au groupe de formateurs initial. Signalons néanmoins que, dans le Département économique, un nouveau groupe de formateurs s'est approprié les résultats de l'enquête administrée en 2004-2005 et travaille actuellement à un enrichissement du recueil de données axé sur le passé scolaire. Ce choix découle des résultats obtenus pour les étudiants du Département (cf. la présentation des résultats ci-dessus).

4. Conclusion

L'ambition du législateur par la promulgation du décret Qualité du 14 novembre 2002 est de développer une culture de l'évaluation partagée par les établissements d'enseignement supérieur de la Communauté française. Nous avons montré comment un établissement, la Haute Ecole de la Ville de Liège, a répondu à cette attente. Nous avons exposé la méthodologie employée pour réaliser une enquête, inédite au sein de la HEL, portant sur les facteurs de réussite et d'échec des étudiants de première année. Nous avons présenté quelques résultats significatifs, résultats pour lesquels nous tirons des conclusions très prudentes : pour leur conférer une plus grande validité et généralisation, l'enquête devra être reproduite et affinée avec d'autres cohortes.

Nous avons ensuite élargi notre réflexion en nous interrogeant sur les conditions d'émergence d'une culture d'évaluation partagée. A la suite de Barbier (1996), Dubar (2000), Beckers, Paquay et al. (2003), nous avons considéré que ces conditions dépendaient de l'acquisition de nouvelles compétences professionnelles, ainsi que du renouvellement de l'identité professionnelle des formateurs. A nouveau, nous restons mesurés sur les résultats auxquels nous aboutissons : nous avons interrogé un groupe de formateurs restreint sur une action unique. Il serait donc indispensable d'étendre cette réflexion à plus large échelle. Néanmoins, nous constatons que si des modifications des comportements professionnels semblent apparaître chez certains acteurs impliqués, une véritable identité professionnelle nouvelle n'est pas construite. Dès lors, le souhait du législateur nous renvoie à deux questions :

- quelles modalités d'accompagnement imaginer (échanges d'outils, organisme de formations…) pour permettre aux membres des Hautes Ecoles d'acquérir une culture d'évaluation partagée ?
- quelle reconnaissance accorder au sein de leurs institutions à ce nouveau rôle rempli par certains membres des Hautes Ecoles ? Leur identité professionnelle ne devrait-elle pas évoluer pour que leurs actions ne restent pas marginales et sans suite ?

5. Références

Barbier, J.-M. (dir.) (1996). *Savoirs théoriques et savoirs d'action*. Paris : PUF.

Bautier, E. (1995). *Pratiques langagières, pratiques sociales*. Paris : L'Harmattan.

Beckers, J. (2002). Préparer de futurs enseignants à développer les compétences essentielles chez tous leurs élèves : les enjeux d'une réforme. *Puzzle*.

Beckers, J., Paquay, L., Scheepers, C. et al. (2003), *Quels dispositifs favorisent le dveloppement des compétences professionnelles dans l'enseignement supérieur ? Et à quelles conditions ?*. Liège/Louvain la neuve : Université de Liège, Université catholique de Louvain.

Bodson, X. (sous la direction scientifique de G. Bajoit) (1999). *La vie d'étudiant. Analyse des manières d'être étudiant et de leurs conséquences sur la réussite et l'échec en première candidature*. Louvain la Neuve : Université catholique de Louvain.

Bourdieu, P. & Passeron J.-C. (1964). *Les héritiers. Les étudiants et la culture*. Paris : Editions de Minuit.

Bourdieu, P. & Passeron J.-C. (1970). *La reproduction. Eléments pour une théorie du système d'enseignement*. Paris : Editions de Minuit.

Coulon, A. (1997). *Le métier d'étudiant*. Paris : P.U.F.

Dal, L. et al. (sous la direction scientifique de Ch. Wattelar) (1994). Etude prospective de la population étudiante des universités de la Communauté française de Belgique. Louvain la Neuve/Bruxelles : Institut de Démographie de l'Université catholique de Louvain, Laboratoire de Méthodologie du Traitement des Données de l'Université Libre de Bruxelles.

Dubar, C. (2000). *La socialisation*. Paris : Armand Colin.

Erlich, V. et al. (2000). La socialisation des étudiants débutants. In *Les Dossiers*, 115. Paris : Ministère de l'Education nationale.

Leclercq, D. et al. (1997). *Franchir le cap des candis. Inscriptions et réussite en première année universitaire en Communauté française de Belgique de 1988 à 1995* : (Rapport du groupe de travail « réussites »). CIUF.

Leclercq, D. et al. (1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.

Leclercq, D. et al. (2000). *Le projet Mohican check up : quelles connaissances chez les étudiants entrant à l'Université* : (Actes du 1er Congrès des chercheurs en éducation). Ministère de la Communauté française de Belgique.

Romainville, M. (1999), Les facteurs prédictifs de la réussite à l'université. In S. Louryan & F. Thys-Clément (eds), *Enseignement secondaire et enseignement universitaires. Quelles missions pour chacun ?*. Bruxelles : Editions de l'Université libre de Bruxelles.

6. Annexe : le questionnaire enseignant

1. Le décret Qualité du 14 novembre 2002 invitait les hautes écoles à mettre en place des procédures d'évaluation de la qualité. Au début du projet, pensiez-vous que cette mission incombait aux hautes écoles ? A présent, qu'en pensez-vous ?

2. Comment avez-vous été informé(e) de la constitution du groupe Qualité ?

3. Quels motifs vous ont amené(e) à vous y inscrire ?

4. En ce qui vous concerne, avez-vous eu le sentiment de développer de nouvelles compétences professionnelles grâce au dispositif Qualité ? Si oui, lesquelles ? Si non, pourquoi ? Ces compétences peuvent être en lien avec une évaluation inédite des étudiants, une réflexion sur leurs méthodes de travail, un partage de pratiques, une posture de chercheur à adopter…

5. En ce qui vous concerne, avez-vous le sentiment que le projet Qualité a modifié votre identité de formateur/trice ? Pensez-vous que votre rôle, votre mission, votre statut… ont été changés ? En quoi ?

6. En ce qui vous concerne, quelles difficultés avez-vous rencontrées ? Quelles stratégies avez-vous mises en place pour les contourner ? Ces difficultés concernaient-elles : la gestion du temps, la déontologie, la dynamique du groupe Qualité, la direction, l'implication des autres collègues, les compétences à mobiliser, les problèmes logistiques… ?

7. Pensez-vous que le projet a favorisé la mise en place d’une nouvelle culture professionnelle ? Ainsi, a-t-il contribué à instaurer une « culture d’évaluation partagée » ? Pensez-vous qu’il a contribué efficacement au pilotage de notre propre système éducatif ?

8. Le projet que nous avons mené se structure en plusieurs phases (choix de la problématique à aborder conjointement, choix de la méthodologie à adopter, élaboration du questionnaire, passation du questionnaire, traitement des données, diffusion des analyses, suites données à l’enquête). Avez-vous des remarques critiques à formuler quant à l’une ou l’autre de ces phases ?

9. Si c’était à refaire ?

EVALUER AUTREMENT LA CAPACITE D'ANALYSER SA PRATIQUE D'ENSEIGNANT AU SECONDAIRE SUPERIEUR SOUS L'ANGLE DE LA PSYCHOLOGIE EDUCATIONNELLE

Delcomminette Séverine, Leclercq Dieudonné, & Sente Isabelle Université de Liège - Belgique

Mots-clés : Transfert, analyse, enseignant

Résumé :
Le cours de psychologie éducationnelle à l'agrégation dispensé au sein de l'Université de Liège s'adresse à des futurs enseignants du secondaire supérieur. En novembre 2006, nous avons renversé notre processus traditionnel d'évaluation certificative pratiqué depuis 20 ans. Au lieu d'inviter les étudiants à répondre à des questions à choix multiple (QCM) sur des cas créés par nous sur base de nos seules représentations personnelles de la vie en classe, nous leur avons demandé de fournir des exemples dans leur domaine (chimie, langues germaniques, sciences sanitaires, … ; 20 domaines au total). Ce renversement de la perspective est comparable au passage du thème (traduire vers la langue étrangère) à la version (traduire vers sa langue). Dans cet article nous analysons l'aspect docimologique, formatif ainsi que l'intérêt perçu par les étudiants et par les encadrants de cette nouvelle formule d'évaluation.

---

1. Le contexte de l'expérience

En Belgique francophone, les enseignants du secondaire supérieur bénéficient d'une formation initiale en deux phases. Dans la première phase, ils acquièrent un diplôme universitaire de maîtrise nécessitant 4 à 5 années d'études selon les spécialités (chimie, langues germaniques, économie, etc.). Dans une seconde phase, les étudiants doivent acquérir un diplôme appelé « Agrégation de l'Enseignement Secondaire Supérieur ». Dans ce cadre, à l'Université de Liège, en plus de stages et de leçons publiques, existent des cours tels que méthodologie générale, didactique spécialisée, connaissance de l'institution scolaire, pédagogie interculturelle, médias, et « Psychologie éducationnelle de l'adolescent et du jeune adulte ». Ce cours est actuellement réduit à 15h (il en a comporté 30 par le passé), soit seulement 7 à 8 réunions de 2 heures dans un amphithéâtre de 300 personnes. Il est donné depuis vingt ans par le professeur D. Leclercq.

Bien que la présente communication soit centrée sur l'évaluation, on se doute que les objectifs sont déterminants ainsi que les méthodes, à la fois sur l'objet de l'évaluation et sur sa forme. Nous rejoignons là les deux préoccupations d'Ebel (1982) dans son article « Solutions proposées à deux problèmes concernant la construction des tests ». Ces deux problèmes sont le **quoi** (mesurer) et le **comment** (le

mesurer). Nous reviendrons à ses réponses pour les comparer aux nôtres (dans la conclusion).

2. Le contenu du cours

En vue de rendre les étudiants capables, en situation réelle, d'interpréter et de concevoir des situations scolaires d'apprentissage et de communication interpersonnelle sous les angles de la psychologie de l'éducation, le contenu du cours est exposé dans un livre et réparti en 8 chapitres :

1. L'équipement de base (cerveau, synapses, épigénèse, attrition, perceptrons, intelligence des animaux et des machines, débat inné / acquis, etc.),
2. Mémorisation et rappel (MCT et MLT, chunking, moyens mnémotechniques, rôle du sommeil, etc.),
3. **A**ffectivité (besoins, motivation),
4. **S**kills ou Automatismes (chaînes Sensori-motrices, automatismes intellectuels, exercisation, guidage, enseignement programmé, etc.),
5. **C**ognition, concepts et réseaux conceptuels (schémas, scripts, irradiation, *advance organizers*, apprentissage verbal significatif, etc.),
6. **I**mage de soi et métacognition (estime de soi, sentiment d'auto-efficacité, attributions causales, dissonance cognitive et sa réduction, auto-évaluation de ses connaissances et de sa compétence, réalisme, analyse et régulation, position épistémologique personnelle, etc.),
7. **D**écisions et engagement (degré d'implication dans diverses activités),
8. Adolescents et adultes (Relation adolescents – parents et adolescents – enseignants, etc.).

Le tableau 1 met ces contenus (colonnes « chapitres ») en relation avec les objectifs du décret (colonne 1) définissant la formation des enseignants sélectionnés par le professeur. Idem pour les méthodes (Leclercq et Poumay, 2005) mises en place. Idem aussi pour l'évaluation (colonnes « Evaluation ») à laquelle se soumettent les étudiants.

Tableau 1 – Situation de départ en termes de concordance entre objectifs-méthodes et évaluation.

Objectifs		**Méthodes (8 évènements d'apprentissage)**							**Chapitres (contenu)**							**Eval.**	
Du professeur	*Du décret définissant la formation initiale des agrégés de l'enseignement secondaire supérieur (M.B. 22-02-2001 modification du M.B. 31-01-02)*	Observation	Réception	Exercisation	Exploration	Expérimentation	Création	Débat	Méta	1 (Equipement de base)	2 (Mémorisation)	3 (Affectivité)	4 (Savoir-faire)	5 (Concepts et réseaux)	6 (Image de soi)	7 (Décision)	8 (Adolescent et rapport à autrui)	Livre fermé	Livre ouvert
Comprendre les processus d'apprentissage, de mémorisation, de "gestion" du succès et de l'échec, de motivation, de décision, d'auto-évaluation	Mobiliser des connaissances en sciences humaines pour une juste **interprétatio**n des situations vécues en classe et autour de la classe et pour une meilleure adaptation aux publics scolaires																		
Saisir leurs implications pour l'éducation, la relation professeur-élèves et les stratégies de formation et d'évaluation	Mesurer les enjeux éthiques liés à sa pratique quotidienne																		
Analyser des situations concrètes en termes de théories exposées	Mobiliser des connaissances en sciences humaines pour une juste interprétation des situations vécues en classe et autour de la classe et pour une meilleure adaptation aux publics scolaires																		
Comprendre les enjeux et mécanismes psychologiques liés à l'adolescence et se situer face à ceux-ci	Développer les **compétences** relationnelles liées aux exigences de la profession																		
Vivre des expériences (d'analyse de cas, d'apprentissage et d'évaluation) et se situer vis-à-vis d'elles	Planifier, gérer et évaluer des situations d'apprentissage Porter un regard réflexif sur sa pratique et organiser sa formation continuée																		

Au survol du tableau 1, on observe dans situation dite de départ des failles (absence de X pour la ligne « Analyser des situations concrètes en termes de théories exposées »). Certains objectifs ne sont pas alimentés par le syllabus, ou ne sont pas entraînés ou encore ne sont pas mesurés (évalués). Nous détaillerons ci-après les éléments – objectifs, méthodes et évaluation.

3. Les objectifs

3.1 Au niveau épistémologique

Les (futurs) enseignants de l'enseignement secondaire général sont des universitaires qui devraient, vis-à-vis de la pédagogie avoir les mêmes critères de jugement que vis-à-vis de leur discipline, à savoir la vérifiabilité des sources, des exigences méthodologiques quant au statut des propositions qui sont énoncées : hypothèses ou données d'observation ? Ils devraient atteindre le plus haut niveau épistémologique possible, notamment se voir comme constructeurs de connaissance. Ils ne doivent pas seulement se voir comme des « solutionneurs de problèmes », mais aussi comme des penseurs des problèmes et de leur solution.

Nous insistons sur cette distinction en rappelant la distinction que Piaget faisait entre l'intelligence (résoudre des problèmes) et la pensée. Pour Piaget, les animaux sont intelligents, mais il n'existe pas de pensée avant le langage. Cette position aura toute son importance au moment de l'évaluation qui fera largement appel au langage. Nous ne testerons pas et n'entraînerons pas seulement l'intelligence des situations, mais aussi la pensée sur les situations.

Les enseignants sont (doivent être) des spécialistes de la pensée. Non seulement de la leur, mais de celle des autres, et notamment des experts. Nous tâcherons de montrer comment nous nous efforçons de faire en sorte que ce cours contribue à cet objectif, poursuivi par d'autres cours, évidemment.

3.2 Au niveau cognitif

Nous tentons de couvrir quasi tous les niveaux de processus mentaux décrits dans la taxonomie de Bloom (1956) des objectifs du domaine cognitif. Nous insistons sur « processus mentaux » parce que, pour nous, connaissance, compréhension, application, analyse, etc. ne sont que des processus mentaux et ne deviennent des objectifs que quand on leur donne des compléments d'objet (la matière à connaître, comprendre, appliquer, etc.). Nous montrerons, dans la rubrique « évaluation » comment nous mesurons ces divers niveaux cognitifs et dans la rubrique « méthode » comment nous y entraînons nos étudiants.

Le défi majeur, parmi ces objectifs cognitifs, est pour nous de rendre les étudiants capables de faire les liens entre théorie (celle du cours) et pratique, donner du sens à leurs apprentissages théoriques en termes d'utilité pour leur métier. Et, sur ce point, nous étions déçus de la situation de départ (cf. le tableau 1). Cette communication présente les efforts que nous avons faits en ce sens.

3.3 Au niveau métacognitif

Tout au long de l'apprentissage, nous soumettons nos étudiants à des évaluations formatives recourant aux degrés de certitude (Leclercq, 1982, 1993, 2003). Ceux-ci sont exploités pour entraîner les étudiants à l'auto-estimation et à la régulation de leurs apprentissages.

4. Les méthodes

Rappelons le défi de toujours pour ce cours en grand groupe : faciliter pour nos (300) étudiants le lien entre la théorie et la pratique, assurer chez eux la perception du sens du cours, tant dans la phase d'apprentissage que dans la phase d'évaluation.

Etant donné la contrainte horaire (15h), le titulaire a, jusqu'à présent, recouru aux méthodes suivantes :

- L'exposé magistral, supporté par des diapositives PowerPoint dont les étudiants ont la copie papier dès le départ (ce qui leur permet de ne prendre note que de leurs réflexions (sur la moitié droite de la feuille). Cet exposé est interrompu par des démonstrations, questions et illustrations, quelques fois par les apprenants eux-mêmes,
- La séquence LQRT : **L**ecture du chapitre X à domicile par les étudiants ; **Q**uestions par mail et au plus tard lors de la séance par les étudiants ; **R**éponse immédiate par le professeur à ces questions et **T**est (annoncé et formatif) clôturant l'échange. (Leclercq, 1998, 165-185),
- La simulation d'examens (un test de 10 QCM en fin de chaque séance avec correction immédiate sur questionnaire spectral). (Leclercq et al., 2003),
- L'analyse de cas (sur l'adolescence et le métier de professeur) apportés par les étudiants.

5. Notre évaluation traditionnelle en deux volets

L'évaluation est, elle aussi est très marquée par les contraintes : plus de 300 étudiants sont inscrits chaque année. Après avoir, pendant plusieurs années, procédé par examens oraux, puis par travaux personnels, le titulaire du cours a fini, depuis une quinzaine d'années, par recourir à un examen écrit en deux volets, corrigé par ordinateur.

L'examen à Livres Fermés (LF) recourt aux Questions à Choix Large ou QCL (Leclercq, 2005, chap. 4, 8-10). Dans cette technique de questionnement, les (nombreux) trous de textes doivent être complétés par un des mots d'une liste de plusieurs centaines (500 par exemple) de termes rangés par ordre alphabétique sur 8 colonnes d'une feuille A4 recto-verso. Cette technique est intermédiaire entre les QCM (quelques solutions proposées) et les Questions à Réponse Ouverte Courte (aucune solution proposée) : l'étudiant doit d'abord évoquer la réponse, puis chercher son code numérique (3 chiffres) dans la liste alphabétique et reporter ce code sur le formulaire destiné à la lecture optique. Les questions sont de pure restitution de mémoire. Nous estimons en effet qu'un minimum de savoirs sont nécessaires pour pouvoir participer à des débats, lire des livres ou tout simplement penser dans des termes lourds de signification et appartenant à des réseaux conceptuels théoriques, car pour penser, comme le disait Kurt Lewin, « Rien n'est plus pratique qu'une bonne théorie ».

L'examen à Livres Ouverts (LO) recourt aux QCM avec Solutions Générales Implicites : solutions « Aucune, Toutes, Manque, Absurdité » (Leclercq, 1986, 127-144). Ces solutions, ajoutées aux QCM sont générales dans ce sens qu'elles ont valables pour toutes les questions du test et implicites parce qu'elles ne sont

annoncées qu'une fois (en début de test) et ne sont pas répétées dans chaque question. Les questions portent sur des principes généraux, décrits par la littérature scientifique et leurs liens avec des exemples imaginés par le professeur.

Ce volet d'examen vise à mesurer 3 des niveaux de la taxonomie des objectifs cognitifs de Bloom : « compréhension » des théories, « application » sur des exemples et « analyse » – vigilance, ce dernier niveau étant représenté par les questions dont la solution correcte est soit « Manque de données », soit « Absurdité dans l'énoncé ».

Les Solutions Générales Implicites (SGI) permettent d'éviter le piège d'un des plus gros défauts des QCM classiques, à savoir de porter sur des détails parce que, comme l'expliquent Noizet et Caverni (1978, 155), elles portent sur des questions « décidables » qui ne laissent pas place à l'opinion : « Cette condition de décidabilité est capitale puisque la réponse sollicitée est de nature « vrai » ou « faux ». Ainsi par exemple, la question de savoir si tel écrivain a écrit ou non telle œuvre constitue un item décidable [...]...Mais la question de savoir si [...] tel écrivain peut être considéré comme un romantique constituerait un item indécidable ».

Les **Degrés de Certitude** (Leclercq, 1982) appliqués aux examens LF et LO sont à choisir sur une échelle à 6 niveaux : 0%, 20%, 40%, 60%, 80%, 100% (Leclercq, 2003, 30). Afin de permettre le calcul immédiat des indices de Confiance et d'Imprudence (voir bas de l'annexe ci-après), et de favoriser l'identification des lacunes, nous recourons à un formulaire spectral (Leclercq et al, 2003 ; voir un exemple en annexe). Les degrés de certitude permettent de mesurer le réalisme dans l'auto-évaluation et le processus « évaluation – jugement », dernier « étage » de la taxonomie de Bloom.

Ces deux volets « traditionnels » ne permettent pas de mesurer l'ensemble des objectifs du cours, aussi avons-nous tenté de pallier cette lacune.

6. Trois innovations majeures entre 2004 et 2006

Poser des questions sur des cas permet, selon nous en référence au tableau 1, de mesurer les objectifs :

- Analyser des situations concrètes (contextualisation des contenus) en termes de théories exposées,
- Vivre des expériences d'analyse de cas, d'apprentissage et d'évaluation et se situer vis-à-vis d'elles.

En 2004, les cas ont été construits **par le professeur seul**. Dans le souci de rendre l'évaluation contextualisée, nous sommes partis de cas (6 situations par examen imaginées par le titulaire du cours selon ses propres représentations), posant 5 questions sur un même cas. C'est cette modalité de questionnement qui est illustrée dans l'annexe (le 4ème cas, donc avec les questions 16 et 17).

La procédure de feedbacks et discussion oraux et collectifs a été appliquée à, propos du questionnaire proposant les cas. Nous avions commencé ce debriefing à 21 heures ; il a duré jusque 23 heures. Les cas, plus ils sont concrets, plus ils sont susceptibles d'interprétations multiples et les nuances (présentes ou manquantes) n'en finissent pas de rendre de nombreuses réponses très différentes potentiellement

correctes. Une très bonne arme en évaluation formative devient une arme à double tranchant en évaluation certificative.

En 2005, les cas ont été rédigés **par le professeur et une étudiante monitrice** et les étudiants volontaires (198 sur les 300) ont été testés par ordinateur. Nous avons posé nos questions en deux temps :

1. Quelle est l'opinion de l'étudiant en regard de telle partie du cas exposé ? (Question à Réponse Ouverte ou QRO).
2. Quelle est la position des chercheurs sur ce même aspect ? QCM avec SGI et degré de certitude, ainsi que la possibilité de justifier son choix.

En cas de réponse incorrecte, nous avons lu les opinions et les justifications. Si ces dernières s'avéraient pertinentes, l'étudiant obtenait le point de bonne réponse. Ce dispositif nous a permis d'évaluer des processus mentaux subtils à partir de cas complexes se déroulant dans l'enseignement secondaire supérieur (Delcomminette & Leclercq, 2006). Cependant ces cas avaient deux inconvénients : ils avaient été rédigés par nous et ne couvraient pas tous les contenus dans chaque discipline.

En 2006, des cas réels sont relatés **par les étudiants**. Ce faisant, nous avons voulu poursuivre deux nouveaux objectifs. Le premier objectif est au bénéfice des formateurs : disposer d'exemples concrets dans des disciplines variées, toutes celles des étudiants inscrits au cours. Ces cas réels devraient être exploités dans les notes de cours et le dispositif de l'année académique suivante et ainsi constituer un pas supplémentaire en termes de liens théorie-pratique. Le second objectif est au bénéfice des étudiants : indiquer à tous que l'expérience des étudiants est intéressante et aussi valable que celle du professeur, la référence restant la théorie présentée lors du cours. C'est pourquoi nous avons invité les étudiants qui le voulaient à **remplacer l'examen** en produisant des exemples pertinents (si possible vécus) en regard des contenus théoriques mais dans **leur** discipline. Ils ont reçu 8 feuilles A4, chacune dédicacée à un des 8 thèmes suivants : (**A**) Affectivité – Attitudes (A1 à A7) ; (**S**) Savoir-Faire Sensori-moteurs Skills (S1 à S7) ; etc.

Dans chaque thème, 7 contenus correspondant au découpage du chapitre en 7 blocs les plus importants, soit 56 points de matière.

Les étudiants ont dû fournir au moins 2 exemples **jugés pertinents par les correcteurs** pour chaque thème et au total, ils devaient avoir fourni 20 exemples jugés pertinents accompagnés chacun d'un des 6 Degrés de Certitude.

Sur 300 étudiants inscrits, 130 étudiants se sont prêtés au jeu. La très grande majorité a réussi (94,6% de réussite). Chaque réponse était notée selon 3 critères : (**1**) Pertinence du contenu ; (**2**) Précision – intelligibilité pour le lecteur ; (**3**) Respect du niveau demandé (secondaire supérieur).

Il s'agit ici d'une innovation en termes de la taxonomie de Bloom (1956) : notre évaluation traditionnelle mesurait la connaissance et la compréhension, alors que notre épreuve de production, elle, mesure l'analyse (résolution de problème) et la synthèse (formulation, création).

Voici un exemple de réponses d'une étudiante dans la première des 56 cellules :

A (Affectivité – Attitudes)

		Avt /Ap	DC	Pertinence Contenu	Précision	Niveau
A1. Désirs ou répulsions à promouvoir et comment	Mon professeur de néerlandais m'a fait découvrir ma passion pour les langues, et principalement pour le néerlandais. Ce prof de secondaire m'a vraiment donné le goût des langues. Il m'a transmis sa passion, son désir de communiquer en langues étrangères.	Avt	80 %			
	Lorsque j'étais en rhétorique, alors que j'avais choisi l'option math, j'ai été dégoûtée des mathématiques par mon professeur. Sa sévérité et son élitisme à former toute la classe à l'examen d'entrée aux ingénieurs m'ont totalement « écoeurée ». Je ne pense pas un jour renouer avec les mathématiques: dégoût total et à vie ! Dommage...	Avt	80 %			

Figure 1 - exemple de réponses d'une étudiante en langues germaniques.

On constate l'importance de l'innovation. C'est comme si l'on était passé du thème (exprimer dans les termes du professeur ce qui est un cas exprimé dans le langage habituel) à la version (donner ses exemples à soi de concepts exprimés dans la langue du professeur).

Dans une perspective d'amélioration de la consigne aux étudiants et de nos critères de correction, nous pourrions recourir aux concepts de Torrance (1965) :

1. La **fluidité** ou nombre d'éléments produits (dans l'exemple ci-dessous on constatera que pour A2, cet étudiant a fourni 2 exemples) Le minimum observé est 20 (le minimum exigé) et certains sont allés jusqu'à plus de 100 (sur les 56 cellules). Cette année nous n'en avons pas tenu compte dans la notation, nous pourrions envisager de le faire plus tard.
2. La **flexibilité** ou le nombre d'éléments produits de catégories différentes. Nous avons pu remarquer des différences entre étudiants, mais notre consigne n'étant pas assez précise à ce sujet (du genre « fournissez les exemples les plus variés possible »), nous n'en avons pas tenu compte non plus dans la notation.
3. L'**originalité** ou rareté (par rapport au groupe ; ce qui en fait un concept normatif). La norme de référence pourrait être sur certains points (exemple la discipline) l'ensemble des situations pédagogiques et pour d'autres (exemple la sécurité dans les laboratoires) propre à un domaine particulier.
4. L'**élaboration**, sophistication ou degré de détail, de précision. Nous en avons fait un critère.

Nous avons annoncé au début de cette section nos objectifs. Il nous semble qu'ils ont été atteints du point de vue des formateurs, mais qu'en est-il chez les étudiants ? Avant de rentrer dans une analyse des résultats, nous explicitons au point suivant nos hypothèses sous-jacentes à l'objectif poursuivi au niveau des étudiants.

7. Nos hypothèses motivationnelles

Nous avons souhaité que notre évaluation ait les caractéristiques suivantes, en référence au modèle de Viau (1994).

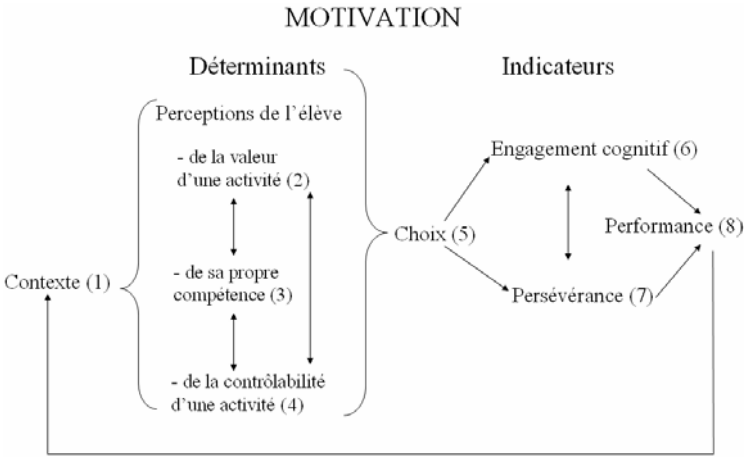


Figure 2 – Modèle de Viau (1994).

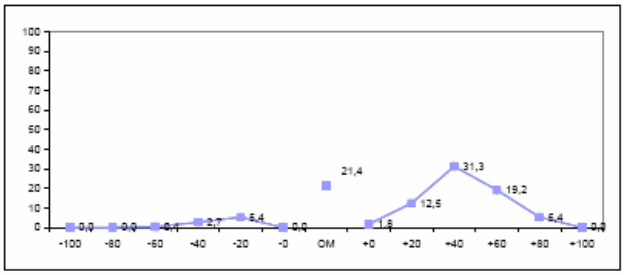
1. Contexte : Au lieu d'évaluer les étudiants de façon convergente (Guilford, 1967) en leur demandant de répondre à des QCM par notre réponse correcte, nous avons voulu réaliser une évaluation divergente.
2. Cette consigne devrait valoriser à ses propres yeux l'expertise qu'a chaque étudiant et que nous n'avons pas. Comme d'ailleurs le fait d'appliquer les théories à sa pratique.
3. Le contrôle de l'activité lui incombait partiellement : il pouvait choisir d'illustrer 20 points matière sur les 56 possibilités.
4. L'image de sa propre compétence devrait être positive vu le taux élevé de Réponses Correctes (RC). En outre, nous faisons l'hypothèse que cette analyse l'a normalement aidé à comprendre certaines situations, prendre du recul et être rassuré sur sa capacité à mieux gérer ces mêmes problèmes ultérieurement.
5. Le choix de présenter cette épreuve dispensatoire a été fait par 130 étudiants sur base de ces perceptions.
6. Ces 130 étudiants se sont pleinement engagés cognitivement : nous avons observé que les travaux étaient très fouillés et rédigés avec sérieux (pertinents en regard de nos 3 critères de correction).
7. Chez les 130 étudiants, la persévération a été élevée : la moyenne d'exemples a été beaucoup plus importante que les 20 « planchers » minimums.
8. Les performances se sont avérées proches de l'excellence, avec une moyenne des scores classiques proche de 15/20 et, « avec bonus métacognitif », de 17/20.

Nous ferons un dernier détour par le feedback que nous avons rendu à chaque étudiant sur sa performance avant de nous plonger dans les résultats observés

8. Le feedback aux étudiants

Un feedback individuel personnalisé leur a été envoyé par mail afin d'alimenter leur développement personnel en leur indiquant quels contenus sont bien maîtrisés et lesquels ne le sont pas. Voici le 1er bloc de feedback pour un étudiant. Il y en a 8 en tout. Chaque étudiant a aussi reçu deux graphiques personnels (l'un représente la distribution spectrale (Leclercq, 2003) sur l'ensemble du travail, l'autre sur les 20 meilleurs exemples), nous vous montrons ici celui relatif à l'ensemble des réponses soumises.

	Avt/Ap	DC	Pertinence Cor	Précision	Niveau	position spectrale de la performance sur l'échelle de qualité			Total /4
Nom :									
Prénom :									
AE88									
Section : Sciences économiques									
Discipline									
Poids			1	1	1	2	1	1	
A1. Désirs ou répulsions à promouvoir et comment	Avt	80	1	1	1	60	60	60	3
A2. Classables selon la pyramide des besoins de Maslow, ou criticables comme le fait Albertini	Ap	40	2	1	1	40	40	40	4
A3. Ex de mesures (hit parade), avec questionnaire chargé (si oui, comment ?)	Ap	20	0	0	0	-20	-20	-20	0
A4. Réflexes conditionnels	Avt	40	1	1	1	40	40	40	3
A5. Stress et de la réduction par information procédurale et/ou sensorielle	Avt	80	2	1	1	60	60	60	4
A6. Lavage de cerveau ou procédure systématique de " réorientation des réflexes "	Avt	20	0	0	0	-20	-20	-20	0
A7. Classification du style de celle de Krathwohl et al.						OM	OM	OM	0



Sur l'ensemble

	-100	-80	-60	-40	-20	-0	OM	+0	+20	+40	+60	+80	+100	
Pertinence	0	0	0	1	3	0	12	1	7	18	11	3	0	56
Précision	0	0	1	3	3	0	12	1	7	16	10	3	0	56
Niveau	0	0	0	1	3	0	12	1	7	18	11	3	0	56
Total	0	0	0,45	2,68	5,36	0	21,4	1,79	12,5	31,3	19,2	5,36	0	100

Figure 3 – Exemple de feedback envoyé à un étudiant

Le graphique du haut représente la distribution spectrale (Leclercq, 2003) sur l'ensemble du travail, celui du bas sur les 20 meilleurs exemples. Ce dernier est meilleur de par le déplacement de la courbe droite ; leur fluidité les a donc renforcés positivement.

9. Résultats

9.1 Scores élevés obtenus par les étudiants

Nous attribuons ces notes plus élevées que les notes habituelles à plusieurs facteurs :

1. Nous avions sous-estimé les étudiants et leur capacité à fournir des exemples pertinents,
2. Les étudiants se sont investis beaucoup plus que le minimum exigé,

3. La consigne « 20 correctes sur 56 cellules » permet à chaque étudiant de ne s'attaquer qu'aux contenus où il se sent fort.

9.2 Informations permettant aux formateurs d'améliorer le cours

Tableau 2 – Les sujets par thème qui ont le plus grand nombre de réponses correctes (RC) dans chaque chapitre.

Thème	Sujet	Nb de RC[1]
Affectivité	A5. Stress et de la réduction par information procédurale et/ou sensorielle	111
Skills	S4. Soufflage dégressif	112
Cognition	C1. Scripts (séquences d'actions) : leur connaissance théorique	97
Réseaux conceptuels	RC1. Carte conceptuelle (exemple, son utilité et son utilisation)	104
Image de soi	I6. Attribution causale	99
Adolescence	Ado3. Influence des pairs	**120**
Métier d'enseignant	ME2. Exigences et ambiance de la classe & ME3. Le travail de sourcier	117

Tableau 3 – Les sujets par thème qui ont le plus grand nombre d'omissions dans chaque chapitre.

Thème	Sujet	Nb d'OM
Affectivité	A3. Ex de mesure (hit parade) avec questionnaire chargé (si oui, comment ?)	76
Skills	S7. Tricotage conceptuel autre que l'enseignement programmé	61
Cognition	C6. Image mentale présentes (ex : le cube) et manquantes (ex : la 4ème puissance)	31
Réseaux conceptuels	RC7. Connotations (par l'image ou par des slogans ou des gestes)	52
Image de soi	I5. Confiance, prudence, nuance – Distribution spectrale de la qualité d'une performance	**83**
Adolescence	Ado6. Le développement moral	53
Métier d'enseignant	ME6. Le travail d'architecte	40

Tableau 4 – Les sujets par thème qui ont le plus grand nombre de réponses incorrectes (RI) dans chaque chapitre

Thème	Sujet	Nb de RI[2]
Affectivité	A4. Réflexes conditionnels	41
Skills	S6. Progression vers la spontanéité – enlèvement des aides	17
Cognition	C1. Scripts (séquences d'actions) : leur connaissance théorique	16
Réseaux conceptuels	RC2 & RC3. Caractère multimédia de la nature d'un réseau : MOTS & IMAGES	8
Image de soi	I7. Dissonance cognitive	21
Adolescence	Ado1. Déséquilibres physique, intellectuel et social	13
Métier d'enseignant	ME7. Le travail de choriste	**56**

[1] Pertinence notée 2 ou 1

[2] Pertinence notée 0

C'est le point relatif à la distribution spectrale (image de soi) qui a été le moins illustré (91 omissions). Nous savons que c'est une « innovation maison » et que, forcément, hormis nos propres cours, les étudiants ne l'ont jamais vécu.

10. Avis des étudiants sur cette forme d'évaluation

Sur 130 étudiants, 55 nous ont renvoyé le questionnaire. Les catégories de réponses ne sont pas exclusives, les étudiants ont parfois évoqué plusieurs raisons

Question 1 *: Y a-t-il un intérêt à mettre des théories et des noms sur des choses vécues (bref à théoriser la pratique), ce que fait le cours de Psychologie Educationnelle ?* **Oui = 50 ; non = 5**. Si oui, lequel ?

Réponses : Analyse du vécu personnel / Mettre des noms sur les choses vécues (22), parler tous le même langage (7), bagage théorique et pratique pour le futur (6), comprendre le bien-fondé de certaines pratiques (5), prendre une distance par rapport à la pratique (5), se rassurer, relativiser son vécu (« cela n'arrive pas qu'à moi ! ») (5), mieux cibler la théorie grâce aux exemples (4), …

Question 2 : *Cet exercice de transfert / d'analyse de votre pratique à la lumière des théories vues au cours a-t-il été utile ?* **Oui = 55.** Si oui, à quoi ?

Réponses : Intégrer la théorie et mieux la comprendre (13), comprendre vraiment au lieu de mémoriser « bêtement » (10), relier théorie et pratique (6), approfondir la matière et y réfléchir en fonction de sa discipline (5), se poser des questions sur la pratique future (enseigner ≠ simplement donner cours) (5), être dispensé d'examen (4), mettre des mots sur les difficultés rencontrées (4), …

Question 3 *: Quelles sont les difficultés que vous avez rencontrées dans le cadre de cet exercice ?*

Réponses : Trouver des exemples et juger de leur pertinence (23), manque d'expérience sur le terrain (13), manque de clarté de certaines notions abordées (pas assez d'explications) (8), cibler les exemples pertinents (6), manque de clarté des items pour lesquels il fallait donner des exemples (5), manque de clarté des consignes (4), …

11. Conclusions et perspectives

Nous pensons avoir progressé en termes de lien théorie-pratique. Nous poursuivons le travail en injectant certains de ces exemples dans le syllabus papier et certains autres dans un testing par cas (Delcomminette & Leclercq, 2006). Les contenus devront être validés par des didacticiens responsables de chaque discipline.

Toutes les formes d'évaluation proposées ci-dessus sont restées verbales (des mots pour les questions, des mots pour les réponses). La **première raison** relève du **quoi** (ce que nous voulons tester) : ce n'est pas seulement la réponse, mais surtout la justification de cette réponse, et même la connaissance qui fonde cette justification. Ici, nous nous rangeons aux arguments d'Ebel (1982, 268) : « *Thinking is a process, knowledge is a product..[...] thinking always involves knowledge"* et à ceux de Hirst (1972, 397-399) : "*To say that schools should teach pupils how to think instead of teaching them knowledge is to urge the impossible*."

La position sage d'Ebel (1982, 269-270) est que "*Il est possible de poser des questions de test qui sont très exigeantes quant à la capacité de l'étudiant de penser, c'est-à-dire de traiter de l'information, et assez peu d'exigences quant à sa capacité d'accumuler de l'information…[...] Bien sûr, savoir n'est pas la même chose que faire. Si le faire implique une manipulation physique cela peut exiger des habiletés sensori-motrices que* ***connaître*** *ne peut donner. Même dans le domaine des tâches purement mentales, la pratique peut augmenter la faciliter à réaliser la tâche. Mais, à part la facilité, réaliser une tâche mentale requiert-il autre chose que* ***savoir*** *parfaitement bien comment la faire ? [...] Mais ces connaissances nécessaires pour réaliser ces tâches cognitives sont-elles toutes de la connaissance verbale ? [...] Probablement pas si les impressions tacites, inarticulées sont considérées comme une forme subjective de connaissance personnelle (Polanyi, 1958). Néanmoins, en enseignant et en testant, nous devons traduire la connaissance implicite en une connaissance articulée, objective et publique. »*

La **deuxième raison** porte sur le **comment**. Nous n'envisageons pas, pour le moment, de remplacer cet examen par un examen simulé qui n'exigerait pas d'explication verbale fournie par l'apprenant, et où celui-ci serait jugé sur sa seule conduite. Comme cela pourrait être le cas en tests par « micro enseignement » s'inspirant des principes des ECOS (Examens Cliniques Objectifs Structurés) de Harden et al. (1975).

Notre conception du lien théorie – pratique, si elle vise à entraîner à résoudre des problèmes, distingue cependant, à la manière de Piaget (1974), « réussir » et « comprendre » et tente de mettre les futurs enseignants en position de « comprendre avant d'agir ».

Références

Bloom, B.S., Hastings, J.T., & Madaus, G. (1969). *Taxonomie des objectifs pédagogiques : tome 1 domaine cognitif.* Montréal : Education Nouvelle. (Version anglaise : 1956).

Delcomminette, S., & Leclercq, D. (2006). Transférer de la théorie à la pratique via un testing par cas en formation d'enseignants (TICAFE). *Actes du XXIIIème Congrès de l'Association Internationale de Pédagogie Universitaire*.

Ebel, R., L. (1982). Proposed solutions to two problems of test construction. *Journal of Educational measurement*, 19, 4, 267-278.

Guilford, J.P. (1967). *The nature of human intelligence*. New-York: Mc Graw-Hill.

Harden, R., Stevenson, M., Downie, W., & Wilson, G. (1975). Assessment of clinical competence using objective structured examination (O.S.C.E.). *Medical Education*, 13, 41-54.

Hirst, P.H.(1972). Liberal education and the nature of knowledge. In R.F. Dearden, P.H. Hirst & R.S peters, *Education and the development of reason.* London : Routledge and Keagan Paul.

Kuhn, D. (1991). The skills of argument. Cambridge, England : Cambridge University Press.

Leclercq, D. (1982). Confidence Marking, its use in testing, in Postlethwaite & Choppin (Eds), *Evaluation in education: An international review series*. (6-2, 161-287). Oxford: Pergamon.

Leclercq, D. (1986). *La conception des QCM*. Bruxelles : Labor.

Leclercq, D. (Ed) (1998). *Pour une pédagogie universitaire de qualité*. Sprimont : Mardaga.

Leclercq, D. (Ed). (2003). *Diagnostic cognitif et métacognitif au seuil de l'université*. Liège : Editions de l'Université de Liège.

Leclercq, D. (2005). *Edumétrie et docimologie pour praticiens-chercheurs*. Liège : Editions de l'Université de Liège.

Leclercq, D & Bruno, J. (1993). *Item banking, self assessment and interactive testing*. NATO ARW Conference. Berlin : Springer Verlag.

Leclercq, D., Rinaldi, A.M., & Ernould, C. (2003). Un questionnaire spectral pour l'évaluation des connaissances chez le patient diabétique, in Gagnayre et al. (Eds). *L'évaluation de l'éducation thérapeutique du patient*. Paris : IPCEM.

Leclercq, D. & Poumay, M. (2005). *The 8 learning Events Model*. Release 2005.1. LabSET, University of Liège. Available: http://www.labset.net/media/prod/8LEM.pdf

Noizet, G. & Caverni, J.P.(1978). *Psychologie de l'évaluation scolaire*. Paris : PUF.

Piaget, J. (1974). *Réussir et comprendre*. Paris : Presses universitaires de France.

Polanyi, M. (1958). *Personal knowledge*. Chicago : University of Chicago Press.

Torrance, E.P. (1965). *Rewarding creative beahavior*. Englewood Cliffs : Practice Hall.

Viau, R. (1994). *La motivation en contexte scolaire*. Bruxelles : De Boeck.

Annexe

Tableau 2 – Exemple de formulaire spectral.

Réponses incorrectes				Réponses correctes
			La rencontre prof-parents Les parents rencontrent individuellement les professeurs une fois par an. C'est l'occasion pour certains d'entre eux de raconter à l'enseignant ce qu'ils considèrent comme problèmes de leur adolescent. Dans ces dialogues, le professeur doit garder à l'esprit un certain nombre de tendances, soit pour rassurer les parents, soit pour les alerter.	
00 0 0 0 0	**6**		L'évolution physique qui angoisse le plus grand nombre d'adolescents est 1. l'augmentation de leur taille (en hauteur) 2. l'augmentation de leur force (puissance musculaire) 3. la transformation de leur aspect (look), spécialement du visage 4. la transformation de leur capacité intellectuelle (plus abstraite)	0 0 0 0 00
00 0 0 0 0	**7**		Les oppositions aux opinions des parents 1. sont une nécessité pour affirmer l'indépendance de l'adolescent 2. peuvent être un jeu dialectique voulu pour tester les limites des raisonnements des uns et des autres 3. peuvent être pour l'adolescent(e) d'autant plus déculpabilisantes qu'il (elle) sera persuadé(e) que les parents sont nuls. 4. sont utiles quand elles contribuent à créer la pensée du futur	0 0 0 0 00

Réponses Correctes x 1 = Réponses Incorrectes x -0,25 = Score classique (sur 20) =.....

Confiance =……% Imprudence =…… %

Plus Métacognitifs : Confiance =…+ .point Prudence =… + .point Score total (sur 20) = ...……

CONSTRUCTION ET UTILISATION D'UN OUTIL D'ANALYSE DES RESULTATS A MI-PARCOURS POUR PROMOUVOIR LA REFLEXION METACOGNITIVE DE L'ETUDIANT DIAGNOSTIQUE EN DIFFICULTE EN PREMIERE ANNEE DE L'ENSEIGNEMENT SUPERIEUR

Didier SALMON, Unité de Physique des Biosystèmes,
Faculté universitaire des Sciences agronomiques de Gembloux

Mots-clés : accompagnement pédagogique, pronostic de réussite, évaluation formative.

Résumé :

A la Faculté universitaire des Sciences agronomiques de Gembloux (FUSAGx) une série complète d'interrogations sont organisés à mi-parcours (janvier). Un des objectifs de ces évaluations est de permettre à l'étudiant de mesurer son adaptation aux normes et aux exigences d'une évaluation universitaire, mais aussi de réfléchir aux démarches d'apprentissage utilisées pendant la première partie de l'année académique. De plus, depuis trois ans, une série de tests est organisée dès le milieu du premier quadrimestre (novembre) afin de susciter une réflexion métacognitive encore plus précoce.

Dans la première partie de la communication, nous présenterons la démarche de construction d'un outil d'information permettant aux étudiants d'analyser, lors d'un entretien individuel, les résultats obtenus aux interrogations à mi-parcours. Les informations fournies par l'outil sont, d'une part, l'appartenance à une des trois populations types définies à partir d'un pronostic de réussite/échec et, d'autre part, un intervalle de confiance de la performance académique finale attendue.

La deuxième partie de l'exposé permettra de décrire l'utilisation de l'outil par l'institution. Après la semaine d'interrogations à mi-parcours, tous les étudiants sont invités à participer à un entretien individuel avec un intervenant pédagogique. Lors de cet entretien, l'outil est utilisé comme base d'information et permet à l'accompagnateur d'inciter l'étudiant en difficulté à réfléchir, en fonction de sa situation individuelle, à sa maîtrise des normes et des exigences universitaires ainsi qu'aux stratégies alternatives qu'il pourrait utiliser dans la deuxième partie de l'année académique, compte tenu des dispositifs d'accompagnement proposés par l'institution.

---

1. La place de l'évaluation formative dans la programmation de l'année académique

Dans les universités de la Communauté française de Belgique, une diversité d'actions est apparue pour promotionner la réussite en première année de l'enseignement supérieur (Salmon, De Ketele & Parmentier, 2003). Certaines de ces actions visent à accompagner l'intégration de l'étudiant face à la rupture de contexte entre les enseignements secondaire et supérieur (Leclercq & al, 1997). Dans ce cadre, les

évaluations programmées en cours d'année, qu'elles soient formatives ou certificatives, correspondent à un moment privilégié où l'étudiant peut évaluer les démarches d'apprentissage utilisées et son adaptation aux normes et aux exigences universitaires.

Les autorités académiques de la FUSAGx ont souvent mis en avant la priorité qu'elles accordent au travail de réflexion et de modification des cursus organisés dans l'institution ainsi qu'au développement d'initiatives novatrices en matière de lutte contre l'échec en première année. « Depuis le début des années 1990, la FUSAGx a entrepris un travail en profondeur sur la promotion de la réussite en première candidature. La société évolue et a de plus en plus besoin de cadres de haut niveau. A nous de les former en s'adaptant à ces besoins nouveaux tout en maintenant les niveaux d'exigence et de rigueur que requiert la complexité des problèmes rencontrés par les ingénieurs » (Lambert & al, 2001).

En ce qui concerne le rythme de l'année académique, dans le but de permettre à l'étudiant primant de s'adapter progressivement, la FUSAGx a choisi d'organiser le cursus de la première année en deux semestres. Ce choix a été défini bien avant que le décret réorganisant l'enseignement supérieur en la Communauté française de Belgique (pris dans le cadre de la mise en place du processus européen d'harmonisation des études) ne l'impose à toutes les institutions d'enseignement supérieur. La programmation de l'année académique en deux parties a été conçue en privilégiant deux modalités d'organisation : la répartition du contenu-matière de chaque cours en modules capitalisables sur l'ensemble de l'année académique et la programmation de séries d'interrogations organisées en janvier et en avril en plus des deux sessions traditionnelles d'examens en juin et en septembre. Le calendrier académique ainsi obtenu permet aux étudiants de percevoir leur première année académique comme la succession de périodes plus courtes avec des quantités de matières à ingérer plus réduites et des compétences à acquérir plus accessibles. De ce fait, cette planification permet d'augmenter le sentiment d'auto-efficacité de l'étudiant débutant face à la tâche proposée (Bandura, 1993). Enfin, la programmation de courtes périodes d'enseignement entrecoupées de périodes d'évaluation permet d'alléger considérablement la ou les sessions d'un étudiant en construisant progressivement sa performance académique finale.

Le décret « Bologne » de 2004 a néanmoins contraint la FUSAGx a opéré un ajustement de son calendrier académique toutes années confondues. Pour la première année, lors de cet ajustement, l'institution a veillé à garder, voire à accentuer, les avantages inhérents à l'organisation initiale de manière à renforcer encore le caractère transitif de la première année. « Le nouveau décret nous donne une relative liberté de définir la première année de bachelier comme une année de franche transition ». (Closset, 2003). Ainsi, une nouvelle période d'évaluation a été introduite au milieu du premier quadrimestre de manière à confronter encore plus précocement les étudiants aux attentes et aux normes d'une évaluation universitaire. Cette période d'évaluation a été conçue sous forme de tests dont les résultats, à la différence des interrogations de janvier et d'avril, sont uniquement informatifs et n'offrent aucune possibilité d'influence sur la note finale.

2. Développement d'outils d'interprétation des résultats obtenus à mi-parcours

2.1. Cadre du développement des outils

Quelle que soit la programmation, les interrogations du mois de janvier correspondent au premier moment où l'étudiant va être soumis à une évaluation collective sanctionnante effectuée par le corps enseignant dans l'ensemble des matières. A l'issue de cette série d'interrogations, une ou deux notes (théorie et/ou exercices) lui seront attribuées pour chaque cours. Pour prendre connaissance de ses résultats, soit l'étudiant se déplace aux différentes valves de chaque unité, soit l'étudiant se rend à la séance d'information entre encadrants (professeurs et assistants intervenant en première année) et étudiants organisée au début du mois de février, soit l'étudiant participe aux séances facultatives programmées par les unités où l'étudiant peut consulter individuellement ses copies et en discuter avec les assistants pédagogiques (correction, critères d'évaluation, dispositifs de remédiation, …). Pour les autorités académiques, ces différentes occasions offertes aux étudiants d'évaluer leur situation à l'entame du second semestre nécessitent de disposer d'outils, pour interpréter solidement les résultats aux interrogations organisées en janvier, outils construits et validés suivant le développement de la pédagogie universitaire. C'est dans ce cadre que nos premiers outils d'interprétation des résultats à mi-parcours ont été élaborés (Salmon, 2001).

Lors de la construction des premiers outils d'interprétation des résultats à mi-parcours, nous avions porté une attention particulière au statut que nous entendions donner à nos travaux, spécifiant et analysant les limites définies par d'autres chercheurs à propos de l'utilisation des modèles de prédiction de réussite. Sans entrer dans le détail de cette discussion (Salmon, 2001), nous aimerions en rappeler les axes principaux.

Dans une étude quantitative sur des facteurs explicatifs de la performance académique, De Ketele (1990) montre, notamment, que les capacités cognitives de base à l'entrée de l'enseignement supérieur n'interviennent que pour 25 % dans l'explication de la réussite. De Ketele précise que les corrélations obtenues avec la performance académique sont faibles et qu'il convient de rester prudent dans leur interprétation surtout vis-à-vis de leur communication aux étudiants. En ce qui concerne l'utilisation de résultats obtenus à mi-parcours à la FUSAGx, en fonction des contenus et des savoirs à y mettre en œuvre, l'élaboration de nos outils d'interprétation nous a permis de démontrer que, les résultats étant liés à des interrogations se situant en janvier, il est raisonnable de penser que situées à mi-chemin dans l'année académique, leurs corrélations avec la performance académique finale sont significativement plus favorables. Néanmoins, la manière dont l'étudiant réagit à la communication des résultats reste aussi importante que le résultat enregistré, en particulier lorsque ce résultat avoisine la fine tranche d'incertitude des modèles de réussite.

Poursuivant l'objectif d'élaborer un indicateur fiable et stable permettant aux encadrants d'orienter les étudiants en difficulté vers les filières de remédiation les plus indiquées, Lebrun (1996) développe un modèle basé sur le caractère prédictif de la session en janvier en terme de probabilité de réussite de l'année académique. Comme

les outils d'interprétation que nous souhaitons développés, le facteur F défini par Lebrun utilise uniquement pour indicateurs des variables disponibles à la suite des évaluations de janvier. Ajoutons, que par rapport à d'autres outils plus classiques, le facteur F réussit également, à travers l'utilisation de ces seules variables, l'intégration d'une mesure de l'implication de l'étudiant en mettant en rapport le nombre d'évaluations présentées en janvier avec le nombre total d'évaluations présentables. Dans leur version initiale, nos outils d'interprétation ne permettaient pas de mettre en avant l'impact de l'implication de l'étudiant (Salmon, 2001). L'intégration dans la construction de nos outils de cette variabilité va permettre, de manière singulière, de leur donner une signification supplémentaire et une plus grande validité.
En conclusion de la définition du cadre dans lequel nos outils d'interprétation sont élaborés, nous voudrions, une fois encore, rappeler le statut des résultats que nous proposons : Ceux-ci ne sont que des éléments parmi d'autres qui, bien souvent ne permettent que d'estimer la probabilité mais pas de prédire la réussite individuelle des étudiants, « parler de prédiction dans le domaine de l'éducation apparaissant sans doute comme largement optimiste » (Romainville, 1998). Néanmoins, les résultats obtenus par les étudiants lors des évaluations de janvier et les interprétations qui y sont liées peuvent servir de base d'information éclairante aux apprenants et initier chez eux un processus de rétroaction à travers leur questionnement vis-à-vis de leurs méthodes de travail et des résultats qui y sont associés.

2.2. Les variables utilisables pour le développement des outils

Dans un premier temps, la construction de nos outils s'est concentrée sur l'exploitation du lien existant entre la performance moyenne obtenue lors des interrogations à mi-parcours et la performance finale obtenue par l'étudiant en première ou en seconde session. Ce lien avéré nous a permis de définir par l'utilisation de régressions polynomiales un intervalle de confiance pour la performance finale à en première session (juin) ou toutes sessions confondues (juin et septembre) mais également un pronostic de réussite exprimé sur la base d'un pourcentage (Salmon, 2001). De plus, à partir des seules variables issues des résultats à mi-parcours, nous souhaitions rencontrer la préoccupation d'obtenir non seulement une information validée sur le résultat que l'étudiant peut espérer en fin d'année mais aussi une indication de l'implication de cet étudiant dans cette série d'interrogations. Nous avons donc cherché à faire évoluer nos outils dans ce sens. La différenciation des distributions du résultat moyen obtenu en janvier en fonction du nombre d'interrogations « relativement réussies » (ou d'insuffisances constatées) nous a permis de mettre en évidence l'utilité de prendre en compte cette variable. A l'inverse, le nombre d'interrogations présentées ne nous a pas permis de mettre en évidence des différences significatives.
Par rapport aux travaux de Lebrun, ces constatations nous amènent à formuler deux hypothèses. Premièrement, les interrogations organisées à mi-parcours étant des évaluations à dispenses (aucun impact si le résultat est négatif, report de note si le résultat est positif) sur des quantités de matières relativement restreintes, nous pensons que le statut particulier de ces épreuves influence fortement à la hausse le taux de participation à chacune d'entre elles. Nous posons de ce fait l'hypothèse que l'implication de l'étudiant dans son cursus est plus efficacement mesurée à travers le nombre d'interrogations « relativement réussies » plutôt qu'à travers un rapport

quelconque entre le nombre d'interrogations réussies et présentées. De plus, il convient de préciser que tous les étudiants de notre échantillon étant dans le même cursus, le nombre d'interrogations présentables n'est pas une variable et n'est donc pas exploitable. Deuxièmement, pour être significatif, le seuil de « réussite relative » est fixé à 50 %. Il peut paraître singulier que ce seuil soit situé à 10 % en- dessous du seuil de réussite finale. Encore une fois, nous trouvons l'explication dans le statut particulier des interrogations organisées en janvier à la FUSAGx. Comme il s'agit d'interrogations à dispenses et pas d'examens finaux, nous posons l'hypothèse que le niveau d'excellence attendu à ce moment par le corps professoral est inférieur à celui attendu en juin. Cet élément conforte l'idée institutionnelle que, « outre l'occasion d'alléger considérablement la session de fin d'année, les épreuves de janvier sont aussi l'occasion pour les étudiants de percevoir plus tôt l'exigence de l'enseignement universitaire, de tester leur maîtrise des pré-requis et d'ajuster leurs méthodes de travail » (Lambert & al, 2001).

Le développement de nos outils est aussi fonction de deux aspirations tirant la construction dans deux directions opposées. D'une part, l'information doit être simple, facile à mettre en œuvre et lisible par chaque intervenant. D'autre part, elle doit pouvoir tenir compte de la différence basée sur le nombre d'interrogations « relativement » réussies ce qui tant à complexifier les outils.
Partant de l'hypothèse posée par Beckers et Boxus (1996) que pour une catégorie d'étudiants de première année universitaire un temps d'adaptation plus long serait nécessaire et du travail de Closset (2002) qui démontre l'existence d'une double population parmi les étudiants de la FUSAGx nous avons confirmé qu'il était possible de scinder la population de première année sur le seul critère du nombre d'insuffisances constatées en janvier.

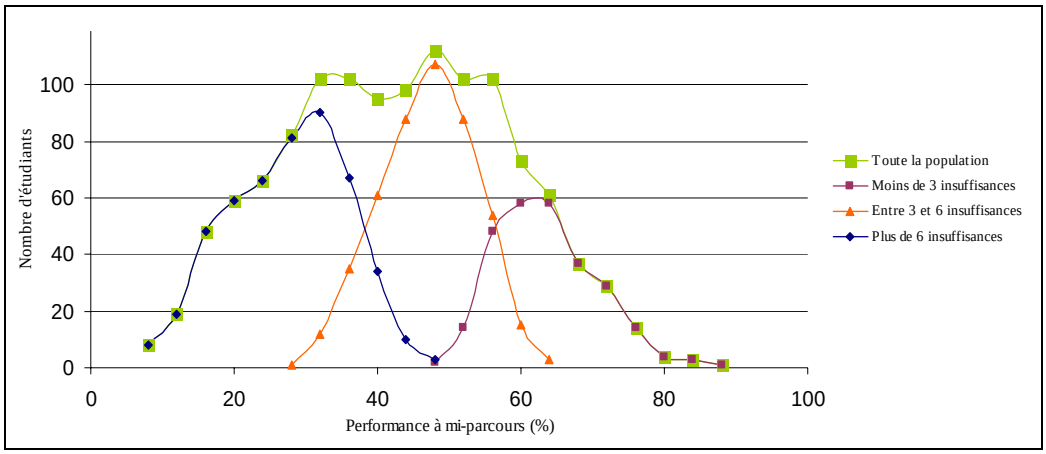


Figure 1 - **Distribution des étudiants en fonction de la performance à mi-parcours et du critère établi sur le nombre d'insuffisances constatées à mi-parcours**

Nous avons ainsi défini une population A qui réunit les étudiants pour lesquels moins de trois insuffisances (sur neuf notes attribuées) ont été constatées à mi-parcours et pour lesquels la probabilité de réussite est très importante (> 98 %), une population B qui réunit les étudiants pour lesquels on constate entre trois et six insuffisances à mi-

parcours et pour laquelle, à ce moment, il n'est pas possible de poser un pronostic de réussite ou d'échec et enfin une population C, qui réunit les étudiants pour lesquels plus de six insuffisances ont été constatées à mi-parcours et qui n'ont quasiment plus aucune probabilité de réussite (< 5 %).

2.3. La construction des outils

Les outils que nous avons développés ont été construits en utilisant les variables dont nous venons de présenter le caractère indicateur. Il s'agit, d'une part, de la moyenne pondérée (le poids de chaque cours est fonction de la charge horaire) obtenue à l'issue des interrogations à mi-parcours et, d'autre part, du nombre d'insuffisances constatées lors de ces interrogations. En pratique, pour la construction des outils correspondant à une année académique déterminée, nous avons pu déterminer que la plus grande validité des outils est obtenue en utilisant les unités d'observation disponibles dans les cohortes des trois années précédentes, ce qui permet d'assurer une certaine évolutivité des outils. Nous avons également pu démontrer que lorsqu'un étudiant ne présentait pas une série complète d'interrogations en janvier, si ses absences sont assimilées à une note nulle, cette opération n'affecte pas sensiblement la validité des outils. Ce constat s'explique d'autant mieux que les étudiants absents à une ou plusieurs interrogations correspondent dans leur grande majorité aux étudiants les moins performants dans les autres matières.

Le premier outil d'interprétation construit permet la détermination d'un intervalle de confiance pour la performance académique en première session. Pour déterminer la valeur centrale et la largeur de cet intervalle, les unités d'observation issues des trois années académiques précédentes sont réparties en trois populations selon le critère du nombre d'insuffisances. Pour chacune de ces populations, la valeur centrale pourra être obtenue en utilisant la fonction issue de la régression linéaire entre le pourcentage à mi-parcours (JMOYE) et la performance académique obtenue en première session. Le pourcentage à mi parcours est obtenu en calculant la moyenne pondérée par le poids relatif en terme de volume horaire des différentes matières de toutes les notes de janvier. La largeur de l'intervalle est définie pour chaque population en calculant la valeur moyenne des écarts entre la valeur prédite pour le centre de l'intervalle et la performance académique réellement obtenue par chaque étudiant. En résumé, l'outil construit peut se présenter sous la forme d'un tableau synthétique reprenant, en fonction de la population de l'étudiant (du nombre d'insuffisances constatées lors des interrogations à mi-parcours), la formule de calcul du centre et la valeur de la demi largeur de l'intervalle de confiance.

Nombre d'insuffisances à mi-parcours	Formule de calcul pour le centre de l'intervalle	Demi largeur de l'intervalle
0, 1 ou 2	**0,807 JMOYE + 17,473**	**7**
3, 4, 5 ou 6	**0,980 JMOYE + 9,424**	**12**
7,8 ou 9	**0,996 JMOYE + 7,018**	**12**

Tableau 1 - Outil de prédiction d'un intervalle de confiance pour la performance en première session (Année académique 2003-2004)

Le deuxième outil d'interprétation construit permet le calcul d'un intervalle de confiance pour la performance académique finale. Selon le pronostic établi par l'outil que nous venons de construire, si la valeur centrale de l'intervalle de confiance obtenu est supérieure à 59 %, l'étudiant peut être considéré en situation de réussite à l'issue de la première session. Dans le cas contraire, il convient de construire un outil pour déterminer un intervalle de confiance de la performance académique en seconde session. La performance académique finale « toutes sessions confondues » sera alors déterminée soit, si le pronostic en première session est positif, par l'intervalle de confiance obtenu par l'outil construit pour la première session, soit, si le pronostic est négatif, par l'intervalle de confiance obtenu par l'outil construit pour la seconde session. La valeur centrale et la largeur de l'intervalle sont déterminées comme pour l'intervalle de confiance à l'issue de la première session mais nous utilisons cette fois les données des étudiants ayant effectivement présenté une seconde session au cours des trois années précédentes. Pour des raisons évidentes de statistiques nous sommes amené à traiter ensemble les deux premières populations. Comme pour la première session, Nous obtenons un tableau synthétique reprenant, en fonction du nombre d'insuffisances constatées, la formule de calcul du centre et la valeur de la demi largeur de l'intervalle de confiance pour la seconde session.

Nombre d'insuffisances à mi-parcours	Formule de calcul pour le centre de l'intervalle	Demi largeur de l'intervalle
0 à 6	**0,671 JMOYE + 27,407**	**15**
7,8 ou 9	**1,076 JMOYE + 9,565**	**11**

Tableau 2 - Outil de prédiction d'un intervalle de confiance pour la performance toutes sessions confondues (Année académique 2003-2004)

Pour le pronostic toutes sessions confondues, l'augmentation de largeur de l'intervalle de confiance permet de rendre compte de la plus grande variabilité de l'évolution de la performance académique entre janvier et septembre pour les étudiants devant présenter une seconde session d'examens. Nous posons l'hypothèse que cette variabilité permet de traduire l'effet de la quantité mais aussi de la qualité de l'investissement entre les sessions de juin et de septembre.

Le dernier outil d'interprétation construit permet de compléter l'information obtenue en établissant un pronostic de réussite/échec en première session ou en fin de première année toutes sessions confondues. L'utilisation des intervalles de confiance construit précédemment et des régressions logistiques construites dans l'outil initial (Salmon, 2001), nous permet d'établir un pronostic et d'en donner un degré de certitude. Lorsque l'entièreté de l'intervalle de confiance de la performance académique finale se situe au-dessus du seuil de réussite (la plupart de ces étudiants appartiennent à la population A), le pronostic posé est positif. Par contre, lorsque l'intervalle se situe entièrement sous le seuil de réussite (population C), le pronostic posé est négatif. Enfin, lorsque l'intervalle contient le seuil de réussite (population B), aucun pronostic ne peut être posé. Néanmoins, dans ce cas, l'intervalle de confiance reste en soi une information éclairante et avérée. En effet, si la validité de l'intervalle de confiance de la performance académique finale est élevée pour l'ensemble des

étudiants (> 88 %), cette validité est maximale (> 95 %) pour la population intermédiaire.
L'information peut encore être utilement complétée en déterminant la probabilité de réussite calculée à partir de la régression logistique construite sur des taux de réussite observés sur l'ensemble de l'échantillon constitué à partir des données des trois années académiques précédentes (Salmon, 2001).

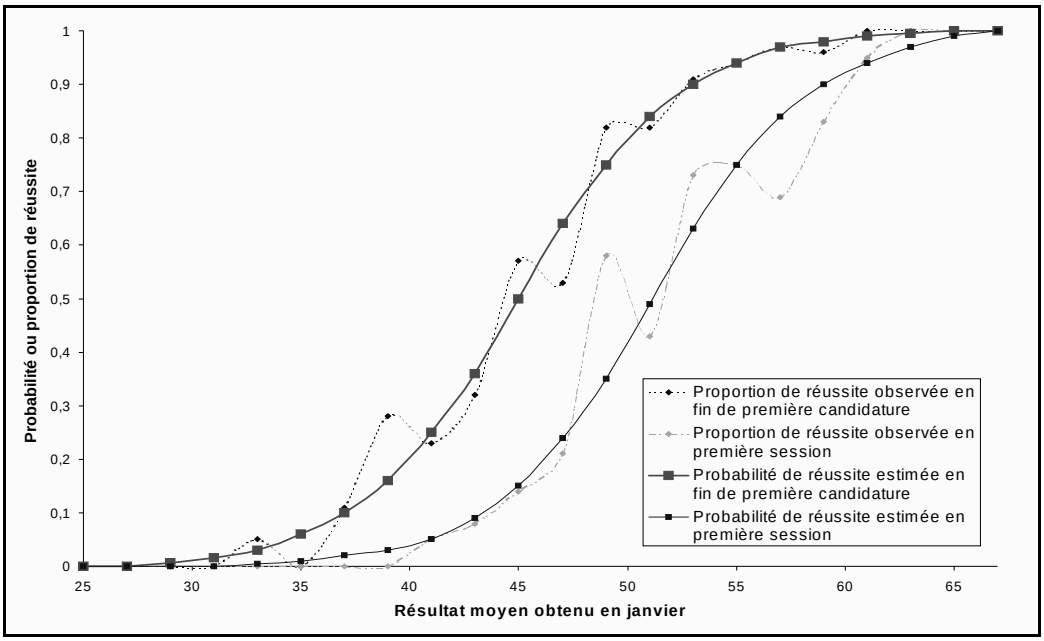


Figure 2 - Evolution des proportions de réussite observées et des probabilités de réussite estimées par les modèles issus de la régression logistique en fonction des sous-classes déterminées par le résultat moyen aux interrogations de janvier

2.4. La validité des outils développés

La validité des informations produites par les outils que nous avons construits est fortement établie et reste constante dans le temps. A titre d'information, nous fournissons les données pour les outils construits pour la première session lors l'année académique 2003-2004.

Pronostic de janvier	Situation de réussite		Situation d'échec		Situation à risque		Dans l'intervalle		Hors de l'intervalle	
Abandon (avant juin)	0 +		54 +		7					
Population A	28 ++	0 ++	0 ++	0 ++	3 +	8 -	31 ++	3 -	5 -	0 -
Population B	0 +	0 +	0 +	12 ++	32 +	44 +	31 ++	45 ++	1 -	11 -
Population C	0 ++	0 ++	2 --	82 ++	0 +	4 +	0 ++	78 ++	2 --	8 -
Résultat de juin	Réussi	Ajourné	Réussi	Ajourné	Réussi	Ajourné	Réussi	Ajourné	Réussi	Ajourné

Tableau 3 - Validation des outils de pronostic de la réussite et de la performance académique en première session

Pour le pronostic de réussite/échec en première session, l'outil ne se prononce que pour 58 % des étudiants et se révèle exact à 98 %. Les pronostics erronés correspondent tous au cas d'étudiants pronostiqués en situation d'échec par l'outil et « repêchés par le jury » lors de la délibération de première session (performance académique légèrement inférieure à 59 % et une seule insuffisance).
Pour le pronostic d'un intervalle de confiance de la performance en première session, le résultat obtenu en juin appartient à l'intervalle dans 87 % des cas sans pouvoir mettre en évidence un particularisme significatif pour l'une des trois populations définies en fonction du nombre d'insuffisances constatées à mi-parcours.
Remarquons que ces mêmes commentaires pourraient être apportés si nous présentions les résultats des outils développés sur l'ensemble des deux sessions avec des taux de validité encore très légèrement supérieurs.

Pronostic de janvier	Résultat à mi-parcours au dessus du pivot		Résultat à mi-parcours en dessous du pivot	
Population A	36 ++	3 --	0 ++	0 ++
Population B	20 ++	4 -	12 -	52 ++
Population C	0 ++	0 ++	2 --	86 ++
Résultat de juin	1ère session réussie	2nd session éventuelle	1ère session réussie	2nd session éventuelle

Tableau 4 - Validation de l'outil de pronostic d'une éventuelle seconde session

Une autre information importante qui peut être transmise en janvier à l'étudiant est la probabilité pour lui de devoir envisager une réussite en deux sessions. La possibilité de transmettre cette information est liée à la validité de la définition des valeurs pivots calculées à partir du premier outil d'interprétation construit. Pour les deux populations les plus typées (A et C), l'homogénéité de ces populations face à la réussite en

première session est clairement confirmée aussi bien par l'observation du résultat réellement obtenu que par le recoupement quasi-systématique (95 %) de l'application du test sur la valeur pivot. Les quelques étudiants de la population A pour lesquels l'application du test sur la valeur pivot se révèle incorrecte sont tous des étudiants qui obtiennent un résultat en janvier très proche de la valeur pivot, échouent de très peu en première session et réussissent très honorablement en seconde session. Pour les étudiants appartenant à la population C et qui mettent en défaut la valeur pivot, il s'agit forcément des cas particuliers déjà détectés précédemment dans la validation du pronostic de réussite en première session. Enfin, pour la population B, les défauts de prédiction sur l'existence d'une seconde session se chiffrent à environ un cinquième de l'échantillon, mais, pour la grande majorité de ces étudiants (90 %), l'outil ne propose pas de pronostic de réussite et la performance académique finale obtenue en première session appartient à l'intervalle de confiance pronostiqué.

3. Utilisation des outils d'interprétation dans le cadre du dispositif d'accompagnement pédagogique des étudiants de première année à la FUSAGx

3.1. Outils d'information utilisés lors des entretiens méthodologiques de janvier

La FUSAGx déploie tout un arsenal de mesures visant la réussite des primo arrivants ou des étudiants répétants. Parmi celles-ci, citons la mise en place d'une propédeutique obligatoire de deux semaines, l'engagement d'une dizaine d'assistants pédagogiques, des travaux pratiques et/ou des travaux dirigés par petits groupes, un tutorat individualisé des étudiants, des séances de monitorat tous les mercredis après-midi, des séances collectives de réflexion sur les méthodes de travail ou encore des feedbacks après les évaluations (Salmon, 2005).

Si ces évaluations ont initialement été introduites pour permettre un allégement de la session de fin d'année par l'obtention d'un certain nombre de dispenses de parties de matières, progressivement un autre objectif institutionnel est apparu. Au-delà des séances de feedbacks, collectives ou individuelles, organisées au sein des unités d'enseignement, d'autres types d'entretiens individuels se sont développés, se centrant principalement sur une analyse globale de l'ensemble des résultats obtenus progressivement par l'étudiant et sur une réflexion à propos des stratégies d'apprentissage et des méthodes de travail mises en place. Depuis une dizaine d'année, les moments d'évaluation et les séances d'analyse des résultats se sont progressivement multipliés avec comme objectif d'informer régulièrement l'étudiant sur sa situation académique et d'inciter, ou de soutenir, tout au long de la première année d'enseignement supérieur une démarche de réflexion métacognitive (Noël & Romainville, 1998).

L'existence et l'évaluation positive des outils d'interprétation des résultats à mi-parcours que nous avons construits nous confortent dans l'hypothèse qu'il est possible de communiquer aux étudiants une information fiable et significative à la suite des interrogations organisées à mi-parcours. Cette information peut prendre un aspect individualisé ou collectif.

MATRICULE	NOMTRI	STATUT	PREJ AN	NBR ABS	MOY	NBR INS	Critère 1	Critère 2	Critère 3	Note de repport éventuelle REP BIC	REP BOT	REP CHI	REP MAT	REP PHY	REP ZOL
		R	1	0	63,1	0	16	A	P.A	12,0		13,0			14,0
		N	1	0	23,9	9	6	C	P.C3						
		N	1	0	11,0	9	3	C	P.C4						
		R	1	0	57,4	3	14	B	P.A				12,0		
		N	1	0	43,2	6	11	B	P.B						
		R	1	0	66,0	0	17	A	P.A	14,5	13,5	13,5			12,0

Résultats de janvier cours par cours (toutes les notes sur 20)																Pronostic juin					Pronostic final				
Bio Cell	Biota	Chi 1	Chi Ex	Math Th	Math Ex	Phys Th	Phys Ex	Zool 1	Zool 2	centre	ecart	max	min	R/E	centre	ecart	max	min	R/E
	10,5			11,0	13,0	12,5	15,0			68	8	60	76	R					
3,0	2,5	8,0	1,0	9,0	9,0	6,0	4,5	1,0	5,0	29	12	17	41	E	44	23	21	67	X
2,0	1,0	4,0	5,5	3,0	3,0	4,0	0,5	0,0	0,0	18	12	6	30	E	38	23	15	61	X
12,0	7,0	9,0	17,5			11,5	15,0	7,5	16,0	65	11	54	76	X	65	13	52	78	X
15,0	11,5	11,0	7,5	3,0	0,0	8,5	8,5	7,0	15,5	49	11	38	60	X	53	15	38	68	X
				16,0	11,0	12,5	14,0			70	8	62	78	R					

Tableau 5 – Présentation des pronostics individualisés pour chaque étudiant.

Sous forme individualisée, les outils permettent de proposer, pour chaque étudiant, l'état de sa « situation académique instantanée ». Dans ce tableau, l'accompagnateur pédagogique et l'étudiant trouvent l'ensemble des données chiffrées disponibles qui permettent d'éclairer un échange autour des stratégies à mettre en place au cours de la seconde partie de l'année académique. Ce tableau autorise également de définir des objectifs, y compris en terme de résultats chiffrés, à atteindre en avril, en juin, voire en septembre. Les données disponibles sont les résultats bruts obtenus pour chaque interrogation (le cas échéant, pour un répétant, remplacés par une note de report obtenue la première année), la position de l'individu dans les différentes populations définies, les pronostics pour la première session et éventuellement les pronostics pour la seconde session.

	Janvier		Première session			Deuxième session		
	Cote	Densité	Performance	R/E	Taux	Performance	R/E	Taux
Entre 3 et 6 insuffisances	10		[0 ; 25]	E		[23 ; 51]	E	0 %
	14		[3 ; 29]	E		[25 ; 53]	E	0 %
	18		[8 ; 34]	E		[27 ; 55]	E	0 %
	22		[12 ; 38]	E		[29 ; 57]	E	0 %
	26		[17 ; 43]	E		[31 ; 59]	E	0 %
	30		[21 ; 47]	E		[33 ; 61]	E	1 %
	34	4 %	[26 ; 52]	E		[36 ; 64]	X	2 %
	38	8 %	[30 ; 56]	E		[38 ; 66]	X	7 %
	42	14 %	[35 ; 61]	X		[40 ; 68]	X	14 %
	44	18 %	[39 ; 65]	X		[42 ; 70]	X	26 %
	50	25 %	[44 ; 70]	X		[44 ; 72]	X	46 %
	54	14 %	[48 ; 74]	X		[46 ; 74]	X	64 %
	58	11 %	[53 ; 79]	X		[48 ; 76]	X	79 %
	62	6 %	[57 ; 83]	X		[50 ; 78]	X	89 %
	66	1 %	[62 ; 88]	R	95 %	[52 ; 80]	X	
	70		[66 ; 92]	R	98 %	[54 ; 82]	X	
	74		[71 ; 97]	R	100 %			
	78		[75 ; 100]	R	100 %			
	82		[80 ; 100]	R	100 %			
	86		[84 ; 100]	R	100 %			

Tableau 6 – Présentation collective des pronostics pour la population B

Sous forme globale, l'interprétation des résultats à mi-parcours se caractérise d'abord à partir du nombre d'insuffisances et ensuite en fonction du pourcentage moyen des notes obtenues en janvier. Dans un tableau produit pour chacune des trois

populations identifiées à partir du nombre d'insuffisances, le pourcentage obtenu à mi-parcours est subdivisé en tranches de 4 %. Pour chacune de ces tranches, le tableau propose le pourcentage d'étudiants appartenant à cette tranche, l'intervalle de confiance de la performance académique en première session, le pronostic réussite/échec en première session, en cas de pronostic de réussite, le taux de probabilité de cette réussite et, le cas échéant (pas de réussite pronostiquée en première session), les même valeurs pour la seconde session. Sous cette forme, l'accompagnateur pédagogique ou l'étudiant peuvent mieux visualiser la « situation académique instantanée » de l'étudiant au sein du groupe.

Dans les faits, lors des entretiens individuels programmés en janvier, deux séances sont organisées. La première est facultative et concerne les étudiants appartenant aux populations A et B. La deuxième est obligatoire mais ne concerne que les étudiants appartenant à la population C.

Pour le premier groupe d'étudiants, il s'agit essentiellement de conforter leurs méthodes de travail et leur gestion du temps. Il est aussi souvent indispensable de les rassurer par rapport à l'interprétation des résultats obtenus. Il n'est en effet pas évident pour tous les étudiants ayant obtenu un résultat compris entre 50 et 60 % en janvier de comprendre que ce résultat peut être considéré comme très positif et annonciateur d'une réussite « facile » en juin. A cet égard, les outils d'interprétation construits permettent de donner des précisions chiffrées mais surtout une légitimité au message transmis. Pour certains étudiants de la population B, il convient, par contre, de leur faire prendre conscience qu'une réussite globale reste très certainement envisageable mais nécessite peut-être une stratégie de réussite en deux sessions.

Pour les étudiants dont le pronostic de réussite est quasiment nul, il convient de les encourager à réfléchir, dès janvier, à la mise en place de stratégies de réussite autrement (Closset, 2005). Pour inciter à cette réflexion métacognitive, tous les étudiants de la population C sont tenus de rencontrer individuellement un des membres de l'équipe pédagogique (coordinateur pédagogique ou assistants pédagogiques) pour faire le point sur leur « situation académique instantanée ». Afin d'assurer une uniformité aux entretiens, les différents messages relayés auprès de ces étudiants sont définis préalablement en concertation par l'ensemble de l'équipe pédagogique notamment en fonction des objectifs et des contraintes définis par l'institution. Bien évidemment, en cours d'entretien, les conseils donnés par l'accompagnateur doivent aussi tenir compte des éléments d'analyse apportés par l'étudiant. En effet, pour créer des conditions propices à ce type d'accompagnement, il importe que l'étudiant qui participe s'y sente intégré, ou à tout le moins pris en compte (Houart & Salmon, 2006). En pratique, les étudiants de la population C sont classés selon leur pourcentage moyen obtenu à mi-parcours et donc de la certitude du pronostic d'échec qui en découle. Pour les quelques étudiants primants les « mieux classés », pour lesquels la probabilité de réussir l'année est quasi nulle mais aussi pour lesquels la probabilité de réussir en redoublant est significative (Mignon, 2006), le conseil est de poursuivre la filière normale avec l'ambition d'une réussite de deux années en trois ans ou encore de répéter la première année avec des reports quinquennaux. Pour les étudiants un peu moins bien classés pour lesquels le pronostic de réussite dans l'année est nul mais qui pourraient envisager une répétition avec fruits, il leur est systématiquement proposé de s'inscrire dans une filière d'étalement de la première année sur deux années académiques (Salmon, Fabry & Closset, 2006).

Pour les étudiants les moins performants en janvier et pour lesquels aucun pronostic ne peut être posé vis-à-vis de la réussite éventuelle de la répétition de la première année, une réflexion en profondeur sur leur projet d'études est conseillée. En fonction de la matérialisation de cette réflexion, un accompagnement individuel peut être mis en place pendant le second quadrimestre soit pour préparer le redoublement soit pour envisager une réorientation vers un autre cursus ou un autre projet de vie. Il en est de même pour les étudiants répétants appartenant à la population C qui envisage une réorientation.

3.2. Détournement des outils dans le cadre d'une réflexion métacognitive précoce

Depuis trois ans, pour promouvoir une réflexion métacognitive encore plus précoce, une série de tests sont organisés dès le milieu du premier quadrimestre (novembre). Comme en janvier et en avril, à la suite de cette semaine de tests, une possibilité d'entretien individuel est proposée à chaque étudiant. A cette occasion, les outils d'interprétation que nous avons développés ont été adaptés pour pouvoir être utilisés pour alimenter la réflexion.

A priori, nous postulons que à ce moment de l'année, à l'inverse de ce que nous avons pu matérialiser en janvier, les résultats obtenus aux tests ne permettent pas, quelle que soit la « situation académique instantanée » déterminée, de poser un pronostic valide par rapport la réussite ou à l'échec en fin de première année. Il sera intéressant lorsque nous disposerons d'un échantillon statistiquement suffisant de valider cette hypothèse.

Références

Bandura, A. (1993). *Perceived self-efficacy in cognitive development and functioning*. Educational Psychologist, 2.

Becquers, J. & Boxus, E. (1996). *La traversée des candidatures universitaires. Une analyse de cas*. Res Adademica, 13.

Closset, J.L. (2002). *Essai d'analyse des interrogations de Noël 2001-2002*. Document interne. Gembloux : Faculté universitaire des sciences agronomiques de Gembloux, Commission de pédagogie universitaire.

Closset, J.L. (2003). *Une première année de baccalauréat en deux ans. Proposition concernant les étudiants bioingénieurs primants*. Document interne. Gembloux : Faculté universitaire des sciences agronomiques de Gembloux, Commission de pédagogie universitaire.

Closset, J.L. (2005). *0 % d'échec à la FUSAGx – Stratégies alternatives de réussite en première année*. Note de travail. Gembloux : Faculté universitaire des sciences agronomiques de Gembloux.

De Ketele, J.M. (1990). *Le passage de l'enseignement supérieur*. Vie Pédagogique, 66.

Houart, M. & Salmon, D. (2006). *Comment développer des actions d'accompagnement pédagogique : synthèse d'échanges et d'analyse de pratiques professionnelles en Communauté française*. Introduction au mini-colloque « L'accompagnement pédagogique des étudiants dans l'enseignement supérieur : conditions, actions et questions sur les critères de qualité ». XXIIIème Congrès

de l'Association internationale de Pédagogie Universitaire. Monastir : Faculté de Pharmacie.

Lambert & al, 2001 C. *Rapport annuel sur la lutte contre l'échec dans le premier cycle de l'enseignement universitaire - Année académique 1999-2000*. Faculté universitaire des Sciences agronomiques de Gembloux.

Lebrun, M. (1996). *Analyse du caractère prédictif de la session de janvier en terme de probabilité de réussite de l'année académique – Le facteur F*. Res Adademica, 13.

Leclercq, D. & al (1997). *Franchir le cap des candis : Rapport du groupe de travail "Réussites en candidature"*. Bruxelles : Conseil interuniversitaire de la Communauté française de Belgique.

Mignon, J. (2006). *Suivi de cohortes*. Communication à la 4ème journée des assistants pédagogiques. Gembloux : Faculté universitaire des Sciences agronomiques de Gembloux

Noël, B. & Romainville, M. (1998). *Accompagner les étudiants*. in Frenay, M., Noël, B., Parmentier, P. & Romainville, M. *L'étudiant-apprenant. Grille de lecture pour l'enseignant universitaire*. Louvain-la-Neuve : Editions De Boeck Université.

Romainville, M. (1998). *La sélection en questions*. in Frenay, M., Noël, B., Parmentier, P. & Romainville, M. *L'étudiant-apprenant. Grille de lecture pour l'enseignant universitaire*. Louvain-la-Neuve : Editions De Boeck Université.

Salmon, D. (2001). *Aspect prédictif des interrogations dispensatrices organisées à mi-parcours en première année universitaire. Etude de la liaison entre les résultats obtenus aux interrogations de janvier et la performance en fin de première année universitaire*. Mesure et évaluation en Education, 24(1).

Salmon, D., De Ketele, J.M. & Parmentier, P. (2003). *Une grille d'analyse des dispositifs d'accompagnement pédagogique. Sept clés pour appréhender les dimensions d'un dispositif*. Communication au XIXème Congrès de l'Association internationale de Pédagogie Universitaire. Louvain-la-Neuve : Université catholique de Louvain.

Salmon, D. (2005). *Assistant pédagogique, une fonction à mi-chemin entre l'assistant chercheur et le professionnel de l'accompagnement. Cas de la Faculté universitaire des sciences agronomiques de Gembloux*. Actes du 3ème Colloque « Questions de pédagogie dans l'enseignement supérieur ». France : Lille

Salmon, D, Fabry, G. & Closset, J.L. (2006). *L'étalement avec remédiation, évaluation d'une filière de réappropriation de son cursus par l'étudiant en difficulté intégrant travail sur la méthodologie universitaire et les compétences langagières*. Communication au XXIIème Congrès de l'Association internationale de Pédagogie Universitaire. Monastir : Faculté de Pharmacie.

PROFILS DE PERFORMANCES ATTEINTS EN ORTHOGRAPHE ALLEMANDE ET FRANÇAISE : LES ENFANTS A DIFFICULTES

Sylvie Bodé, Université du Luxembourg

Joyce Serres, Ecole Primaire Differdange

Mots-clés : orthographe, modèles de traitement de l'information, profils de performances.

Résumé :

La recherche présentée s'intéresse aux profils de performances atteints en orthographe allemande et française par les enfants de langue luxembourgeoise fréquentant la 5e année primaire. Cette étude vise la comparaison des profils de performances d'un groupe d'enfants présentant des difficultés en orthographe (n=37) avec ceux d'un groupe d'enfants sans difficultés (n=156). Dans ce but, des épreuves d'orthographe française et allemande ont été élaborées et appliquées. Ces épreuves se basent sur les modèles théoriques du traitement de l'information développés pour rendre compte de l'apprentissage du langage écrit. Des épreuves tiennent compte du facteur 'fréquence de mot' et 'transparence des règles de traduction phonèmes-graphèmes'. Les résultats montrent que les performances des enfants avec difficultés se situent significativement en dessous de celles des enfants sans difficultés. Tel est le cas pour l'orthographe française et pour l'orthographe allemande. En orthographe française les stratégies des enfants avec difficultés ne diffèrent pas significativement de celles des enfants sans difficultés. Tel n'est pas le cas pour l'orthographe allemande. En orthographe allemande, les enfants avec difficultés semblent se fier davantage aux règles de correspondances phonèmes-graphèmes les plus transparentes (graphies consistantes acontextuelles, graphies consistantes contextuelles). Les enfants avec difficultés ont plus recours aux représentations orthographiques des mots dont l'orthographe a été apprise de manière explicite en classe.

Profils de performances atteints en orthographe allemande et française : Les enfants à difficultés

Les modèles de traitement de l'information

Depuis le début des années 80, une série de modèles théoriques du traitement de l'information ont été élaborés pour rendre compte de l'apprentissage du langage écrit : les modèles à deux voies et les modèles développementaux.

Les modèles à deux voies se caractérisent, comme le nom l'indique, par la présence de deux voies de traitement (p.ex. Frith, 1985): la procédure orthographique et la procédure phonologique (Mousty, Leybaert, Alegria, Content, & Morais, 1994). La procédure orthographique permet de récupérer une représentation orthographique d'un mot connu à partir d'une représentation orthographique stockée en mémoire. Seule l'orthographe des mots familiers peut être accédée par cette procédure. L'orthographe de mots irréguliers (p.ex. femme) doit être récupérée par cette procédure. La procédure phonologique par contre permet d'écrire des mots peu ou pas familiers. Elle se base sur un système sub-lexical de règles de conversion phonologie-orthographe. Cette procédure est sensible au degré de régularité des relations phonèmes-graphèmes (pour plus de détails voir Mousty, Leybaert, Alegria, Content, & Morais; 1994).

Les modèles développementaux se caractérisent par les différentes stratégies dominantes qui se succèdent au cours de l'apprentissage du langage écrit (pour une vue d'ensemble voir Marx, 1997). Au début de l'apprentissage du langage écrit des stratégies tenant compte des relations régulières entre phonèmes et graphèmes dominent. Avec l'expérience, les stratégies évoluent et permettent de plus en plus de rendent compte des relations plus complexes entre phonèmes et graphèmes.

Modèles de traitement de l'information et évaluation

Les modèles du traitement de l'information offrent de nouvelles possibilités pour l'évaluation. D'une part, la variation de la fréquence des mots permet de récolter des indices en rapport avec le développement de la procédure orthographique et de la procédure phonologique. D'autre part, la variation de la régularité des relations phonèmes-graphèmes permet de recueillir des informations sur le développement des différentes stratégies. La 'Batterie d'évaluation du langage écrit et de ses troubles' (BELEC, Mousty, Leybaert, Alegria, Content & Morais, 1994), le 'Salzburger Lese- und Rechtschreibtest. Verfahren zur Differentialdiagnose von Störungen des Lesens und Schreibens für die 1. bis 4. Schulstufe' (Landerl, Wimmer, & Moser, 1997) et la 'Hamburger Schreibprobe' (May,1996) ne sont que quelques exemples de batteries de tests qui se basent sur les modèles du traitement de l'information.

En éducation, ce type d'évaluation est souvent appliqué au début de l'apprentissage du langage écrit pour détecter le plutôt possible toute difficulté de développement d'une des stratégies de traitement de l'information. Appliqué après quelques années d'apprentissage du langage écrit, cette forme d'évaluation permet surtout de dresser un tableau nuancé des stratégies déjà développées et des stratégies qui restent à développer. En cas de rééducation, la formulation des objectifs pédagogiques peut découler directement des profils de performances dressés. A une plus grande échelle, cette forme d'évaluation permet de connaître les stratégies cognitives que les programmes scolaires en cours tendent à favoriser.

Problématique

En première année primaire, les enfants scolarisés au Luxembourg sont alphabétisés en allemand. L'enseignement systématique de l'orthographe française commence en 3e année primaire. Comme nous l'avons déjà indiqué, les batteries d'évaluation basées sur les modèles de traitement de l'information sont soit en langue allemande, soit en langue française. Vu l'enseignement parallèle de l'orthographe allemande et française au Luxembourg, il y a évidence de l'intérêt d'une batterie d'évaluation contenant deux épreuves comparables, l'une en langue allemande et l'autre en langue française. Disposer de deux épreuves basées sur les mêmes modèles théoriques et construites selon les mêmes critères permet de mieux comparer les stratégies développées dans chacune des deux langues.

Méthodologie

Epreuves

L'épreuve d'orthographe française est une adaptation de l'épreuve ORTHO de la batterie BELEC (Mousty, Leybaert, Alegria, Content, & Morais, 1994). L'épreuve ORTHO est une épreuve présentée sous forme de texte lacunaire à compléter par des mots dictés. Les mots dictés tiennent compte de deux facteurs: la fréquence des mots (fréquent, rare) et de la nature des correspondances entre phonèmes et graphèmes. Ainsi par exemple les graphies consistantes acontextuelles se basent sur les règles de correspondance phonèmes-graphèmes systématiques et indépendantes du contexte. D'autres catégories de graphies étudiées sont : les graphies consistantes contextuelles (c.à.d. les règles de correspondance sont systématiques dans le contexte ou elles s'inscrivent), les graphies inconsistantes contextuelles (c.à.d. les correspondances phonèmes-graphèmes ne sont pas systématiques et dépendent du contexte dans lequel elles s'inscrivent) et les graphies dérivables par la morphologie (certaines graphies muettes peuvent êtres dérivées par la morphologie).

Vu l'importance du facteur 'fréquence des mots', l'adaptation tient compte du vocabulaire transmis dans le cadre du programme scolaire officiel. En conséquence, 8 mots de l'épreuve ORTHO ont été remplacés. En plus, 6 mots irréguliers (p.ex. auto*mn*e) ont été ajoutés à la liste. Quatre-vingts huit graphies ont été retenues pour l'étude des stratégies appliquées en orthographe française. L'épreuve d'orthographe allemande a été constituée en tenant compte des mêmes facteurs (fréquence des mots, nature des correspondances entre phonèmes et graphèmes). Vu le nombre réduit (2 items) de mots utilisés fréquemment dans le programme officiel et présentant soit une graphie non-dérivable par la morphologie soit une graphie irrégulière, ces types de graphies ont été étudiés uniquement pour les mots rares. L'étude finale des stratégies appliquées en orthographe allemande se base sur 60 graphies.

Echantillon

La recherche présentée porte sur un groupe de 193 enfants de langue maternelle luxembourgeoise fréquentant la 5e année d'étude. Elle a pour objectif de comparer le profil de performances développé par un groupe d'enfants présentant des difficultés au niveau de l'orthographe allemande et française (1er quartile de l'ensemble de la population ; n=37) au profil développé par un groupe d'enfants sans difficultés en orthographe (2e, 3e et 4e quartile; n=156). Tous les enfants ont été alphabétisés en allemand dès leur première année primaire. Dès la troisième année primaire ils ont été initiés à l'orthographe française.

Principaux résultats et interprétations

Les analyses statistiques (ANOVA) montrent que pour les différents types de graphies de l'épreuve française et de l'épreuve allemande la moyenne des enfants avec difficultés (1er quartile) se situent significativement ($p<0.001$) en dessous de celle des enfants sans difficultés (2e, 3e et 4e quartile).

Pour l'épreuve d'orthographe française le profil de performance des enfants avec difficultés est proche de celui des enfants sans difficultés. Les résultats obtenus pour les mots fréquents (graphies consistantes contextuelles, graphies inconsistantes contextuelles dominantes, graphies morphologiquement dérivables) se situent significativement au dessus de ceux réalisés pour les mots rares. L'importance de cet effet de fréquence est comparable pour les enfants avec difficultés et les enfants sans difficultés (GLM). Ces premiers résultats montrent que le recours à la procédure orthographique joue un rôle important pour les mots dont l'orthographe a été introduit de manière explicite en classe. L'application de règles plus ou moins complexes et acquises de manière plus ou moins implicite reste difficile. Tel est le cas aussi bien pour les enfants sans difficultés que pour les enfants avec difficultés.

Même si le recours à la procédure orthographique joue un rôle important, l'effet de dominance observé au niveau des graphies inconsistantes contextuelles suggère que les deux groupes d'enfants ont acquis (probablement de manière implicite) une règle dominante qui tend a être appliquée même en cas de graphies inconsistantes contextuelles minoritaires. Quant à la dérivation par la morphologie, les graphies indérivables par la morphologie sont mieux réussies que les graphies dérivables par la morphologie. Cette tendance est plus prononcée chez les enfants avec difficultés que chez les enfants sans difficultés. Ces résultats permettent de supposer que les deux groupes d'enfants ne tendent pas à utiliser la stratégie de dérivation par la morphologie.

Pour l'orthographe française les principaux résultats concernant les stratégies développées peuvent se résumer comme suit : Le recours à la procédure orthographique joue un rôle important. Par suite, l'écriture de mots peu ou pas familiers reste difficile. Les deux groupes d'enfants ne tendent pas à utiliser la stratégie de dérivation par la morphologie. Reste pourtant à noter que les deux groupes d'enfants tendent à utiliser des règles normalement acquis de manière

implicite. Même si le niveau de performance des enfants avec difficultés se situent clairement en dessous de celui des enfants sans difficultés, ces deux groupes ne diffèrent guère quant à la nature des stratégies de traitement de l'information développées.

Pour l'orthographe allemande, le profil de performances des enfants avec difficultés diffère clairement de celui des enfants sans difficultés. Le profil de performance des enfants sans difficultés se caractérise par un plafonnement (pourcentage de réponses correctes entre 90 et 100 %) pour touts les types des graphies à l'exception des graphies peu familiers nécessitant le recours à la procédure orthographique (graphies morphologiquement indérivables et graphies irrégulières). Les enfants avec difficultés atteignent des performances élevées (entre 90 et 100%) uniquement pour les graphies basées sur des règles de correspondances phonèmes-graphèmes très transparentes (graphies acontextuelles et graphies consistantes contextuelles).

L'effet de fréquence est significativement plus important pour les enfants avec difficultés que pour les enfants sans difficultés. Ceci permet de supposer que les enfants avec difficultés tendent à recourir davantage à la procédure orthographique pour les mots dont l'orthographe a été appris de manière explicite en classe et que l'application de règles plus ou moins complexes reste à développer. Aucun des deux groupes étudiés présente un effet de dominance (graphies inconsistantes contextuelles dominantes et minoritaires). Les enfants sans difficultés atteignent des performances moyennes élevées pour les graphies inconsistantes contextuelles dominantes et minoritaires. Ces données permettent de supposer qu'ils ont recours à la procédure orthographique. Pour les enfants avec difficultés les résultats réalisés pour les graphies inconsistantes contextuelles minoritaires dépassent ceux obtenus pour les graphies inconsistantes contextuelles dominantes. Ceci peut s'expliquer par le recours à des règles de correspondances consistantes acontextuelles via la procédure phonologique. Quant à la dérivation par la morphologie, l'importance de l'effet est comparable pour les enfants avec difficultés et les enfants sans difficultés. Le fait que pour les deux groupes d'enfants les graphies dérivables par la morphologie sont mieux réussis que les graphies indérivables par la morphologie montre que les deux groupes ont recours à la dérivation par la morphologie.

Pour l'orthographe allemande, les principaux résultats concernant les stratégies développées peuvent se résumer comme suit : Contrairement à l'orthographe française, les stratégies utilisées par les enfants avec difficultés diffèrent de celles des enfants sans difficultés. Les enfants sans difficultés. Comparé aux enfants sans difficultés, les enfants avec difficultés semblent avoir plus de difficultés à écrire des mots peu ou pas familiers. Ceci peut s'expliquer par une difficulté à appliquer les règles de correspondance plus ou moins systématiques. Si les enfants avec difficultés ont recours à la procédure phonologique, l'application de règles de correspondances phonèmes-graphèmes transparentes domine sur l'application de règles de correspondance plus complexes. La dérivation par la morphologie est appliquée dans une même mesure par les enfants avec difficultés que par les enfants sans difficultés.

Conclusion finale

Au delà du niveau de performance, l'analyse des profils de performance informe sur les stratégies développées par les différents groupes d'enfants étudiés. Les résultats présentés montrent que l'analyse des profils de performances permet de dresser un tableau plus nuancé que la simple analyse des niveaux de performances.

Quant aux implications pratiques, les résultats de l'épreuve française ont mis en évidence l'importance du rôle joué par la procédure orthographique. Par suite, il semble favorable de mettre d'avantage l'accent sur la conversion phonologie-orthographe via la procédure phonologique et sur la dérivation par la morphologie. En effet, ces stratégies facilitent l'écriture de mots peu ou pas familiers.

L'épreuve allemande montre que pour le groupe d'enfants avec difficultés il est intéressant de développer d'avantage la conversion phonologie-orthographe via la procédure phonologique notamment pour les règles de conversion plus complexes.

Références

Frith, U. (1985). Beneath the surface of developmental dyslexia. Are comparisons between developmental and acquired disorders meaningful ? In K.E. Patterson, J.C. Marshall & M. Coltheart (Eds.), *Surface Dyslexia* (pp.301-330). London: Erlbaum.

Landerl, K., Wimmer, H., & Moser, E. (1997). Salzburger Lese- und Rechtschreibtest. Verfahren zur Differentialdiagnose von Störungen des Lesens und Schreibens für die 1. bis 4. Schulstufe. Bern: Huber

Marx, H. (1997). Erwerb des Lesens und des Rechtschreibens : Literaturüberblick. In F.E. Weinert & A. Helmke (Eds.), *Entwicklung im Grundschulalter* (p.85-111). Weinheim: Psychologie Verlags Union.

May, P. (1996). Hamburger Schreibprobe zur Erfassung der grundlegenden Rechtschreibstrategien. Hamburg: verlag für pädagogische medien.

Mousty, P., Leybaert, J., Alegria, J., Content, A., & Morais, J. (1994). BELEC : Une batterie d'évaluation du langage écrit et de ses troubles. Dans J. Grégoire & B. Piérart (Eds.), *Evaluer les troubles de la lecture : Les nouveaux modèles théoriques et leurs implications diagnostiques* (pp.127-145). Bruxelles : De Boek Université.

La Validation des Acquis de l'Expérience : vers un renouvellement des regards sur l'auto-évaluation

Borhène Chakroun, ENESAD

Mots-clés : VAE, auto-évaluation, pratiques réflexives

Résumé:

L'étude des processus d'auto-évaluation s'inscrit principalement dans le cadre des travaux sur l'évaluation formative. La revue des travaux réalisés dans le cadre de l'ADMEE-Europe permet de dégager diverses perspectives (Allal, 2004). Dans ce cadre, la VAE apporte un éclairage différent sur les formes et processus auto-évaluatifs en proposant d'une part, d'inscrire l'auto-évaluation dans un processus d'analyse de l'expérience en vue d'obtenir un diplôme et d'autre part, d'envisager un nouveau rapport entre l'auto-évaluation et la pratique réflexive.

En analysant la façon dont les candidats à la VAE reviennent sur leurs acquis expérientiels, le chercheur est confronté à la place prépondérante des processus auto-évaluatifs et réflexifs. L'auto-évaluation est une partie du travail prescrit aux candidats et les dossiers VAE élaborés par les Universités Françaises incitent les candidats à adopter une posture « réflexive » à travers la présentation d'une situation-problème.

L'activité réflexive, dans le cadre de la VAE, constitue un objet d'étude parce qu'elle semble accompagner des formes d'auto-évaluation qui nécessitent, entre autres, des actions de sélection, de hiérarchisation et d'organisation (Mayen, 2005). Certains auteurs considèrent que la réflexion est souvent imbriquée dans une auto-évaluation ((Paquay et al. 2001, Jorro 2004).

Dans cette communication, je m'attacherai, à partir du cadre théorique de la didactique professionnelle, à mettre en évidence les formes auto-évaluatives et réflexives mobilisées par les candidats à la Validation des Acquis ainsi que les conceptions organisatrices qui les sous-tendent (Vergnaud, 1996 ; Mayen, 2005).

Le travail présenté tire parti de recherche, en cours, sur le travail des candidats dans le cadre de la VAE. Le corpus analysé porte sur les versions intermédiaires et finales du dossier de VAE élaborés par les candidats ainsi que les retranscriptions des entretiens d'accompagnement.

La contribution expose les premiers résultats issus d'une analyse thématiques de ces traces. Le choix initial des thèmes est fondé sur les travaux de Campanale (1995).

L'AUTORITÉ DU "BON PROFESSEUR": EVALUATION DES ATTITUDES ET COMPORTEMENTS FAVORABLES AU RESPECT DES REGLES

Françoise Welbes, Ecole Primaire Cents

Robert A.P. Reuter, Université du Luxembourg

Mots-clés: autorité, respect des règles, représentations

Résumé :

L'autorité est une caractéristique d'un rapport social et elle présuppose le respect réciproque. La présente recherche vise à relever les attitudes et comportements d'un enseignant favorables à son autorité. Selon Rey (1999), ni une attitude de supériorité, ni une attitude de proximité ne sont favorables à l'autorité d'un enseignant, mais plutôt celle qui permet de résoudre de façon convaincante et satisfaisante cette dialectique entre éloignement et proximité, appelé ici « positionnement à distance ». Cette attitude est favorable au sentiment de l'élève d'être respecté par son professeur, ce qui engendre le respect des règles établies par l'enseignant et donc la reconnaissance de son autorité. Nous proposons ici d'évaluer laquelle de ces trois attitudes est perçue par les élèves comme la plus favorable à l'autorité. Nous avons ainsi établi des définitions opérationnelles de ces attitudes en les associant à trois catégories de comportements: communication congruente/non-congruente, mode de réaction non-affirmatif/affirmatif/hostile et traitement juste/injuste. Afin de vérifier l'hypothèse d'un lien étroit entre positionnement à distance et le couple respect-autorité, nous avons réalisé des entretiens semi-dirigés auprès de jeunes adolescents d'un établissement d'enseignement secondaire technique du Grand-Duché de Luxembourg. Les résultats obtenus suggèrent que les adolescents interrogés accordent de l'importance au fait que le professeur les respecte et qu'il adopte un positionnement à distance. Selon leur propre évaluation, c'est le sentiment d'être respectés par le professeur qui les conduit à respecter ses règles. Cependant, les propos des jeunes soulèvent également l'importance d'un rapport de négociation, d'une zone de tolérance et de flexibilité par rapport à la loi et au savoir, ainsi qu'un ensemble de représentations du système scolaire qui semblent s'éloigner de la définition des missions et des finalités de l'école et rapprochent l'autorité à leur interprétation du « bon professeur ».

Introduction

Qu'est-ce qui dans l'attitude d'un enseignant fait que les élèves respectent les règles et les commandements auxquels il les soumet? L'obéissance aux règles est souvent rattachée à la personnalité de l'individu, de son physique, de son âge ou de la matière enseignée. L'autorité est ainsi conçue comme naturelle ou innée, comme un don que certains possèdent et d'autres pas.

Dans le présent article, nous allons développer les hypothèses que l'autorité (A) est une caractéristique d'un rapport social (Rey, 1999), et que l'autorité exige le respect de règles (R), c'est-à-dire que si l'élève respecte les règles de l'enseignant, ce dernier a de l'autorité.

En nous basant sur la théorie d'Eirick Prairat (2004) que l'autorité exige le respect réciproque, nous supposons que pour que l'élève respecte les règles énoncées par son professeur, il doit éprouver un sentiment de respect (SR) pour celui-ci. Seul l'élève qui respecte son professeur est aussi en mesure de respecter ses règles et son autorité. Corollairement, il est important que l'élève se sente respecté par son professeur. Le sentiment d'être respecté par une personne mène au respect de l'autre, et donc à la reconnaissance de son autorité. Le respect réciproque joue donc un rôle important dans le processus de l'autorité.

Pour développer ce sentiment de respect, l'enseignant doit s'adresser au jeune en adoptant une attitude respectueuse, qui est ainsi favorable à son autorité. Qui dit respect dit limite (Prairat, 2004). Une attitude dite « de limite » semble donc la plus propice pour que l'enseignant développe le sentiment d'être respecté chez l'élève et pour que ce dernier éprouve du respect pour le professeur et par conséquent reconnaisse son autorité.

Concernant notre questionnement initial « Qu'est-ce qui dans l'attitude de l'enseignant fait que les élèves respectent les règles et les commandements auxquels ils sont soumis? », Rey (1999) distingue deux types d'attitudes – de supériorité et d'égalité. Il conclut cependant qu'aucune des deux dispositions à elles seules ne permet d'obtenir de l'autorité. D'où notre proposition de rajouter une troisième manière de se positionner en tant qu'enseignant à l'égard de l'élève, que nous avons qualifiée de « positionnement à distance » (PD), basé sur la dialectique éloignement/supériorité et proximité/égalité. Nous émettons l'hypothèse que ce positionnement à distance est un positionnement respectueux, c'est-à-dire entraînant un sentiment de respect auprès des jeunes, et qu'il est ainsi le seul à favoriser l'autorité du professeur.

Afin de définir clairement ces trois types de positionnements, nous nous sommes basés sur les apports de Ginott (cité par Charles, 1997), de Canter (cité par Charles, 1997) et de Prairat (2004). Nous avons repris les apports de Ginott sur la communication congruente coïncidant avec une bonne discipline en classe, et correspondant à une manière respectueuse de s'adresser aux élèves. La

communication congruente renvoie par ailleurs à la vertu morale de la notion de respect (Prairat, 2004), à savoir le respect de la dignité d'une personne. Canter, quant à lui, insiste sur la réaction adéquate au mauvais comportement, parce que seulement une intervention face au comportement inapproprié indique qu'il y a des limites. Il faut cependant veiller à ce que cette façon de réagir soit respectueuse, décisive et conséquente, d'où l'importance d'un mode de réaction par affirmation de soi. Enfin, nous nous inspirons de Prairat (2004) et de la vertu sociale de la notion de respect, la justice.

Ainsi, nous avons réussi à définir une attitude respectueuse d'un enseignant face à ses élèves, basée sur la dialectique éloignement – proximité. Le sentiment de respect suscité auprès du jeune fait que ce dernier respecte son professeur et donc les règles auxquelles il est soumis. Comme le respect des règles équivaut à l'autorité, ce positionnement à distance implique véritablement l'autorité du professeur.

Méthode

Sur base de ces considérations théoriques, nous avons réalisé des entretiens semi-dirigés avec 10 jeunes adolescents d'un établissement d'enseignement secondaire technique du Grand-Duché de Luxembourg, âgés de 15 ans. Ces entretiens se divisent en deux parties, d'une part des questions ouvertes sur les conceptions des jeunes sur l'autorité et le respect, et d'autre part des situations-exemples qui demandent aux jeunes de choisir une proposition parmi d'autres. Ces alternatives représentent des illustrations des trois dimensions du respect (communication - mode de réaction - justice) mises en rapport avec les trois types de positionnements (égalité – distance – supériorité). Pour chaque situation, nous avons demandé aux élèves, à quel type d'enseignant ils pensent obéir, quel type les respecte le plus et comment eux en tant que professeur aurait réagi dans la situation en question.

Ces entrevues permettent, d'une part, de détecter l'efficacité d'un PD, d'un PE et d'un PS par rapport à l'autorité du professeur et au sentiment de respect suscité auprès de l'élève et, d'autre part, elles permettent de dégager la relation entre autorité et respect. Nous cherchons ainsi à savoir si le sentiment d'être respecté est une condition suffisante et/ou nécessaire à l'autorité voire au respect des règles et quels autres comportements/propriétés peuvent intervenir dans le processus de l'autorité.

Dans une première phase, nous avons analysé les résultats obtenus aux deux parties de l'entretien (questions ouvertes et situations exemples). Les réponses à la première partie de l'entretien sont classées en fonction d'une grille d'analyse qualitative des contenus, définie précédemment. Les réponses à la deuxième partie sont d'abord soumises à une analyse quantitative, puis à une analyse plus approfondie des propos des jeunes.

Dans une deuxième phase, nous avons mis ensemble les résultats obtenus aux deux parties du questionnaire afin d'en dégager des associations et des dissociations. Cette deuxième phase se prolonge par l'analyse des propos des jeunes qui ne peuvent être classés à l'aide de notre grille d'analyse préétablie.

Résultats

Les résultats des entretiens confirment l’ensemble de nos prédictions. Les jeunes que nous avons interrogés pensent effectivement qu’il est important que le professeur les respecte et adopte un positionnement à distance. Donc l’attitude d’un professeur de se mettre à distance par rapport à ses élèves n’est pas seulement une manière respectueuse de se positionner par rapport aux jeunes, mais que cette disposition est également favorable au respect des règles et instructions auxquelles il les soumet. Retenons comme comportements révélateurs une prise en compte réelle des capacités, des potentialités et des difficultés des élèves, une attitude bienveillante et sympathique, une disponibilité à l’écoute des élèves mais évitant toute sorte d’intimité, le respect de la dignité, une réaction adéquate, décisive et conséquente au comportement inapproprié, et une attention égale réservée à chaque membre de la communauté. Les propos des jeunes soulignent en effet le positionnement de supériorité comme n’étant ni favorable à l’autorité, ni au sentiment d’être respecté par le professeur. Retenons ainsi une attitude antipathique, un non-respect de la dignité, des réactions hostiles, inadéquates et injustes par rapport aux élèves.

Le positionnement à distance étant favorable aussi bien au respect des règles qu’au sentiment de respect implique que le sentiment d’être respecté est une condition favorable à l’autorité. Les jeunes considèrent cette association comme importante et suffisante, cependant pas toujours nécessaire. Des élèves admettent obéir à un professeur qui ne les respecte pas, cependant pas « par respect » mais par peur et en vue d’éviter toute conséquence négative. La soumission à l’autorité du professeur n’est plus à base volontaire mais à base forcée.

Les élèves relèvent cependant aussi d’autres propriétés du professeur liées à l’autorité comme sa personnalité, son caractère, son âge, son physique ainsi que sa façon de s’habiller et de se soigner. La matière enseignée et la capacité du professeur à la transmettre jouent également un rôle important dans le processus de l’autorité. De plus, les réponses des élèves suggèrent qu’il peut avoir de l’autorité par d’autres moyens, qui sont surtout liés à la réussite scolaire et à leurs représentations du « bon professeur ».

Discussion

Une attitude basée sur la dialectique éloignement - proximité, donc le positionnement à distance, n’est pas seulement une manière respectueuse de la part du professeur de s’adresser aux jeunes, mais cette attitude est également favorable au respect des règles et instructions auxquelles il les soumet. Le respect de l’enseignant pour son élève, et donc, le sentiment de l’élève d’être respecté par son professeur, est une condition importante et suffisante à son autorité.

Nous avons essayé de définir les attitudes (et comportements) d'un enseignant favorables au respect des règles auxquelles il soumet ses élèves en partant de l'axiome que « l'autorité exige le respect réciproque » (Prairat, 2004). L'attitude d'un enseignant - déterminant sa manière de se positionner face à ses élèves - peut se définir par une série de conduites que nous avons regroupées sous forme de trois positionnements (égalité – distance – supériorité), En effet, selon les représentations des jeunes, une attitude de positionnement à distance serait favorable aussi bien au respect et donc à l'autorité de l'enseignant, qu'à leur propre sentiment d'être respectés. Ainsi une « communication congruente » (Ginott, cité par Charles, 1997), un « mode de réaction par affirmation de soi » (Canter, cité par Charles, 1997) et un traitement juste (Prairat, 2004) permettent aux jeunes d'éprouver le sentiment d'être respectés par leur professeur. Nos résultats confirment également l'inefficacité d'un positionnement de supériorité en vue de l'obtention d'un bon climat en classe, favorable aux apprentissages scolaires.

Les propos des jeunes indiquent également d'autres variables comme intervenant dans le processus de l'autorité, telles que la matière enseignée, les compétences et les connaissances du professeur, ainsi que l'âge et l'expérience de l'enseignant (Gordon cité par Charles, 1997; Rey, 1999). Selon Gordon, l'autorité basée sur la compétence découle « des connaissances spécialisées, de l'expérience, de la formation, des habiletés, de la sagesse et de l'éducation d'une personne » (p.205). Gordon considère que ce type d'autorité a une influence positive sur les élèves. Les jeunes aperçoivent également l'âge de l'enseignant comme important dans le processus de l'autorité. Ils considèrent un professeur plus âgé comme plus expérimenté, mieux capable donc de leur faire acquérir le savoir nécessaire à la réussite scolaire. L'âge et l'expérience du professeur peuvent être rapprochés du type d'autorité 'C' (compétence) de Gordon. Rey (1999) suppose aussi que l'autorité du professeur repose sur la compétence que l'élève lui reconnaît. Rey considère que la capacité de transmettre le savoir est plus déterminante pour les élèves que le fait que l'enseignant a des compétences et des savoirs importants. La légitimité du professeur ne tient pas à « l'étalage du savoir », mais à sa capacité de partager ce savoir. Dans ce sens, Rey soulève encore un autre aspect de la compétence, celui de la persévérance. Il est important que les élèves pensent que l'enseignant a envie de leur faire apprendre des choses. Pour Rey, la compétence est la forme de légitimité la plus conforme à la fonction de l'enseignant. L'étalage du savoir n'est en effet pas l'aspect le plus important pour le jeune, ni son acquisition. Il suffit que l'enseignant dispose des quelques connaissances dont il a besoin afin de faire partager le savoir nécessaire, prévu au programme, et plus spécialement le savoir faisant l'objet de l'évaluation du jeune. D'où la nécessité du professeur d'être doté de la capacité de transmettre ce savoir au jeune, et comme le dit Rey, de posséder une certaine persévérance, visant à les aider à apprendre. Il y a un élève qui parle à ce propos de la « patience » nécessaire au professeur pour expliquer la matière.

Ce qui ressort particulièrement des propos des jeunes est l'importance que l'enseignant réussisse à établir un climat favorable à l'étude et plus précisément à l'apprentissage de la matière prévue au programme et faisant objet d'une interrogation. Ainsi le professeur doit tenir compte des potentialités de chaque élève,

en adaptant le cours à ses capacités et faiblesses. Les élèves obéissent au professeur principalement en vue de réussir leur année scolaire. Il est aussi très important pour eux que leur professeur les interroge uniquement sur la matière traitée au cours. Ils détestent devoir étudier des pages et des pages non-traitées parce que le professeur n'a pas eu le temps de faire tout son programme. Nous pouvons ainsi constater que les descriptions d'un professeur à qui les jeunes pensent obéir correspondent à la description d'un professeur idéal qui à son tour correspond au professeur qui respecte ses élèves. Les adolescents mettent en relation l'obéissance au professeur avec la réussite scolaire et donc moins avec leur propre sentiment d'être respectés. Ce dernier n'est cependant pas totalement exclu, parce que le souhait que l'enseignant amène le jeune à la réussite scolaire, par une interrogation de la matière traitée au cours, est une attitude de respect pour l'élève attendue de sa part, et ce respect engendre en retour celui accordé au professeur. Les élèves n'énumèrent donc pas des comportements qui sont vraiment nécessaires à l'autorité, mais plutôt des attitudes favorables à être désigné comme un « bon » enseignant. L'entrevue que nous avons eu avec les élèves, mobilise par conséquent principalement les représentations des jeunes par rapport au « bon » et/ou « mauvais » professeur et donc leurs représentations générales du système scolaire, plutôt que les attitudes et comportements favorables à l'autorité proprement dite.

L'obéissance au professeur se fait donc dans le but de réussir l'année scolaire, plutôt par séduction, afin d'obtenir de bons points ou de plaire à l'enseignant, et donc pas vraiment dans le but d'acquérir du savoir, du savoir-faire et du savoir-être, ce qui constitue quand même la mission fondamentale de l'école. La représentation que le jeune a du système scolaire n'est donc pas en accord avec la définition des missions et des finalités de l'école. Nous nous demandons cependant s'il n'existe pas un certain parallélisme entre la représentation du jeune et ce qui se passe actuellement dans nos écoles. Est-ce que l'école actuelle ne tend pas à se positionner à proximité de l'élève? Ne cherche-t-elle pas à séduire le jeune en vue de le faire entrer dans le monde des savoirs?

Rey (1999) soulève notamment que les jeunes professeurs sont souvent tentés d'effacer les signes de supériorité à l'égard des élèves afin d'établir une relation plus amicale avec eux. Ils se font tutoyer et appeler par leur prénom. Nous avons interrogé les jeunes sur cette attitude du professeur et ils considèrent le tutoiement comme bizarre, parce qu'inhabituel, et le vouvoiement d'un professeur est mis en parallèle avec le respect d'une personne. En tenant compte de ces résultats nous pouvons affirmer à nouveau notre hypothèse d'une distanciation comme favorable à l'autorité de l'enseignant. Ce dernier doit proposer ce que nous appelons une relation partiellement ouverte. Nous percevons également une contradiction entre les propos du jeune adolescent et ceux du jeune enseignant. L'élève ne cherche pas cette proximité exprimée par le jeune professeur.

Une question que nous nous posons à cet égard, et en même temps suite aux résultats obtenus à la présente recherche, est de savoir si le fait de manifester une préférence claire pour le positionnement à distance ne reflète pas une attitude nécessaire au développement psychologique du jeune adolescent?

Psychologiquement, ce n'est qu'en se détournant de l'adulte que l'adolescent peut construire sa propre identité. Supposons que parallèlement l'adolescent doit se détourner de l'enseignant pour construire le savoir. En effet, nous pouvons émettre l'hypothèse que le jeune ne souhaite pas être instruit par un maître-camarade. Celui-ci ne semble pas réussir à créer un climat favorable à l'étude. Le jeune ne pense pas pouvoir acquérir le savoir nécessaire à l'étude dans de telles conditions. Du point de vue des connaissances, le maître-camarade serait ainsi moins estimé par le jeune. Bernard Rey (1999) soulève également un danger inhérent au positionnement de proximité à l'égard des élèves: la sous-estimation par les élèves des exigences propres au savoir. En effet, l'acquisition du savoir nécessite l'obéissance à des procédures rigoureuses. Dans une relation amicale, il est souvent difficile aux élèves de se mettre au travail. Ils sont tentés de reprendre les activités ludiques qu'ils partagent habituellement avec leurs camarades.

Une relation de confiance construite sur la base du respect réciproque est importante et suffisante, mais pas nécessaire à l'autorité du professeur. Le respect est une aide, mais ce qui importe pour être un professeur idéal, voire un professeur ayant de l'autorité, ce sont la capacité et la volonté de transmettre des savoirs adaptés au jeune.

Le savoir n'est cependant pas toujours considéré comme une donnée intrinsèque, mais son acquisition est basée sur un affrontement avec l'adulte. Le savoir est considéré comme une série d'affirmations arbitraires. Rey (1999) souligne que le rapport que les jeunes entretiennent avec les savoirs joue un rôle important dans la discipline en classe. Le but n'est plus dans certains cas de comprendre le savoir mais de plaire ou de résister à l'enseignant. Ce rapport de force peut être mis en parallèle avec un positionnement d'égalité, respectivement de supériorité. L'enseignant peut aussi tenter de se faire respecter par ses élèves en essayant de plaire ou de résister aux jeunes. Ces deux attitudes engagent en tout cas l'élève dans une dérive relationnelle par rapport à l'enseignant – transmetteur du savoir. Yves Lorvellec (2003, p.116) souligne que l'élève du maître charismatique acquiert le savoir par l'effet de la fascination et de la séduction exercée par le maître. L'élève est soumis à l'admiration et les connaissances acquises sont ambiguës. Certes, il faut qu'un maître arrive à passionner ses élèves, mais à condition qu'il y arrive par ce qu'il enseigne et non par sa personne.

Les jeunes mettent en relation la discipline et l'ordre en classe avec l'acte d'enseigner. L'établissement de cet ordre nécessite des limites, ce qui exclut un positionnement basé purement sur l'égalité. Cette distanciation demandée au professeur est souvent décrite par une sorte de tolérance et de flexibilité, aussi bien par rapport à la loi que par rapport au savoir. Le rapport à l'autorité doit souvent être négociable. L'enseignant ne doit pas réagir au mauvais comportement de façon invariable, mais laisser à l'élève une marge de liberté. La loi apparaît au jeune comme une atteinte à sa liberté. Les élèves ont tendance à négocier la loi et à la considérer comme étant la force de l'enseignant. Selon Rey (1999), il faut amener les jeunes à interpréter la loi comme étant une nécessité. Certains élèves parviennent à prendre conscience du fait que la loi est la « convention qui permet de cohabiter ». La loi est

ce qui me donne une limite, ce qui m'empêche de faire tout ce que j'ai envie de faire, mais en même temps, cette même loi empêche autrui de faire ce qu'il veut et par là me protège (Rey, 1999, p.113). Pour que les élèves perçoivent mieux le rapport à la loi, il faut les faire participer à l'élaboration des règles de discipline en classe. Ainsi l'adolescent apprend à se conformer à une règle parce qu'elle permet la vie en commun, et parce que le jeune en comprend la nécessité.

Les élèves attachent également l'autorité d'un enseignant et leur sentiment d'être respectés par cet enseignant au « bon » et/ou « mauvais » professeur: le rapport à l'autorité est lié à une préférence relationnelle. Au stade de l'adolescence, le jeune a parfois du mal à distinguer amour et respect. Dans son langage courant, les deux notions semblent souvent équivalentes. Nous observons ainsi à travers des réponses des jeunes concernant la description d'un professeur auquel ils pensent obéir, respectivement ne pas obéir, que le sentiment d'être respectés et l'autorité d'un professeur sont souvent mis en parallèle avec des éléments de la personnalité de l'enseignant (gentillesse, politesse, humour). Kounin (cité par Charles, 1997, p.60) a également soulevé que la sympathie éprouvée à l'égard du professeur a une influence importante sur le comportement des élèves. S'ils ont beaucoup d'estime pour leur professeur, alors ils adoptent moins souvent un « comportement répréhensible » et commettent moins d'infractions. Ils ont tendance à s'appliquer avec une plus grande efficacité au travail.

Une autre caractéristique du professeur favorable aussi bien au sentiment d'être respecté de l'élève qu'à l'autorité du professeur est son écoute pour des problèmes scolaires et personnels. Gordon (cité par Charles, 1997) évoque aussi dans son modèle le concept d'écoute active. Il la définit comme l'attention prêtée « à ce que dit une personne et de lui faire savoir que l'on comprend ce qu'elle tente d'exprimer » (p.203). L'enseignant doit accorder une attention particulière à ses élèves, il doit les écouter. Cela ne signifie pas qu'il doit résoudre leurs problèmes à leur place. Le jeune doit « résoudre lui-même ses problèmes, mais le professeur doit le persuader qu'il est quelqu'un avec qui le dialogue est possible ». Ceci requiert de « l'authenticité, de la disponibilité et du respect de la vie privée » (Estrela, 1996, p.76). Gordon définit certaines techniques nécessaires à l'enseignant: l'écoute active, la résolution de problèmes, la confrontation, la résolution de conflits et les messages à la première personne. L'écoute active « n'implique ni l'évaluation ni les jugements de valeur ». Elle aide à clarifier les positions et à faciliter la discussion (Gordon, cité par Estrela, 1996, p.76). Le jeune demande à l'adulte de lui apporter soutien, sécurité et connaissances permettant la quête de soi et la poursuite de son développement, et en même temps il souhaite que l'enseignant le laisse tranquille pour découvrir de nouvelles choses. La tâche de l'enseignant, entre ses demandes contradictoires de l'élève, exige donc une certaine flexibilité et tolérance. Il doit de plus en plus avoir recours à des compromis en vue de fonder son autorité.

Problèmes conceptuels

Au sein de notre recherche, nous avons eu du mal à définir et à distinguer certains concepts. D'abord, nous avons eu du mal à distinguer deux positionnements, celui **de distance** et celui **d'égalité**. En effet, tous les deux sont définis par des concepts dont

les interprétations varient selon le lecteur. Il serait donc important, en vue de prolonger cette recherche, de redéfinir les positionnements d'égalité et de distance pour mieux pouvoir les distinguer.

Nous avons également eu du mal à définir quelques concepts liés aux différents positionnements comme celui du **professeur gentil**. Qu'est-ce que cela signifie vraiment pour les élèves? Il fallait donc insister sur des définitions plus claires afin que chaque lecteur puisse interpréter de la même manière les différentes notions employées. Il aurait été important de réduire cette subjectivité par la relecture des interviews par une personne extérieure. Nous avons essayé de déterminer la signification des notions dans la partie concernant les explications de notre grille d'analyse. Ces commentaires pourraient être retravaillés ou complétés.

Une autre distinction qui posait particulièrement problème était celle des **concepts d'autorité et de respect**. Nous avons défini l'autorité d'un enseignant comme équivalant au respect de ses règles. Pour se soumettre volontairement aux règles établies par son enseignant, le jeune doit les respecter, parce qu'elles sont fonctionnelles, constituant un discours réglé. Il doit respecter en plus la « supériorité » du professeur qui le soumet à ses règles, vu qu'il s'agit d'une relation d'autorité inégale et hiérarchique. Rappelons qu'il est clair que l'élève doit se sentir respecté par son professeur pour éprouver du respect pour lui. Au sein de cette recherche, nous nous intéressons particulièrement à cette troisième condition, parce que l'autorité de l'enseignant nous semble se fonder particulièrement dans le respect de sa personne, surtout lors de la période de l'adolescence. Comme l'autorité exige le respect réciproque (Prairat, 2004), nous avons défini une attitude qui suscite chez le jeune le sentiment d'être respecté par le professeur. Ce sentiment de respect éprouvé, engendre de la part de l'élève du respect pour le professeur et donc le respect de ses règles. Le positionnement à distance, par lequel nous avons défini cette attitude respectueuse de la part de l'enseignant, devrait favoriser ce sentiment de respect de l'élève, et par là constituer le fondement de l'autorité de l'enseignant, donc l'obéissance volontaire aux règles établies par le professeur.

Ajoutons une dernière remarque concernant le sujet de **l'autorité**. C'est un concept difficile à explorer avec les adolescents et comme les résultats l'indiquent, c'est un concept qui n'est pas encore vraiment aperçu par le jeune. Il y a souvent des contradictions à éclairer. Il reste toujours des hésitations entre maître-camarade et maître distant. Nous pensons a fortiori qu'il s'avérerait difficile d'interroger des élèves du primaire à ce sujet, puisque leur capacité de réflexion n'est pas assez développée pour s'exprimer sur les concepts d'autorité et de respect.

Problèmes méthodologiques

Concernant la **méthodologie** du présent travail, nous considérons la technique de l'entrevue semi-dirigée comme étant un bon moyen pour dégager dans un premier temps les représentations d'autorité et de respect des jeunes de 15 ans. Pourtant nous constatons que notre entretien permet plutôt de dégager la représentation du jeune par rapport au système scolaire et à son professeur idéal, que les attitudes et comportements favorables à l'autorité proprement dite. Afin de remédier à cette

confusion au niveau de ce que notre outil de recherche mesure réellement, il serait donc favorable soit (1) de partir d'un cadre conceptuel différent de celui de Ginott – qui définit deux types d'enseignant: le professeur « idéal » et le « pire » enseignant – (cité par Charles, 1997), soit (2) de recourir à un autre outil méthodologique, comme l'observation directe sur le terrain. Celle-ci nous permettrait mieux de dégager ce qui se passe vraiment au sein des classes et de savoir ainsi quels types de comportements conduisent effectivement au respect des règles. Nous proposons à cet égard l'observation d'une panoplie de situations filmées au sein de classes. Parmi ces situations doivent figurer les trois types de positionnements (égalité – distance – supériorité). Des observateurs indépendants, ne connaissant pas notre plan et objectif de recherche, devraient observer la situation qui leur est présentée et compléter une grille d'observation comprenant une série de comportements. L'analyse à l'aide de cette grille d'observation permettrait de vérifier si l'autorité et le respect sont présents dans la situation filmée et quels comportements respectivement quel(s) positionnement(s) du professeur mènerai(en)t à son autorité et au respect des règles. De plus, nous pourrions observer à la fois élèves et enseignants et vérifier la cohérence de leurs propos.

Concernant notre **outil de recherche** proprement dit, nous constatons que les résultats obtenus aux deux parties de l'entretien sont équivalents ce qui réduit l'utilité de faire passer les deux options. Mieux vaut alors se concentrer sur l'une des deux et creuser pour obtenir plus d'informations. Nous percevons la technique des situations-exemples comme plus efficace parce que cette méthode est plus rapidement administrable et parce qu'elle permet une interprétation plus rapide et moins subjective. En effet, les situations-exemples peuvent entre présentées sous forme d'un questionnaire à cocher. Ainsi elles pourraient être administrées à un échantillon plus large. D'autre part, dans le cadre d'une entrevue semi-dirigée, les situations-exemples mettent les élèves plus timides ou un peu angoissés plus à l'aise. Il faudrait cependant retravailler les situations-exemples afin d'arriver à une interprétation commune et de réduire le caractère caricatural de certaines situations.

Notons finalement qu'il pourrait être intéressant de prolonger cette recherche en ayant recours à une **méthodologie expérimentale**, qui consisterait à introduire un « traitement » spécifique - supposé favorable à l'autorité du professeur - et à observer les effets auprès des enseignants et des élèves. Au groupe expérimental, on donnerait des instructions claires de se comporter dans la lignée du positionnement à distance, alors qu'au groupe contrôle, on ne donnerait aucune instruction spécifique. L'évaluation des effets d'un tel « traitement » peut se faire soit par observation directe sur le terrain, soit par questionnaire. Ainsi nous pourrions aboutir à une démonstration plus directe de l'efficacité du positionnement à distance.

Références

Charles, C. M., (1997). *La discipline en classe: modèles, doctrines et conduites*, Paris, Bruxelles, De Boeck & Larcier.

Estrela, M. T., (1996). *Autorité et discipline à l'école*, Paris, ESF.

Lorvellec, Y., (2003). Savoir et autorité, in J. Lombard, *L'école et l'autorité: Etudes réunies et présentées par Jean Lombard*, Paris, L'Harmattan, 103-126.

Prairat, E., (2004). *Questions de discipline à l'école et ailleurs…*, Ramonville Saint-Agne, Éditions érès.

Rey, B., (1999). *Les relations dans la classe: au collège et au lycée*, Paris, ESF.

LES PROCESSUS DE PENSEE DES ENSEIGNANTS DURANT LA PLANIFICATION ET L'IMPLEMENTATION DE LEUR ENSEIGNEMENT

Philippe Wanlin, Université du Luxembourg (FLSHASE, EMACS, chercheur)

Cindy Bodeux, Université de Liège (FAPSE, étudiante)

Mots-clés : Teacher thinking, Planification, Interaction, Décisions, Antécédents.

Résumé:
Cet article subdivisé en trois parties tente d'apporter un éclairage complémentaire sur les processus de pensée des enseignants lors des phases de planification et de mise en œuvre de leur enseignement. Suite à une brève revue de la littérature internationale relative au *teacher thinking*, les processus de pensée lors de la planification et de l'implémentation des activités pédagogiques de futurs instituteurs et d'instituteurs en activité de l'enseignement primaire de la Communauté française de Belgique sont analysés. Les données rassemblées auprès de ces deux publics permettent d'analyser les types et la fréquence des décisions planificatives et interactives, de même que les arguments qui les expliquent. Enfin, les résultats de cette recherche sont discutés en regard des constats répertoriés dans la littérature internationale notamment dans l'optique d'examiner le dispositif méthodologique utilisé.

Introduction

Cet article, qui présente une étude qualitative de plusieurs cas, poursuit deux objectifs principaux. Il tente, d'une part, d'apporter des connaissances supplémentaires sur la manière dont les maîtres s'y prennent pour mener à bien leur travail et, d'autre part, de montrer pourquoi et comment l'activité d'enseignement est, sur le plan cognitif, complexe et exigeante. L'ambition ultime étant de fournir éventuellement aux chercheurs, aux formateurs d'enseignants, que ce soit pour la formation initiale ou continuée, aux décideurs politiques et aux enseignants eux-mêmes, une perspective complémentaire et une approche nouvelle relatives à la professionnalité enseignante. La visée est également de montrer la complexité de l'acte d'enseignement auquel sont confrontés les enseignants qui prennent une légion de décisions différentes dans l'exercice de leur profession. A aucun moment il ne fut projeté de présenter quelque recommandation spécifique sur la manière de dresser un plan de leçon ou de prendre une décision en cours d'enseignement.

Il est nécessaire de garder en mémoire, durant la lecture de cet article, que deux postulats de base ont transcendé les observations. Selon le premier postulat, des activités de lecture renvoyant à des situations d'enseignement relativement différentes entraînent des patterns décisionnels similaires entre enseignants. Le deuxième postulat pose que ces patterns décisionnels ne diffèrent guère en fonction du cycle dans lequel l'activité est implantée. En effet, ces postulats reposent sur les constats issus de la littérature de recherche sur le *teacher thinking*. Comme nous le verrons

plus loin, qu'il s'agisse de la pensée planificative ou interactive, les recherches ont abouti à des constats remarquablement similaires et ce, qu'elles aient été menées dans des situations (petits groupes, grands groupes, collectivement, …), dans des branches (mathématique, français et sciences, …) ou à des degrés d'enseignement (secondaire, primaire, …) différents. Ces postulats restent valables tant pour les enseignantes que pour les futures enseignantes (normaliennes) qui ont participé à notre recherche.

Ceci dit, cette étude peut éventuellement autoriser des constats au niveau de la comparaison selon le statut (enseignantes *vs*. normaliennes), d'un côté, et au niveau du dispositif de recherche, de l'autre. Soulignons que le but n'est pas de généraliser nos constats à l'ensemble de la population enseignante, mais plutôt de mettre en lumière quelques tendances intéressantes en vue d'alimenter les débats scientifique, didactique et politique relatifs à la formation initiale et continuée des instituteurs, à la méthodologie de la recherche sur la pensée enseignante et à l'innovation pédagogique et curriculaire.

Enfin, les observations autorisent la formulation de quelques hypothèses qui poussent notamment à s'interroger sur l'intérêt potentiel de redéfinir les stages pédagogiques, leur déroulement et leurs critères d'évaluation ainsi que le rôle des encadrants des normaliens (maître de stage, formateur d'instituteur, …) au sein de la formation initiale des enseignants, en vue de permettre une certaine liberté en matière de prise de décisions de la part des normaliens œuvrant en faveur du développement optimal de leur professionnalité.

Le courant du *teacher thinking* : une brève revue de la littérature

Les chercheurs qui se sont intéressés au *teacher thinking* ont tenté de décrire la pensée enseignante notamment lors de la planification des activités pédagogiques et de l'interaction en classe. Les paragraphes qui sont proposés dans cette partie présentent une palette non exhaustive de recherches ayant traité de ces deux aspects.

La pensée enseignante durant la planification

Types et fonctions de la planification

Les enseignants consacrent une énergie considérable à la structuration, à l'organisation et au management préalables du temps d'instruction en classe (Clark & Peterson, 1986). Ils utilisent simultanément plusieurs types différents de planification (Yinger, 1977 ; Clark & Yinger, 1979b) qui sont emboîtés les uns dans les autres et qui interagissent de manière complexe (Clark & Lampert, 1986). Ces types de planification renvoient à une unité temporelle (planification journalière, hebdomadaire, annuelle, …) ou à une unité de contenu (l'activité, l'unité ou le chapitre).

Les recherches montrent que les enseignants planifient leur action pédagogique pour répondre à des besoins personnels et immédiats tels que la réduction de leur incertitude face à la non-prédictibilité inhérente aux situations d'enseignement ou la recherche d'une sensation de maîtrise du contenu et de l'interaction (Ben-Perez, 1975 ; Clark & Yinger, 1979b ; Clark & Elmore, 1981). Ils planifient aussi pour trouver les moyens d'atteindre les objectifs d'apprentissage (Clark & Yinger, 1979b),

pour répondre à un besoin de simplification et d'efficience du management temporel (Yinger, 1977 ; Bromme, 1982 ; Cremers & Westerhof, 1982) et pour provoquer des gains d'apprentissage dans le chef de leurs élèves (Morine, 1976 ; Clark & Peterson, 1986).

Les plans couchés sur papier ne reflètent que rarement la totalité de la planification enseignante (Morine-Dershimer, 1977, 1979 ; May, 1986). Cette dernière serait un processus d'emboîtement de routines permettant de structurer une image mentale de la leçon (Yinger, 1977 ; Morine-Dershimer, 1977, 1979 ; Smith & Sendelbach, 1979 ; Clark & Elmore, 1981 ; May, 1986). Ce plan mental, suppléé par des notes relativement télégraphiques des points importants que l'enseignant veut être certain de ne pas oublier (Morine-Dershimer & Vallance, 1976), serait reconstruit de mémoire lors de l'instruction (Smith & Sendelbach, 1979). Les intentions et l'image mentale sont alors adaptées en fonction des activités et de leur flux (May, 1986).

Contenu de la pensée planificative

Selon Yinger et Clark (Yinger, 1977 ; Clark & Yinger, 1979b ; Yinger & Clark, 1982, 1983), le processus de planification commence par une idée générale (cycle exploratoire d'interaction des représentations de l'enseignant concernant les objectifs, ses connaissances, son expérience, le matériel disponible) produisant une conception initiale d'un problème, puis par la formulation de ce problème ainsi que sa résolution, ensuite, par la visualisation de cette solution comme étant jouée dans le contexte spécifique de la classe et, enfin, par l'implémentation, l'évaluation et, éventuellement, la routinisation de cette solution. Ils continuent que cette évaluation-routinisation des plans contribue ensuite au développement du répertoire de connaissances et d'expérience de l'enseignant qui, à son tour, joue un rôle majeur dans les délibérations planificatives ultérieures.

Pour certaines recherches (3 études/5 ayant traité de cet aspect), le facteur principal de la planification sont les élèves (**tableau 1**). Les enseignants s'intéressent aussi au contenu de matière à enseigner (9/9) en y appliquant des considérations contextuelles (temps, matériel, …), en se référant à leurs connaissances et expériences personnelles, ainsi qu'en réfléchissant à la façon de rendre la situation d'enseignement intéressante pour les élèves en prenant en compte leurs intérêts et leurs attitudes. En outre, ils pensent aux objectifs des activités pédagogiques (8/10) de même qu'à la méthode ou au processus d'enseignement à utiliser pour les atteindre (3/3). Pour May (1986), nuancé par McLeod (1981), les enseignants pensent aux objectifs tant durant la phase de planification (où ils sont surtout de type cognitif) que pendant l'interaction (où ils sont surtout de type socio-affectif). De plus, les enseignants sélectionnent et organisent les activités d'enseignement-apprentissage et le matériel manipulable, gèrent les transitions d'entrée et de sortie entre les activités et l'environnement physique de la classe. Enfin, les procédures d'évaluation et leurs critères peuvent être un sujet décisionnel mais leur statut est discutable. A l'opposé, très peu d'attention serait accordée à la planification de procédures d'enseignement alternatives et aux comportements verbaux interactifs probables car les enseignants estimeraient que ces derniers sont trop incertains (Peterson, Marx & Clark, 1978 ; Carnahan, 1980).

Tableau 1 – Facteurs ou contenus de la pensée planificative des enseignants.

	Tyler (1950)	Taylor (1970)	Zahorik (1970)	Zahorik (1975)	Morine-Dershimer & Vallance (1976)	Yinger (1977)	Peterson, Marx & Clark (1978)	Carnahan (1980)	Mintz & Yarger (1980)	Clark & Yinger (1979b)	Hill, Yinger & Robbins (1981)	McLeod (1981)	Bromme & Juhl (1984)	Charlier (1986)	May (1986)	Bromme & Dobslaw (2003)	Tillema (2003)
Élèves (Besoins intérêts, habiletés et activités)	o	+	o	+	-	o	o	o	o	o	o	o	o	-	o	o	+
Contenu de matière à enseigner	o	+	+	+	o	o	+	o	+	o	o	o	+	+	o	+	+
Objectifs des activités pédagogiques	+	+	o	+	-	+	+	o	o	+	o	+	o	-	+	o	o
Méthode d'enseignement à utiliser pour atteindre les objectifs	o	+	o	o	+	o	o	o	o	o	o	o	o	+	o	o	o
Sélection et organisation des activités et du matériel manipulable	+	o	o	o	o	+	o	o	o	+	+	o	o	o	o	o	o
Gestion des transitions d'entrée et de sortie entre les activités	o	o	o	o	o	o	o	o	o	o	+	o	o	o	o	o	o
Gestion de l'environnement physique de la classe	o	o	o	o	o	o	o	o	o	o	+	o	o	o	o	o	o
Procédures d'évaluation et critères d'évaluation	+	-	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
Procédures alternatives à celles planifiées	o	o	o	o	●	o	o	o	o	o	o	o	o	●	o	o	o
Connaissances et expériences personnelles et professionnelles de l'enseignant	o	o	o	o	o	+	o	o	o	+	o	o	o	o	o	o	o
Comportements verbaux interactifs	o	o	o	o	o	o	-	-	o	o	o	o	o	o	o	o	o

Légende : + attention lors de la planification ; – pas d'attention ; ● attention rare ; o non traité dans l'étude

Soulignons enfin que les décisions planificatives que prennent les enseignants sont fortement interreliées et qu'elles s'organisent en un réseau où chaque composante entretient, avec les autres, des relations de but, de conséquence, de cause ou d'opposition, qui la dynamisent (Charlier, 1986 ; May, 1986 ; Tillema, 2003). Elles sont un déterminant majeur de ce qui sera enseigné en classe (Clark & Lampert, 1986).

Antécédents des décisions planificatives

Les antécédents décisionnels des planifications sont des représentations ou des théories personnelles relatives aux variables de la situation éducative (matière en général, compétence et intérêt des élèves, …), des éléments environnementaux, des expériences antérieures et/ou des décisions antérieures (Charlier, 1986). Les théories pédagogiques ne sont jamais mentionnées (Huberman, 1986 ; Charlier, 1986).

Pour certains (Clark & Elmore, 1981 ; Fuchs, Fuchs, Hamlett & Stecker, 1991 ; Fuchs, Fuchs, Karns, Hamelett & Katzaroff, 1999 ; Calderhead, 2003), les ressources curriculaires publiées (programmes, manuels, …) et les formations pédagogiques ont une influence puissante sur les contenus enseignés et les processus d'enseignement, pour d'autres non (Taylor, 1970 ; Smith & Sendelbach, 1979). Les manuels auraient, par exemple, un impact limité sur les contenus enseignés, car même les enseignants

qui les suivent le plus scrupuleusement ne couvrent pas la totalité de leur contenu (Freeman & Porter, 1989) et restructurent leurs connaissances, leurs décisions et le processus d'enseignement les concernant en fonction de leurs élèves (May, 1986).

Les efforts de l'enseignant pour garder le contrôle du processus d'enseignement-apprentissage, le « work ethos » (flux d'activités, participation des élèves et maintien de leur occupation), l'attachement d'une grande valeur aux relations sociales, la sensibilité de l'enseignant aux différences sociales ainsi que les tentatives d'empathie sont aussi des critères avancés par les enseignants pour expliquer leurs décisions planificatives (Halkes & Deijkers, 2003).

La pensée enseignante durant l'interaction

Contenu et fréquence des décisions interactives

Toutes les études ayant décrit le contenu des pensées interactives (Marland, 1977 ; Semmel, 1977 ; Conners, 1978 ; McNair, 1978 ; Marx & Peterson, 1981 ; Colker, 1982) ont, malgré l'utilisation de méthodologies différentes, malgré qu'elles aient été réalisées sur des branches et des niveaux scolaires différents, abouti à des constats remarquablement similaires. La pensée interactive semblerait être axée, d'abord, sur les apprenants et, ensuite, sur le processus d'enseignement (procédures et stratégies instructionnelles). Les enseignants penseraient cependant peu au contenu et à la matière lors de cette phase. Davantage de réflexion y serait accordée à des décisions managériales, à la façon de développer l'enthousiasme des élèves envers les sujets abordés et à la manière de leur proposer de l'aide afin de les amener à organiser leurs idées (Calderhead, 2003).

Le nombre estimé de décisions interactives de la part des enseignants se situe entre 0,5 et 0,7 par minute selon l'étude considérée. Il semble donc qu'en moyenne, les enseignants prennent une décision interactive environ toutes les deux minutes (Clark & Lampert, 1986 ; Clark & Peterson, 1986). À noter que le fait d'envisager des stratégies d'enseignement alternatives en cours d'action est rare (MacKay & Marland, 1978 ; Lowyck, 1980 ; Dunkin, 1986) : les enseignants restent relativement fidèles à leurs plans préalables (Clark & Lampert, 1986 ; Clark & Peterson, 1986). Ces alternatives ne sont décidées que si les multiples ajustements et petits changements habituels ne réussissent pas à maintenir le rythme et que l'activité se noie (Peterson & Clark, 1978).

Antécédents des décisions interactives

Les recherches qui ont examiné les antécédents aux décisions interactives des enseignants ont mis en évidence deux grands types d'antécédents : les indices relatifs aux élèves, d'un côté, et les indices qui ne les concernent pas, de l'autre. Pour certaines recherches, les indices relatifs aux élèves sont les antécédents de la majorité des décisions interactives rapportées par les enseignants (Cone, 1978 ; Fogarty, Wang & Creek, 1982 ; Housner & Griffey, 1983 ; Clark & Lampert, 1986 ; Altet, 1994), pour d'autres, c'est moins le cas (Marland, 1977, Shavelson, Atwood & Borko, 1977, Wodlinger, 1980). À noter que dans toutes les recherches, au moins un des antécédents aux décisions interactives renvoie aux élèves.

Les indices relatifs aux élèves renvoient soit à leur performance, soit à d'autres facteurs. Pour ce qui est de la performance, l'enseignant se réfère, quand il pense en cours d'instruction, à l'élève comme étant membre d'une unité cognitive (Bromme, 1989). Dahllöf et Lundgren (1970) ont montré que les enseignants suédois cadencent leur instruction en fonction d'un *steering group*[1] constitué des élèves les plus faibles de leur classe. Néanmoins, Wanlin (2004) a montré que l'enseignante de première primaire qu'il a observée ne mentionne pas spontanément sa centration sur la performance de certains élèves durant l'interaction. Ainsi, lorsqu'elle est interrogée à ce sujet, il ressort que, si *steering group* il y a, qu'il est constitué, dans son cas, des élèves les plus forts de la classe ou de ceux qui se situent dans la moyenne, même si elle a conscience que certains élèves ont des difficultés à suivre le rythme. Cette enseignante mentionne d'autres indices tels que la participation des élèves et leurs comportements, d'un côté, mais aussi, de l'autre, le temps alloué à l'enseignement, l'ennui mutuel, l'importance des activités et l'atteinte des objectifs. Outre la performance, Marland (1977) cite aussi trois types d'indices relatifs aux élèves, qui, dans son étude, représentent environ 44 % des antécédents des décisions interactives : les indices comportementaux (inattention, perturbations[2], …), les indices de fluidité de l'activité (réponses ou travail temporellement et qualitativement insatisfaisants des élèves) et les indices relatifs à la perception de la compréhension des élèves.

Néanmoins, les décisions interactives rapportées par les enseignants n'apparaissent pas toujours en réponse à une observation du comportement des élèves, même si ceux-ci sont un point de centration non négligeable (Marland, 1977). La sélection d'une technique d'enseignement spécifique, la sélection d'exemples appropriés, certains éléments propres à l'enseignant lui-même ou à l'environnement de la classe, les états cognitifs ou affectifs de l'enseignant, ses pensées, perceptions et/ou sensations, l'environnement y compris les contraintes temporelles, les interruptions par d'autres adultes, le matériel et l'équipement d'instruction ainsi que les jugements de la part de l'enseignant concernant l'évolution de la leçon et de la stratégie qui est en cours d'emploi auraient autant, voire plus, d'influence sur les décisions interactives (Marland, 1977 ; Shavelson et *al*., 1977 ; Wodlinger, 1980 ; Clark & Lampert, 1986 ; Wanlin, 2004). Pour nuancer, les élèves seraient surtout les antécédents des décisions interactives relatives au management d'enseignement et moins de celles concernant l'instruction (Wodlinger, 1980).

Méthodologie

Description du dispositif méthodologique

Cette recherche est composée de plusieurs études de cas qui se sont déroulées, bon an mal an, de la même manière au sein d'un dispositif méthodologique de base constitué de deux étapes-clés. La première étape des études de cas consistait en une prise de données avant interaction où le sujet était invité à noter les étapes-clés de l'activité qu'il projetait de donner (Protocole). Ce travail servait ensuite de base

[1] Le *steering group* est un petit sous-ensemble de la classe que les enseignants utiliseraient comme une référence informelle du groupe pour décider du rythme à suivre durant l'instruction.

[2] Cone (1978) a aussi montré que le type de comportement d'élève qualifié de déviant a un effet significatif sur les décisions des enseignants concernant les stratégies managériales à utiliser.

idéalement à deux types de méthodes de recueil. Il était demandé, d'une part, de commenter les notes sur une bande audio en traitant de certains éléments fixes (Réflexions), et/ou, d'autre part, de répondre à une série de questions lors d'une interview semi-dirigée (Entretien). Il s'agit là de variantes des techniques du « *thinking aloud* » et du « *journal keeping* » (Clark & Peterson, 1986).

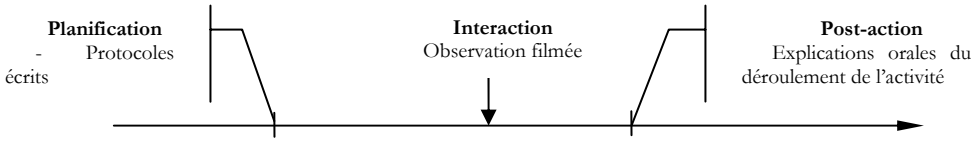


Figure 1 – Dispositif méthodologique de base des recherches présentées : cycle de recherche complet.

La seconde étape des études de cas était composée du relevé des explications du déroulement de l'activité par l'enseignant réalisé lors d'un entretien semi-dirigé généralement appuyé de l'observation filmée des activités planifiées et implantées (Clark & Peterson, 1986 ; Charlier, 1989 ; Tochon, 1993 ; Wanlin, 2003).

Tableau 2 – Description du public et de la participation ou non aux différentes phases de recherche.

	Planification			Interaction	Post-action	Nombre d'activités Observées	Expérience	Année primaire
	Protocole	Réflexions	Entretien semi-dirigé	Film	Rappel Stimulé (Entretien semi-dirigé)			
Futur enseignant 1	Oui	Oui	Oui	Oui	Oui	3 (L)	2e A.F.I.	1
Futur enseignant 2						3 (L)	2e A.F.I.	1
Futur enseignant 3						3 (L)	3e A.F.I.	2
Enseignante 1	Oui	Non	Oui	Non	Oui	2 (M ; L)	5 ans	4-5-6
Enseignante 2	Oui	Non	Oui	Non	Oui	1 (L)	5 ans	3
	Oui	Non	Oui	Oui	Oui	3 (L)		
Enseignante 3	Non	Non	Non	Oui	Oui	10 (L)	10 ans	1

Légende : L = Lecture ; M = Mathématique ; A.F.I. = Année de formation initiale

Ces deux étapes ont été répétées plusieurs fois par sujet observé. A noter que l'enseignante 3 n'a participé qu'à la deuxième étape qu'elle a répété 10 fois, que l'enseignante 1 a été rappel-stimulée sans film pour les deux activités étudiées et que l'enseignante 2 a été rappel-stimulée sans film pour l'une des quatre activités qu'elle a présentées. Ces différences dans les méthodologies utilisées nous ont permis de distinguer 3 groupes expérimentaux : le groupe 1 (G1) ayant participé aux cycles complets (les deux phases), le groupe 2 (G2) qui a participé aux deux phases sans toutefois avoir été filmé et, enfin, le groupe 3 (G3) qui a uniquement participé à la phase 2.

Description du public

Les sujets des études de cas sont tous de sexe féminin et se répartissent selon deux statuts différents qui se distinguent selon l'entrée ou non dans la vie active. D'une part, nous avons les normaliennes (N), c'est-à-dire les futures enseignantes, qui ont participé à l'étude dans le cadre de l'un de leurs stages pédagogiques au premier cycle de l'enseignement primaire de la Communauté française de Belgique. Ces normaliennes étaient en deuxième ou en troisième année de formation initiale au moment de la prise de données et ont participé aux cycles complets de la présente recherche qu'elles ont répétés 3 fois. D'autre part, nous avons les enseignantes (E) en exercice qui ont entre 5 et 10 ans d'expérience professionnelle dans la même Communauté. L'enseignante 1 se distingue des deux autres par le fait qu'elle enseigne en classe rurale multiple (cycle supérieur du primaire). A noter que les comparaisons selon le statut (N/E) sont organisées comme suit : d'une part, les normaliennes *vs.* les enseignantes 1 et 2 pour la planification et, d'autre part, les normaliennes *vs.* les enseignantes 1 et 3, pour l'interaction.

Traitement des données

Les données des entretiens, qu'elles portent sur la phase de planification ou d'interaction, ont été traitées selon une analyse de contenu de leurs retranscriptions : les extraits significatifs ont été relevés, ils ont été rangés selon les thèmes communs qu'ils abordaient et, enfin, ces thèmes ont, à leur tour, été rassemblés dans des catégories génériques. Les décisions d'enseignement ont été rangées selon les trois catégories suivantes :

- Managériales (DM) : choix de contrôler la situation d'enseignement, d'assurer des conditions idéales à son bon déroulement et de gérer les interactions maître-élèves,
- Pédagogiques (DP) : énoncés du sujet concernant la stratégie d'enseignement à adopter, l'organisation sociale et physique de la classe, les objectifs pédagogiques, les prérequis et les besoins des élèves, le temps d'instruction, l'évaluation, la sélection des activités d'apprentissage, le contenu, etc.,
- Transitionnelles (DT) : décisions de continuer ou non les activités ou la séquence d'activités, d'y insuffler des changements rythmiques, d'improviser, d'appliquer le protocole, bref, il s'agit de passages, liaisons et enchaînements entre les activités et les séquences d'activités.

Les antécédents aux décisions d'enseignement sont les raisons que les enseignants évoquent pour expliquer leurs décisions planificatives et interactives. Nous les avons regroupés dans le système de catégories suivant :

- Elèves (El) : performance, participation, attention, comportement, motivation, connaissances préalables, aide à leur fournir, etc.
- Enseignant (En) : décision antécédente, routine ou habitude, besoin personnel et état d'esprit, envie, intuition, expérience antérieure, théorie personnelle (conviction), formation initiale (maître de stage), collègues, etc.,

- Activité (Ac) : envie d'atteindre l'objectif, objectif atteint, importance ou inutilité de la tâche, difficulté, suite logique indissociable, évaluation, envie de clore en beauté, etc.,
- Interaction (In) : déroulement et conditions, incertitude, rectification possible et nécessaire, besoin éventuel d'un changement rythmique, etc.,
- Contexte (Co) : travail parallèle isolément ou avec des élèves (correction, remédiation, autre classe), temps alloué et disponible, adaptation par rapport à des leçons précédentes, parents, bulletin, ressources matérielles, organisation de la classe, contexte externe (exigence de l'inspection, de la ville), etc.

Les extraits ont été ventilés selon les décisions d'enseignement qui faisaient référence aux antécédents (de 0 à 5 antécédents selon les décisions considérées) lorsqu'il s'agissait de l'analyse des types de décision ainsi que du nombre moyen d'antécédents et, selon les antécédents lorsqu'il s'agissait de l'étude de leur type. Les analyses que nous proposons ici portent donc sur les décisions (Np=419 ; Ni=458 ; N=877) ou les antécédents (Np=2095 ; Ni=2290 ; N=4385) qui sont caractérisés selon le statut des sujets (normalienne ou enseignante) et le groupe expérimental d'appartenance des sujets qui les ont prononcés. Signalons que les antécédents sont également caractérisés selon le type de décision qu'ils expliquent.

Résultats

La pensée enseignante durant la planification

Types décisionnels selon le statut

Le **tableau 3** montre une même distribution des décisions, que l'on observe les normaliennes ou les enseignantes : beaucoup de décisions pédagogiques, peu de décisions managériales et une fréquence de décisions transitionnelles qui se situe entre ces deux extrémités. La comparaison en fonction du statut indique que les décisions managériales et transitionnelles sont principalement émises par les normaliennes alors que les décisions pédagogiques proviennent surtout des enseignantes.

Tableau 3 – Fréquence des types de décisions d'enseignement selon le statut et les groupes expérimentaux.

			N	DM (%)	DP (%)	DT (%)
P	***N/E***	***N***	165	75 (5,5)	35,5 (79,4)	65,8 (15,2)
		E	254	25 (1,2)	64,5 (93,7)	34,2 (5,1)
	G1/G2	***G1***	249	75 (3,6)	56,9 (84,3)	78,9 (12)
		G2	170	25 (1,8)	43,1 (93,5)	21,1 (4,7)
I	***N/E***	***N***	122	69,6 (26,2)	27 (32,8)	33,6 (41)
		E	221	30,4 (6,3)	73 (48,9)	66,4 (44,8)
	G1-3/G2	***G1-3***	343	95,8 (13,4)	60,4 (43,1)	90,3 (43,4)
		G2	115	4,2 (1,7)	39,6 (84,3)	9,7 (13,9)
	G1/G3	***G1***	207	95,7 (21,3)	67,6 (48,3)	42,3 (30,4)
		G3	136	4,3 (1,5)	32,4 (35,3)	57,7 (63,2)

Antécédents décisionnels selon le statut et les décisions

On constate une certaine similarité entre les patterns des fréquences d'apparition des différents antécédents, que ce soit au niveau des normaliennes ou des enseignantes (**tableau 5**). Malgré cela, les deux populations doivent être considérées comme différentes puisque les enseignantes justifient moins souvent leurs décisions que les normaliennes et, quand elles le font, elles mentionnent essentiellement des antécédents décisionnels qui leur sont propres ou relatifs au contexte. Les normaliennes, de leur côté, expliquent plus souvent leurs décisions par des antécédents relatifs aux élèves, à l'activité et à l'interaction. Les normaliennes de la présente étude mentionnent, en moyenne, davantage d'antécédents décisionnels de planification que leurs aînées (**tableau 4**).

Globalement, les décisions d'enseignement sont peu fréquemment justifiées par les sujets. Dans l'ensemble, les antécédents mentionnés renvoient principalement aux élèves et rarement à l'interaction. Néanmoins, on peut observer deux pics pour les décisions transitionnelles : au niveau de l'antécédent « activité » et au niveau de l'antécédent « contexte ». Peu importe l'antécédent analysé, on observe toujours une distribution identique en faveur des décisions pédagogiques et aux dépens des décisions managériales.

Tableau 4 – Nombre moyen d'antécédents selon le statut et les groupes expérimentaux (ANOVA).

			N	M	e-t
P	*N/E*	*N*	165	1,25	1,21
		E	254	0,86	1,07
	G1/G2	*G1*	249	1,02	1,20
		G2	170	1,01	1,07
I	*N/E*	*N*	122	2,36	1,46
		E	221	1,87	1,47
	G1/2/3	*G2*	115	0,97	1,15
		G1	207	1,97	1,47
		G3	136	2,16	1,49

Types et antécédents décisionnels selon le dispositif de recherche utilisé

On observe un même pattern de fréquence des décisions d'enseignement pour les groupes 1 et 2 : d'abord les décisions pédagogiques, ensuite, les transitionnelles et enfin, les managériales (**tableau 3**). Toutefois, ces deux groupes doivent être différenciés l'un de l'autre car le groupe 1, à tout égard, mentionne plus de décisions.

Mis à part pour l'antécédent « activité » mentionnés par les sujets du groupe 1, on peut dire que les patterns de fréquence des antécédents sont similaires (**tableau 5**). Néanmoins, le groupe 1 se distingue du groupe 2 par le fait qu'il mentionne plus d'antécédents à ses décisions de planification. A noter, toutefois, que c'est également le groupe qui justifie le moins souvent ses décisions.

Le **tableau 4** montre que l'on ne peut pas distinguer le nombre moyen d'antécédents mentionnés selon le plan expérimental utilisé. Les sujets mentionnent le même nombre moyen d'antécédents, qu'ils en aient fait état lors de réflexions, d'entretiens ou de la combinaison de ces deux techniques.

Tableau 5 – Fréquence des types d'antécédents selon le statut, les décisions d'enseignement et les groupes expérimentaux.

			N	Né (%)	El (%)	En (%)	Ac (%)	In (%)	Co (%)
P	***N/E***	***N***	1435	34,1 (65,6)	53,3 (20,7)	37,5 (3,6)	63,2 (5,9)	55,8 (1,7)	40,4 (2,6)
		E	2270	65,9 (80,2)	44,7 (10,6)	62,5 (3,7)	36,8 (2,2)	44,2 (0,8)	59,6 (2,5)
	DEC	***DM***	290	6,4 (61)	14 (25,9)	7,4 (3,4)	10,5 (4,8)	11,6 (1,7)	9,6 (3,1)
		DP	2830	79,8 (77,9)	65,5 (12,4)	77,9 (3,7)	57,9 (2,7)	69,8 (1,1)	64,9 (2,2)
		DT	585	13,8 (65,1)	20,5 (18,8)	14,7 (3,4)	31,6 (7,2)	18,6 (1,4)	25,5 (4,1)
	G1/G2	***G1***	2630	68,2 (71,7)	80,8 (16,5)	61 (3,2)	88 (4,4)	86 (1,4)	78,7 (2,8)
		G2	1075	31,8 (81,6)	19,2 (9,6)	39 (4,9)	12 (1,5)	14 (0,6)	21,3 (1,9)
I	***N/E***	***N***	1435	41,4 (65,6)	52,3 (20,7)	44 (3,6)	40,2 (5,9)	55,8 (1,7)	36,9 (2,6)
		E	1875	58,6 (70,9)	47,7 (14,5)	56 (3,5)	59,8 (6,7)	44,2 (1)	63,1 (3,5)
	DEC	***DM***	285	7,6 (60,7)	13,2 (26,3)	7,8 (3,2)	7,2 (5,3)	11,6 (1,8)	7,8 (2,8)
		DP	2060	68,5 (75,5)	51,1 (14,1)	57,8 (3,3)	35,4 (3,6)	62,8 (1,3)	45,6 (2,3)
		DT	965	23,9 (56,3)	35,7 (21)	34,5 (4,1)	57,4 (12,4)	25,6 (1,1)	46,6 (5)
	G1-3/G2	***G1-3***	2065	86,5 (61,9)	96,2 (20,6)	96,9 (4,5)	97,1 (8)	100 (1,3)	98,7 (3,6)
		G2	225	13,5 (88,4)	3,8 (7,6)	3,1 (1,3)	2,9 (2,2)	0 (0)	1,3 (0,4)
	G1/G3	***G1***	2630	83 (71,7)	76,4 (16,5)	71,6 (3,2)	56 (4,4)	86 (1,4)	71,8 (2,8)
		G3	680	17 (56,8)	23,6 (19,7)	28,4 (4,9)	44 (13,5)	14 (0,9)	28,2 (4,3)

La pensée enseignante durant l'interaction

Types décisionnels selon le statut

Les patterns de distribution des fréquences d'apparition des décisions sont différents selon le statut des sujets : les enseignantes mentionnent beaucoup plus de décisions pédagogiques et transitionnelles que les normaliennes qui, par rapport à ces premières, rapportent davantage de décisions managériales.

Antécédents décisionnels selon le statut et les décisions

A nouveau, on peut observer, grâce au **tableau 5**, que les raisons des décisions d'enseignement sont rarement avancées par les sujets. Les principaux antécédents des décisions interactives, que celles-ci aient été rapportées par les enseignantes ou par les normaliennes, concernent d'abord les élèves et puis l'activité. En outre, les normaliennes et les enseignantes se distinguent en fonction du nombre moyen d'antécédents mentionnés pour justifier leurs décisions interactives puisque ces premières en citent davantage (**tableau 4**).

En règle générale, les patterns de fréquence des antécédents selon les décisions sont similaires, à l'exception de celui des décisions transitionnelles, qui présente deux

pics : l'un au niveau « activité » et l'autre au niveau « contexte ». Les antécédents peuvent être distingués selon les décisions d'enseignement : d'une part, les décisions pédagogiques renvoient surtout aux élèves, à l'interaction et à des éléments liés aux sujets eux-mêmes et, d'autre part, les décisions transitionnelles font principalement référence à l'activité et à son contexte. Les décisions managériales, quant à elles, semblent être moins expliquées, même si les élèves et l'interaction y sont plus souvent évoqués.

Types et antécédents décisionnels selon le dispositif de recherche utilisé

On remarque que les patterns de fréquence des décisions interactives sont différents selon les groupes expérimentaux : d'une part, l'association du groupe 1 et du groupe 3 mentionne souvent autant de décisions pédagogiques que de décisions transitionnelles et, plus rarement, des décisions managériales, alors que le groupe 2 mentionne essentiellement des décisions pédagogiques (**tableau 3**). La fréquence de prises de décisions rapportées par les groupes qui ont été filmés (G1-3) est plus élevée que celle de l'autre groupe. Les patterns des fréquences de décisions du groupe 1 et du groupe 3 sont différents. Le groupe 1 prend davantage de décisions pédagogiques, alors que pour le groupe 3, ce sont les décisions transitionnelles qui dominent. La fréquence des prises de décisions relevées est généralement plus importante pour le groupe 1.

Le **tableau 5** montre que les fréquences des antécédents rapportés par les groupes filmés répondent au même pattern que ceux du groupe qui n'a pas été filmé. Les décisions sont plus fréquemment expliquées dans les groupes 1 et 3. Soulignons que le groupe 2 ne mentionne aucun antécédent relatif à l'interaction. La comparaison du groupe 1 et du groupe 3 aboutit à un même pattern. Cependant, le groupe 1 justifie plus souvent ses décisions. Le **tableau 4** montre aussi que le nombre moyen d'antécédents mentionnés par les sujets diffère selon le plan expérimental utilisé.

Discussion

Rappelons que deux postulats de base ont transcendé la présente étude qualitative. D'une part, des activités de lecture renvoyant à des situations d'enseignement relativement différentes entraînent des patterns décisionnels similaires entre enseignants et, d'autre part, ces patterns décisionnels ne diffèrent guère en fonction du cycle dans lequel l'activité est implantée. Sans vouloir généraliser nos constats à l'ensemble de la population enseignante, quelques tendances intéressantes peuvent néanmoins être dégagées. Ainsi, cette étude autorise des constats au niveau de la comparaison selon le statut, d'un côté, et au niveau du dispositif de recherche, de l'autre.

En ce qui concerne la comparaison selon le statut, nous pouvons remarquer une différence en fonction de la phase d'enseignement. Même si un profil décisionnel similaire peut être dégagé pour la planification, les normaliennes prennent avant tout des décisions managériales et transitionnelles et expliquent surtout leurs choix par des antécédents renvoyant aux élèves, à l'activité et à l'interaction, alors que les enseignantes prennent davantage de décisions pédagogiques et justifient leurs options essentiellement à l'aide d'arguments relatifs à un certain feeling personnel routinisé et

à des éléments propres au contexte. Pour ce qui est de l'interaction, nous ne pouvons pas mettre en exergue un même profil décisionnel puisque les enseignantes prennent plus de décisions pédagogiques et transitionnelles que les normaliennes qui se soucient surtout du management de la classe. Néanmoins, les antécédents mentionnés répondent à un pattern très semblable faisant principalement référence aux élèves et à l'activité. Signalons que nous sommes confrontés à un grand nombre de décisions qui ne sont pas expliquées, quelle que soit la phase d'enseignement considérée.

Le fait que les normaliennes renseignent, lors de la planification, plus d'antécédents décisionnels que les enseignantes peut éventuellement s'expliquer de deux manières. Premièrement, elles ont eu une occasion supplémentaire de s'exprimer puisqu'elles ont d'abord décrit individuellement leur protocole sur une bande audio et qu'ensuite, une rencontre avec le chercheur était prévue. Néanmoins, une étude qualitative de la masse de données rassemblées dans les réflexions montre que la quantité d'informations différentes est peu abondante, voire insignifiante, par rapport au nombre de renseignements fournis par les entretiens. Deuxièmement, les enseignantes mentionnent souvent le fait qu'elles connaissent leurs élèves et rapportent beaucoup de routines et de choix antérieurs comme raisons décisionnelles. Ceci nous amène à conjecturer qu'elles ont une image simplifiée de la multidimensionnalité de la vie dans leur classe, qui allège leur réflexion pédagogique. Au niveau de la phase interactive, deux hypothèses peuvent être avancées. D'une part, la différence dans les types de décisions mentionnés peut éventuellement être expliquée par l'utilisation de routines qui déchargent l'enseignante de tous les aspects managériaux de l'interaction en classe. Dans la même optique, les normaliennes prennent plus de décisions managériales car leur préoccupation principale est d'implanter coûte que coûte l'activité prévue et exigée pour se conformer aux critères d'évaluation de stage. D'autre part, la centration sur l'activité dans des conditions où la participation, l'attention et la compréhension des élèves sont assurées, peut également être interprétée comme un souci d'adéquation aux exigences des stages pédagogiques car, à l'opposé, les enseignantes justifient leurs décisions par des routines et des éléments contextuels notamment liés aux parents d'élèves, à la direction, etc.

Le maître de stage, le contexte institutionnel du lieu de stage, les formateur de l'institut de formation initiale, etc. avaient des exigences liées à l'évaluation des normaliennes ayant participé à notre étude que ces dernières mentionnent souvent, que ce soit lors de la planification ou de l'interaction. Fortement attachées au respect de ces exigences, car indissociables de leur réussite, ces dernières semblent cependant vivre cet attachement comme une sorte d'entrave à leur prise de décision libre et, partant, au développement de leur professionalité. Ces hypothèses nous poussent à interroger sur la potentielle nécessité de redéfinir les stages pédagogiques, leurs critères d'évaluation, … ainsi que les rôles du maître de stage, du formateur d'instituteurs, en vue d'œuvrer en faveur, d'une part, de la prise de décisions d'enseignement plus libre pour les normaliens et, d'autre part, du développement de leur background professionnel.

Pour ce qui est du dispositif de recherche, malgré le fait que les sujets des différents groupes mentionnent des patterns de décisions et d'antécédents relativement semblables, les techniques utilisées pour le groupe 1 ont permis de

relever une masse d'informations plus importante, tant au niveau de la planification que de l'interaction. Ces constats nous rallient à l'avis de Clark et Peterson (2006) pour qui il est hasardeux d'étudier le processus de planification enseignante isolément et indépendamment de l'enseignement interactif et de la réflexion postactive. Alors que pour la phase planificative, la multiplication de méthodes semble quelque peu moins féconde, le recours au film pour la phase interactive permet de rassembler davantage de données que si l'on n'y a pas recours.

Les méthodes que nous avons utilisées ont, pensons-nous modestement, permis de montrer la complexité de l'acte d'enseignement auquel étaient confrontés les sujets de notre étude. Ainsi, qu'il s'agisse de la phase planificative ou interactive, les sujets faisaient face à une kyrielle de situations et d'événements, non seulement complexes, mais aussi, incertains, qui les ont amenés à prendre une légion de décisions différentes où il n'existait pas de solution optimale.

Rappelons, en définitive, que cet article n'ambitionne pas de présenter des recommandations spécifiques sur la manière de dresser un plan de leçon ou de prendre une décision en cours d'enseignement, ni même sur la manière de former les futurs enseignants, mais plutôt de montrer, avec plus ou moins de précision, pourquoi et comment l'activité d'enseignement est, sur le plan cognitif, complexe et exigeante, ainsi que d'apporter des connaissances supplémentaires sur la manière dont les maîtres s'y prennent pour mener leur travail. Ceci dans l'optique de fournir éventuellement aux chercheurs, aux formateurs d'enseignants, que ce soit pour la formation initiale ou continuée, aux décideurs politiques et aux enseignants eux-mêmes, une perspective complémentaire en proposant des constats issus de la recherche sur la pensée enseignante. En effet, les recherches sur le *teacher thinking* sont d'un grand intérêt pédagogique puisqu'elles permettent, d'une part, d'apporter des éclaircissements concernant les types de connaissances que les enseignants utilisent notamment pour offrir aux chercheurs un cadre conceptuel permettant de décider quel genre d'informations, de conseils et de soutiens peuvent être utiles à la classe (Clark & Lampert, 1986). Cette connaissance des savoirs contextualisés est importante pour les professionnels de la formation des enseignants et/ou les enseignants eux-mêmes (s'ils s'en servent pour mieux comprendre leur propre activité professionnelle) et pour tous ceux qui, en formation initiale ou continuée des maîtres, ont pour objectifs d'encourager leur public et de le doter de moyens pour réfléchir, analyser et critiquer de manière constructive leur propre façon d'enseigner. Enfin, les décideurs politiques responsables en matière d'innovations pédagogiques et de modifications curriculaires seront aussi probablement intéressés car, sans la prise en considération de la pensée enseignante, les recommandations diverses ne seront vraisemblablement pas d'une grande utilité dans la pratique (Clark & Lampert, 1986).

Références

Altet, M. (1994). *La formation professionnelle des enseignants.* Paris : P.U.F.

Ben-Perez, M. (1975). The concept of curriculum potential. *Curriculum Theory Network, 5*, 151-159.

Bromme, R. (1982). *How to analyze routines in teachers' thinking processes during lesson planning.* Paper presented at the annual meeting of the American Educational Research Association, New-York.

Bromme, R. (1989). The « collective student » as the cognitive reference point of teachers' thinking about their students in the classroom. In J. Lowyck, & C. M. Clark (Eds.). *Teacher thinking and professional action.* Studia paedagogica, no 9, Leuven university press, 209-222.

Bromme, R, & Juhl, K. (1984). *Students' understanding of tasks in the view of mathematics teachers.* (Occasional Paper No 58). Bielefeld: Institut für Didaktik der Mathematik.

Bromme, R., & Dobslaw, G. (2003). Teachers' instructional quality and their explanation of students' understanding. In M. Kompf, & P. M. Denicolo (Eds). (2003). *Teacher Thinking twenty years on: Revisiting persisting problems and advances in education*. The Netherlands, Lisse: Swets & Zeitlinger B.V., 25-36.

Calderhead, J. (2003). Planning and thinking in junior high school writing lessons: an exploratory study. In M. Kompf, & P. M. Denicolo (Eds). (2003). *Teacher Thinking twenty years on: Revisiting persisting problems and advances in education.* The Netherlands, Lisse: Swets & Zeitlinger B.V., 53-60.

Carnahan, R. S. (1980). *The effects of teacher planning on classroom processes.* (Tech. Rep. No. 541). Madison: Wisconsin R & D Center for Individualized Schooling.

Charlier, E. (1986). *Étude des décisions pédagogiques prises par les enseignants lors de la planification, caractéristiques et facteurs explicatifs des décisions*. Thèse de doctorat non publiée en sciences de l'éducation. Liège : Université de Liège.

Charlier, E. (1989). *Planifier un cours, c'est prendre des décisions*. PED, Nouvelles pratiques de formation, Editions universitaires. Bruxelles : De Boeck.

Clark, C. M., & Elmore, J. L. (1981). *Transforming curriculum in mathematics, science, and writing : A case study of teacher yearly planning* (Research Series No. 99). East Lansing : Michigan State University, Institute for Research on Teaching.

Clark, C. M., & Lampert, M. (1986). Quel savoir sur l'enseignement pourrait être utile aux maîtres ? In M. Crahay & D. Lafontaine (Eds). (1986). *L'art et la science de l'enseignement*. Liège : Éditions Labor.

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. C. Wittrock (ed.) *Handbook of research on teaching*. (third edition). New York: Mac Millan, 255-296.

Clark, C. M., & Yinger, R. J. (1979b). *Three studies of teacher planning.* (Research Series No. 55). East Lansing: Michigan State University, Institute for Research on Teaching.

Colker, L. (1982). *Teachers' interactive thoughts about pupil cognition*. Unpubliched doctoral dissertation, University of Illinois at Urbana-Champaign.

Conners, R. D. (1978). *An analysis of teacher tought processes, beliefs, and principles during instruction*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Creemers, B. P. M., & Westerhof, K. (1982). *Routinization of instructive and management behaviour of teachers* (research paper). Haren, The Netherlands: Educational Research Institute in the North.

Dahllöf, U.S., & Lundgren, U.P. (1970). *Macro- and micro approaches combined for curriculum process analysis : a Swedish educational field project*. Reports from the Institute of Education, University of Göteborg, no 10, April.

Dessus, P. (2002). Les effets de la planification sur l'activité de l'enseignant en classe. In P. Bressoux (Ed.). *Les stratégies de l'enseignant en situation d'interaction.* Note de synthèse pour Cognitique, Programme école et sciences cognitives.

Dunkin, M. J. (1986). Concepts et modèles dans l'analyse des processus d'enseignement. In M. Crahay & D. Lafontaine (Eds). (1986). *L'art et la science de l'enseignement*. Liège : Éditions Labor.

Fogarty, J. L., Wang, M. C., & Creek, R. (1982, March). *A descriptive study of experienced and novice teachers' interactive instructional decision processes*. Paper presented at the annual meeting of the American Educational Research Association, New York City.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of Curriculum-Based Measurement and Consultation on Teacher Planning and Student Achievement in Mathematics operations. *American Educational Research Journal*, *28*(3), 617-641.

Fuchs, L. S., Fuchs, D., Karns, K., Hamlett, C. L., & Katzaroff, M. (1999). Mathematics performance assessment in the classroom: effects on teacher planning and student problem solving. *American Educational Research Journal*, *36*(3), 609-646.

Halkes, R., & Deijkers, R. (2003). Teachers' teaching criteria. In M. Kompf, & P. M. Denicolo (Eds.). (2003). *Teacher Thinking twenty years on: Revisiting persisting problems and advances in education*. The Netherlands, Lisse : Swets & Zeitlinger B.V., 3-13.

Hill, J., Yinger, R. J., & Robbins, D. (1981, April). *Instructional planning in a developmental preschool*. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Huberman, A. M. (1986). Répertoires, recettes et vie de classe : comment les enseignants utilisent l'information. In M. Crahay & D. Lafontaine (Eds.). (1986). *L'art et la science de l'enseignement*. Liège : Éditions Labor.

Lowyck, J. (1980). *A process analysis of teaching* (Report No. 21). Leuven, Belgium: Katholieke Universiteit te Leuven, Departement Pedagogische Wetenschappen.

MacKay, D. A., & Marland, P. W. (1978, March). *Thought processes of teachers.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada.

Marland, P. W. (1977). *A study of teachers' interactive thoughts*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Marx, R. W., & Peterson, P. L. (1981). The nature of teacher decision making. In B. R. Joyce, C. C. Brown, & L. Peck (Eds.). *Flexibility in teaching: An excursion into the nature of teaching and training*. New York: Longman.

May, W. T. (1986). Teaching students how to plan: the dominant model and alternatives. *Journal of teacher education*, November-December, 6-12.

McCutcheon, G. (1980). How do elementary teacher plan? The nature of planning and influences on it. *Elementary school Journal, 81*. 4-23.

McLeod, M. A. (1981). *The identification of intended learning outcomes by early childhood teachers : An exploratory study*. Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

McNair, K. (1978). Capturing inflight decisions. *Educational Research Quarterly, 3*(4), 26-42.

Mintz, S., & Yarger, S. J. (1980). *Conceptual level and teachers' written plans.* Paper presented at the AERA meeting, Boston.

Morine, G. (1976). *A study of teacher planning.* Beginning teacher evaluation study, San Francisco: Far West Laboratory.

Morine-Dershimer, G. (1977). *What's in a plan? Stated and unstated plans for lessons.* Paper presented at the annual meeting of the American Educational Research Association, New-York.

Morine-Dershimer, G. (1979). *Teacher plan and classroom reality: The south Bay study: Part 4.* (Research Series No. 60). East Lansing : Michigan State University, Institute for Research on Teaching.

Morine, G., & Vallance, E. (1975). *Special study B: A study of teacher and pupil perceptions of classroom interaction* (Tech. Rep. No. 75-11-6). San Francisco: Far West Laboratory.

Morine-Dershimer, G., & Vallance, E. (1976). *Teacher planning* (Beginning Teacher Evaluation Study, Special Report C). San Francisco : Far West Laboratory.

Peterson, P. L., & Clark, C. M. (1978). Teachers' reports of their cognitive processes during teaching. *American Educational Research Journal, 15*, 417-432.

Peterson, P. L., Marx, R. W., & Clark, C. M. (1978). Teacher planning, teacher behaviour, and student achievement. *American Educational Research Journal, 15*, 417-432.

Semmel, D. S. (1977). *The effects of training on teacher decision making*. Paper presented at the annual meeting of the American Educational Research Association, New York City.

Shavelson, R. J., Atwood, N. K., & Borko, H. (1977). Experiments on some factors contributing to teachers' pedagogical decisions. *Cambridge Journal of education, 7*, 51-70.

Shroyer, J. C. (1981). *Critical moments in the teaching of mathematics: What makes teaching difficult?* Unpublished doctoral dissertation, Michigan State University, East Lansing.

Smith, E. L., & Sendelbach, N. B. (1979). *Teacher intentions for science instruction and their antecedents in program materials.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Sutcliffe, J., & Whitfield, R. (1979). Classroom-based teaching decisions. In J. Eggleston (Ed.). *Teacher decision-making in the classroom: A collection of papers*. London: Routledge & Kegan Paul.

Taylor, P. H. (1970). *How teachers plan their courses*. Slough, Bershire, England : National Foundation for Educational Research.

Tillema, H. (2003). Categories in Teacher Planning. In M. Kompf, & P. M. Denicolo (Eds.). (2003). *Teacher Thinking twenty years on: Revisiting persisting*

problems and advances in education. The Netherlands, Lisse : Swets & Zeitlinger B.V., 61-69.

Tochon, F. V. (1993). *L'enseignant(e) expert(e)*. Les repères pédagogiques. Québec-Paris : Nathan pédagogie.

Tyler, R. W. (1950). *Basic principles of curriculum and instruction.* Chicago: University of Chicago Press.

Wanlin, P. (2003). *Quels sont les indices qu'une enseignante de première année primaire prend en compte pour cadencer son instruction en lecture ? Une étude de cas*. Mémoire de licence non publié. Liège : Université de Liège.

Wanlin, P. (2004). Présentation des indices qu'une enseignante de première année primaire avance pour expliquer ses prises de decision lors de la phase interactive de son enseignement de la lecture. *Les cahiers du Service de Pédagogie expérimentale, 19-20*, Liège : Université de Liège.

Wodlinger, M. G. (1980). *A study of teacher interactive decision making.* Unpublished doctoral dissertation, University of Alberta, Edmonton, Canada.

Yinger, R. J. (1977). *A study of teacher planning: Description and theory development using ethnographic and information processing methods.* Unpublished doctoral dissertation, Mishigan State University, East Lansing.

Yinger, R. J., & Clark, C. M. (1982). *Understanding teachers' judgments about instruction: The task, the method, and the meaning.* (Research series No 121). East Lansing: Michigan State University, Institute for Research on Teaching.

Yinger, R. J., & Clark, C. M. (1983). *Self-reports of teacher judgments*. (Research series No 134). East Lansing: Michigan State University, Institute for Research on Teaching.

Zahorik, J. A. (1970). The effects of planning on teaching. *Elementary School Journal, 71*, 143-151.

Zahorik, J. A. (1975). Teachers' planning models. *Educational Leadership, 33*, 2, 134-139.

UNE CLASSE COOPERATIVE VERTICALE DANS L'ENSEIGNEMENT SUPERIEUR : EVALUATION DES IMPACTS DU DISPOSITIF AUPRES DE FUTURS ENSEIGNANTS.

Voz Grégory
Université du Luxembourg – Educational Measurement and Applied Cognitive Science
Université de Liège – Service de Pédagogie expérimentale
gregory.voz@uni.lu

Mots-clés : Evaluation, Formation des Maîtres, Pédagogie Institutionnelle

Résumé :

Un établissement de formation d'enseignants met en place une classe coopérative verticale, avec des buts philosophiques engagés et l'intention de développer plusieurs compétences essentielles, recouvertes par les pôles définis par Paquay notamment : chercheur, acteur social et personne en relation.

La formation doit apporter les preuves de son efficience, et donc permettre de certifier ces compétences actuelles. C'est dans cette perspective que le groupe d'enseignants a décidé de mettre en place un groupe chargé d'évaluer les impacts de cette formation sur les étudiants. C'est ainsi qu'en partenariat avec l'ULg, l'équipe a investigué des techniques et méthodes ainsi que des outils et défini les évaluations de ces compétences au fil de l'année.

Elles se font de diverses manières :

- Double interviews semi-dirigées d'un échantillon d'étudiants au cours de l'année sur leur projet personnel afin d'évaluer l'évolution de celui-ci et la congruence avec le programme de formation.
- Observation lors de moments-clefs (où la participation de chacun est la base même du travail) avec ou sans enseignant, à l'aide de grilles construites sur la base de travaux précédents pour le pôle « acteur social », sur la base de confrontations entre membres pour le pôle « personne en relation ». Ces deux pôles n'étant nullement indépendants, une partie des grilles est commune, de même que les observations.
- Interviews dirigées d'étudiants, avec comme support un travail de recherche récent, afin d'évaluer le pôle « chercheur » au sens de l'entretien d'un rapport critique et autonome au savoir.

Cette communication a deux buts principaux :

- Faire profiter de cette initiative mise en place et des outils utilisés pour évaluer des compétences parfois neuves, voire originales.
- Profiter de l'expérience et de la réflexion d'un public de professionnels afin d'améliorer encore le dispositif d'évaluation et de formation de ces futurs enseignants.

Le contexte d'enseignement : un système de formation « extra-ordinaire »

En Belgique, la formation des futurs enseignants pour les élèves de 3 à 15 ans est classiquement de trois années successives au cours desquelles les étudiants suivent des enseignements et stages prévus dans un ordre préétabli allant du début de la première année à la fin de la troisième. Les enseignants sont répartis selon leur discipline et dans des classes définies par année d'études.

Hors de cet ordinaire, un établissement a décidé de donner une formation différente sur plusieurs plans aux futurs régents en sciences humaines, qui seront donc amenés à travailler dans l'enseignement secondaire inférieur.

Nous allons rapidement décrire ce contexte de pédagogie institutionnelle. L'organisation des cours n'y est plus réglée par années mais de manière verticale selon des travaux individuels et en commun au travers des trois années de formation. Les étudiants étant en 3e année sont ainsi en classe avec des étudiants des deux autres années et réalisent ensemble différents projets au cours desquels les enseignements en termes de contenus sont différenciés selon les expériences de l'étudiant. Ces enseignements ont aussi pour objectifs énoncés de développer différentes compétences relationnelles et comportementales. Les enseignants y travaillent en équipe interdisciplinaire au sens le plus prononcé du terme. En effet bien qu'ayant des formations différentes (quatre licenciés respectivement en Sciences de l'éducation, Histoire, Géographie et Sociologie), les enseignants sont des personnes ressources pour toutes les activités, et cela pour chaque élève sans aucune distinction d'année.

Les étudiants sont invités à prendre une part active dans la planification et réalisation du curriculum, mais peuvent aussi critiquer ou contester, de manière formelle ou informelle la formation en général, ou bien un ou plusieurs aspects en particulier.

L'organisation des rôles de chacun est révélatrice de ce système extraordinaire dans lequel les étudiants sont loin d'être réduits à des réceptacles mais sont clairement investis de positions et de responsabilités démocratiquement établies.

L'objet de cette communication n'est pas d'expliquer le dispositif innovant de formation et nous ne nous étendrons donc pas plus longuement sur cette partie. Les personnes intéressées par cette description pourront se manifester via l'adresse Internet trouvée en tête de document.

Evaluer ce système : pourquoi ?

Bien que fondé et réfléchi, ce dispositif d'enseignement demande une évaluation. En effet le groupe d'acteurs à la base de ce travail a planifié dans la mise en œuvre de la formation une évaluation. Cette évaluation doit répondre à trois missions principales par rapport à ce système de formation :

- Le cautionner,
- Connaître son efficacité,
- Le réguler si nécessaire l'année suivante.

Cautionner le système

Le cautionnement peut apparaître comme une raison originale, peu banale du moins, mais pour comprendre cette mission il est nécessaire de se rappeler le contexte européen actuel de formation dans le supérieur.

Depuis les réformes souvent nommées « Bologne », les pratiques qui en découlent sont principalement orientées vers une individualisation des apprentissages afin de permettre à chacun de réussir au meilleur de ses compétences et de son investissement, indépendamment pratiquement d'une institution, ou d'un pays, de formation.

Le système ici rencontré va, et de manière revendicatrice, à l'encontre de cette individualisation compétitive et qui ne tient que peu compte des unités groupe, au profit de l'étudiant seul dans son travail d'apprentissage. Ainsi ce système s'insère par des voies parfois étroites dans la politique et le système d'enseignement préconisé au niveau de la Communauté Française de Belgique. Dans ce contexte, il est nécessaire d'apporter une caution à ce système marginal par rapport aux autres établissements délivrant le même diplôme.

Connaître l'efficacité du système

Comme toute innovation, il est important de connaître les effets à termes des aménagements ou, comme dans notre cas, des modifications quasiment complètes du système. Le groupe souhaite éviter l'écueil d'un raccourci entre l'espoir que la proposition soit belle et le fait qu'elle soit bonne. En pédagogie, les exemples de ces croyances absolues s'avérant fausses à l'épreuve de l'évaluation sont légion. Ainsi au bout de trois ans, l'équipe devra fournir la preuve que son système est au moins aussi efficace que le précédent, plus traditionnel. Sans cela, l'investissement humain considérable et l'étroitesse des chemins empruntés pour accéder aux critères de formation établis ne seront, d'une part, plus de rigueur mais surtout, d'autre part, devront disparaître pour revenir à une formation plus familière. Dans le cas contraire, les résultats pourront donner des perspectives pour d'autres systèmes de formation et tenter de contribuer à l'amélioration de la formation des maîtres en général.

Réguler le système

Si l'évaluation du système doit porter sur trois ans, nous avons pris le parti de pouvoir l'aménager dans son courant, à chaque rentrée, si un quelconque élément amenait le groupe d'enseignants et le comité d'accompagnement à souhaiter modifier le système. Nous pensons à la découverte d'un effet non désiré et indésirable qui ne ferait que s'amplifier sur un ou plusieurs étudiants ou enseignants, ou bien à l'inverse un effet imprévu mais bénéfique qui serait jugé opportun de stimuler. Si d'un point de vue méthodologique le système ne devrait subir aucune modification pour en assurer l'évaluation la plus correcte possible, nous avons plutôt pris le parti de donner toutes les garanties possibles aux étudiants, qui ne peuvent bien sûr pas être réduits à des sujets d'expérience.

Que souhaite-t-on évaluer ?

Quelle référentialisation pour cette évaluation ?

Pour débuter nous avons réfléchi aux points de repères avec lesquels nous allions effectuer l'évaluation. Il a alors été question de faire une évaluation comparative des effets de la formation avec des étudiants d'une formation donnant accès au même diplôme mais dans un établissement fonctionnant de manière plus habituelle, ou bien des étudiants d'une autre formation d'enseignants dans le même établissement. Toutefois pour des raisons méthodologiques (difficultés de comparabilité des éléments humains notamment) et également affectives cette option a rapidement, et prudemment, été abandonnée.

Nous avons choisi d'autres points de repères pour effectuer notre évaluation :

- Les objectifs du système affichés et présentés aux étudiants, dans l'institution et en-dehors
- Les performances d'étudiants de la même section dans le système précédent.

Evaluer l'atteinte des objectifs spécifiques de cette formation

Parmi les objectifs de la formation nous en distinguerons deux : les compétences habituellement évaluées dans la formation de maîtres et les compétences nouvelles[1] jugées actuellement nécessaires.

Pour la première partie, les compétences habituelles sont identifiées dans le programme de formation. On y trouve entre autres :

- Mettre en œuvre des compétences de traitement de l'information
- Exercer des compétences méthodologiques
- Pratiquer des compétences instrumentales
- Elaborer des savoirs
- Utiliser des savoir-faire
- …

Le système évalue ces apprentissages via des méthodes traditionnelles (la grille d'évaluation des stages par le maître de stage notamment) et des examens individualisés, pour lesquels les étudiants établissent avec les enseignants les compétences à certifier, en fonction de celles travaillées dans les différents projets et cours de l'année académique.

En ce qui concerne la seconde partie, elle se rapporte aux compétences, plus transversales, que doit acquérir un enseignant selon les travaux de Paquay repris par l'Administration Générale de l'Enseignement et de la Recherche Scientifique [AGERS] en 2001.

Ainsi on peut lire sur la figure qui suit une représentation de ces aspects attendus de l'enseignant du 21e siècle en Communauté Française de Belgique

[1] Ces compétences, plus que d'être neuves, bénéficient actuellement d'une mise en emphase en Communauté Française de Belgique, comme dans d'autres régions.

Figure 1 – Les axes de formation des enseignants selon les travaux de Léopold Paquay (AGERS, 2001, p.15)

Concernant les volets « Maître instruit » et « Pédagogue », ce sont les évaluations concernant les compétences du référentiel dont nous avons parlé ci-dessus qui ont été jugées suffisantes. De la même manière, nous conservons les maîtres en stages et les évaluations réalisées durant ceux-ci comme les meilleures appréciations concernant le pôle « Praticien ».

Partant, c'est sur les volets « chercheur », « personne en relation » et « acteur social » que nous avons porté notre attention. Au vu des objectifs de la pédagogie institutionnelle, on peut espérer voir se développer les compétences d'un acteur social, les possibilités nombreuses de communiquer et d'entrer en relation, sans oublier bien sûr les travaux communs et individuels de recherche. Ces trois points seront décrits individuellement plus bas. Un quatrième élément fut soumis à l'évaluation : le projet personnel. Il est apparu essentiel au groupe d'évaluation de l'aborder car il est actuellement discuté dans les facteurs de réussite et d'échec notamment dans le supérieur (Service de Pédagogie Universitaire, 2001) et que la pédagogie institutionnelle se base celui-ci pour former des enseignants. Ainsi le projet personnel des étudiants fut investigué afin de connaître son adéquation avec la formation. Dans son prolongement nous avons voulu mettre l'accent sur l'identité d'enseignant que se construisent ces étudiants.

Comparer les productions avec les années précédentes

Cette deuxième partie ne pourra se faire que lorsque des étudiants auront complètement effectué les trois années de formation. Ils seront alors comparés avec des productions d'étudiants formés complètement dans l'ancien système. Les éléments communs sur lesquels porteront les comparaisons sont, actuellement, la qualité des mémoires et les rapports de stage, au sein desquels les maîtres de stage sont invités à se prononcer sur les différentes facettes du futur enseignant.

L'évaluation cette année

En somme, le travail présenté ici se limitera à l'évaluation de cette année 2005-2006, par rapport à trois des composantes de l'enseignant selon la figure 1 ainsi que sur le projet personnel de l'étudiant. Les thèmes seront donc :

- L'enseignant-chercheur,
- L'enseignant-personne en relation,
- L'enseignant-acteur social,
- Le projet personnel de l'étudiant et son identité enseignante.

Comment évaluer ?

Remarques générales

Choix et confection des outils

La construction des outils fut réfléchie et décidée au sein du sous-groupe chargé de l'évaluation, que nous nommerons à la suite « groupe évaluation ». Le cheminement fut relativement long et les critiques de part et d'autres nous ont amené à générer les outils qui seront décrits ci-dessous.

Choix de l'échantillon

Les observations et interviews de chaque étudiant étant impossible à réaliser, nous avons pour deux axes (enseignant-chercheur et projet personnel) stratifié les étudiants selon deux critères : le sexe et la période précédant leur inscription. Ainsi, quatre groupes d'étudiants ont été créés : les filles ou les garçons arrivant directement de l'enseignement secondaire / les filles ou les garçons s'étant d'abord inscrit(e)s au moins une année dans d'autres études supérieures. Ceci ne représente effectivement que trois groupes car en 2005-2006, aucun des garçons de cette option ne provenait directement du secondaire. Les étudiants ont ensuite été tirés au hasard dans ces catégories, puis contactés pour approbation. Nous avons pris soin de ne pas solliciter les étudiants pour deux axes différents, afin de ne pas les étouffer d'évaluation et également d'éviter les contaminations entre ces deux axes.

Tableau 1 - Répartition totale des étudiants de 1re année selon l'âge et la période précédant leur inscription

	Arrivant du secondaire	Ayant déjà été inscrits dans d'autres études
Filles	9	6
Garçons		6

Choix des « évaluateurs »

Pour tous ces dispositifs évaluatifs, il nous a semblé préférable de faire intervenir des personnes externes. Les étudiants ont donc été observés ou interviewés par des personnes n'intervenant en rien dans leur formation. Ce recul était pour nous la garantie de pouvoir s'investir dans ces tâches sans s'écarter des activités d'enseignement et d'apprentissage, respectivement pour les professeurs et pour les étudiants. Pour obtenir ce groupe « d'évaluateurs » nous avons fait un appel auquel quatre étudiantes en seconde licence en Sciences de l'éducation ont répondu dans le cadre d'un stage au sein de leur propre formation. Ce sont donc le plus souvent elles qui ont effectué les rencontres et observations. Ainsi intégrées dans le « groupe évaluation » elles ont été encadrées par celui-ci mais lui ont également apporté documents et critiques afin de permettre la mise en place du dispositif. Elles ont aussi fourni un grand travail d'analyse et de synthèse, base des résultats présentés ci-après.

L'enseignant-chercheur

Objectifs

L'objectif, double, était de vérifier si les étudiants présentent des compétences de chercheur

- au sens de cette définition : *Entretenir un rapport critique et autonome avec le savoir scientifique passé et à venir*.
- en accord avec la démarche de recherche utilisée dans leur formation : *imprégnation – exploration et rupture – cadre théorique et hypothèses – observation contrôlée – analyse – synthèse*[2]

Outils

Pour ce faire nous avons choisi de questionner des étudiants sur des travaux qu'ils avaient effectués et dans lesquels les étudiants sont amenés à exercer leur pôle chercheur. Le travail par excellence aurait été le TFE[3], mais les étudiants qui l'effectuent actuellement n'ont été formés que partiellement dans ce système. Nous

[2] Cette méthode est inspirée de Quivy et van Campenhoudt (1995).

[3] Le Travail de Fin d'études est un travail de recherche original réalisé individuellement, avec un promoteur, durant la dernière année.

avons donc opté pour un autre travail, celui d'AFP[4]. Les étudiants ont été interrogés individuellement et oralement avec un questionnaire réalisé par le groupe évaluation, portant sur :

- Les ressources utilisées
- L'aspect autonome
- L'aspect critique
- Le plan de recherche et la rupture
- La qualité de la synthèse

L'outil devait donc être utilisé comme un soutien à l'introspection. Afin de ne pas se limiter aux souvenirs spontanés des étudiants, la chercheuse avait avant chaque entretien, observé la bibliographie de chaque travail ainsi que les synthèses et prenait note des éléments importants de l'entretien afin de faire le plus de relances possibles.

Public

Tableau 2 - Répartition des étudiants interviewés sur le pôle « enseignant chercheur » selon l'âge et la période précédant leur inscription

	Arrivant du secondaire	Ayant déjà été inscrits dans d'autres études
Filles	3	1 (1 refus[5])
Garçons		2

L'évaluatrice était une étudiante en Sciences de l'éducation, encadrée d'un membre du groupe évaluation n'ayant aucun contact avec les étudiants de ce système de formation.

Déroulement

Les étudiants ont été rencontrés individuellement et ce n'est qu'avec leur accord qu'ils ont participé au questionnement.

La technique fut des interviews individuelles semi structurées sur la base du questionnaire.

Les analyses se sont déroulées de la manière suivante :

- Pour chaque étudiant une analyse selon les cinq aspects relevés dans le questionnement, pointait ceux qui étaient atteints, ceux en développement et ceux étant encore totalement négligés par l'étudiant,
- La chercheuse a ensuite ajouté sa perception lors des rencontres et les éléments choquants qui n'émanaient pas directement du questionnaire initial mais qui lui semblaient importants,

[4] Les Activités de Formation Professionnelle sont des préparations de stage pour lesquelles les étudiants sont amenés à réaliser, individuellement, une leçon qu'ils pourraient gérer en classe sur un thème défini, cette année : « l'évolution d'un quartier belge »

[5] Sept étudiants avaient accepté la rencontre, mais une d'entre eux n'a finalement pas souhaité effectuer l'interview. Les entretiens se déroulant en fin d'année il n'a pas été possible de pallier ce manque.

- Suite à ces six analyses individuelles une ligne de base transversale a été dressée afin de donner une idée plus générale des étudiants de cette année.

Suite

- Les interviews seront reconduites en 2e et 3e sur des travaux semblables (AFP), pour observer l'évolution de ces étudiants,
- Les interviews se porteront aussi en 3e sur le TFE,
- Une grille d'observation des TFE sera construite afin de comparer les travaux réalisés par des étudiants de l'ancien système avec des TFE réalisés après trois ans dans celui-ci.

Le projet personnel

Objectifs

Vérifier que le système…
- Permette la prise en compte du projet personnel des étudiants,
- Fasse émerger une identité professionnelle d'enseignant en Sciences humaines en adéquation avec les principes de ce système de formation.

Outils

Ici encore c'est par interviews semi dirigées individuelles que le travail s'est réalisé. Le canevas de questions, dont l'interrogateur pouvait s'écarter selon la tournure de l'interview portait sur :
- Le choix professionnel (le passé),
- Le vécu actuel (le présent),
- Les perspectives personnelles (l'avenir).

Le questionnaire n'a pas changé pour les secondes interviews, mais l'insistance a toutefois été différente sur certaines questions.

Public

Selon la même distinction, sept étudiants[6] ont été tirés au sort et interrogés deux fois durant l'année, une première fois durant octobre ou novembre, la deuxième durant le mois de mai.

Tableau 3 - Répartition des étudiants interviewés sur le pôle « Projet Personnel » selon l'âge et la période précédant leur inscription (1re interview – 2e interview)

	Arrivant du secondaire	Ayant déjà été inscrits dans d'autres études
Filles	3 – 2 (dont la dernière est incomplète)	2 – 2
Garçons		2 – 1

6 Malheureusement, pour des raisons techniques, nous ne disposons des secondes interviews complètes pour uniquement quatre étudiantes.

Pour ce pôle, c'est un duo d'évaluateurs qui a agit. Un enseignant hors du système de formation a effectué les interviews et une étudiante en sciences de l'éducation a réalisé une part des retranscriptions ainsi que la majeure partie des analyses. Celles-ci ont été revues par l'intervieweur ainsi qu'une tierce personne afin d'y apporter amendement ou acceptation.

Déroulement

Les étudiants ont été rencontrés individuellement et ce n'est qu'avec leur accord qu'ils ont participé au questionnement.

Les premières interviews ont été réalisées au premier quadrimestre et analysées en milieu d'année. Les secondes ont eu lieu au mois de mai. Les interviews ont toutes été entièrement retranscrites et ensuite analysées.

Les analyses se sont déroulées de la manière suivante :

- Une chronologie des dires, sorte de ligne du temps de ces propos pour chaque étudiant, a été élaborée après les 1res interviews
- Les liens qui semblaient pertinents entre évènements, éléments factuels, prises de décision et états affectifs ont été mis en évidence par la chercheuse,
- Les chronologies ont toutes été revues deux fois par la chercheuse, le premier tour servant à déterminer les liens possibles, le second à affiner chaque interview,
- Ces travaux ont été revus par l'intervieweur ainsi qu'un membre du groupe évaluation et ensuite acceptés ou amendés.
- Lorsque les profils ont été dressés, une ligne de base transversale a été établie.

De cette première analyse nous avons souligné les questions pour lesquelles les secondes interviews devraient être plus fouillés.

La deuxième partie de l'analyse s'est développée comme suit :

- Chacune des lignes du temps propres à un étudiant est modifiée et prolongée. Ce second travail était guidé par ces questions :
    - ce qui demeure au fil de l'année ?
    - ce qui a changé et à quoi l'attribue-t-on (formation, éléments personnels, etc.)…
        - selon les dires de l'étudiant ?
        - selon le sens ressenti par la chercheuse, l'intervieweur ou la tierce personne ?
- Enfin, on tire une synthèse transversale afin de donner une idée plus générale concernant les étudiants de cette année.

Suite

- Les interviews seront reconduites auprès des mêmes personnes en 2e et 3e pour examiner l'évolution de ces étudiants dans la construction de leur identité et observer l'adéquation, ou non, du système avec les attentes des étudiants.

Acteur Social et Personne en relation

Ces deux pôles ayant des points communs dans les comportements à observer, nous avons choisis de les présenter conjointement car tout le travail s'y rapportant fut réalisé de concert.

Objectifs

Vérifier que le système…

- Permette le développement des compétences de l'acteur social : « *Une personne qui coopère et conteste là où elle est.* »,
- Permette le développement des compétences de la personne en relation : « *Quelqu'un qui travaille en équipe, entre en relation avec tous, concerte les autres et mesure les enjeux éthiques.* »,
- Forme effectivement des acteurs sociaux et des personnes qui pourront facilement entrer en relation.

Outils

Pour ces deux pôles, le groupe a opté pour des observations d'étudiants durant diverses activités influencées peu ou prou par la pédagogie institutionnelle. Pour ce faire deux grilles ont été utilisées. Celle concernant l'acteur social fut réalisée en 1999 par des étudiants de l'ancien système[7]. L'autre fut inspirée de cette grille contenant une partie commune et aménagée pour les autres comportements par le groupe évaluation.

Les cours transmissifs traditionnels et non modifiés par la pédagogie institutionnelle furent écartés des observations. Celles-ci se sont déroulées lors des Projets Collectifs Horizontaux, Projets Collectifs Verticaux, Minipex, présentations, ateliers-classe, quoi d'neuf et assemblée générale, ce qui représentait 28h réparties sur quatre mois.

Public

Six étudiants de première année furent choisis selon un critère : l'aisance de discours et de prise de parole versus la discrétion, le peu de manifestation.

Un groupe de quatre étudiants de seconde année furent observés. Il fut choisi par opportunité parce qu'ils travaillaient ensemble dans un PCH et que cela rendait les observations plus aisées.

Les évaluatrices étaient deux étudiantes de licence en Sciences de l'éducation.

Déroulement

Les étudiants observés n'ont pas eu connaissance qu'ils l'étaient précisément. Les observatrices étaient présentes dans la classe suite à l'acceptation de cette demande lors du conseil, mais n'ont pas prévenu les observés de leurs choix afin de biaiser le moins possible les comportements.

- Les étudiantes se sont appropriées au mieux les grilles avant de se rendre en observation,

[7] TenterPlus, 1999. *Recherche Sociologique sur la participation étudiante à l'Isell Ste-Croix*, Rapport Final (non diffusé).

- Les observations ont toutes été retranscrites,
- Elles ont ensuite été codées « à froid » selon les indicateurs des grilles,
- Les codages ont été vérifiés respectivement par les deux étudiantes et ensuite par une tierce personne non observatrice afin d'apporter des amendements ou acceptations,
- Une description qualitative et quantitative des comportements selon les étudiants ainsi que selon les moments de formation observés a été réalisée,
- Ces comportements ont été synthétisés sous forme d'un tableau général par pôle suivi de tableaux individuels propres à chaque étudiant observé,
- Une synthèse transversale pour chaque pôle a été dressée afin de donner une idée plus générale des étudiants de ces deux années.

Suite

- Les observations seront reconduites auprès des mêmes étudiants en 2e et/ou 3e selon l'année dans laquelle ils rentrent pour observer l'évolution des comportements des étudiants concernant ces deux pôles.

Déroulement de l'évaluation durant 2005-2006

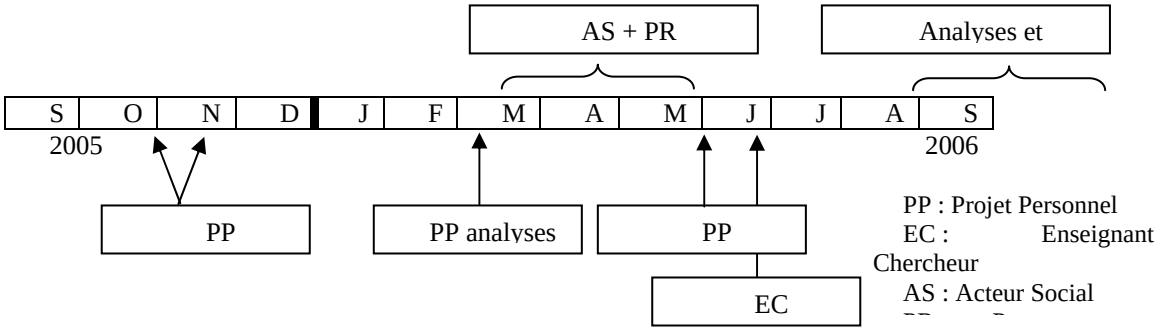


Figure 2 – Chronographe de l'évaluation de septembre 2005 à septembre 2006

Planification de l'évaluation pour 2006-2007

Septembre 2006	Modification du programme de formation
	Appel à « évaluateurs »
Octobre 2006	Formation des évaluateurs
	Premières interviews PP
Novembre 2006 – avril 2007	Observations AS + PR
Mai 2007	Interviews EC
	Secondes interviews PP
	Observations AS+PR
Juin – août 2007	Analyses et synthèses

Quelle suite pour ce travail ?

L'avis des évaluateurs

Suite au travail accompli cette année une première évaluation du travail a été effectuée, donnant place à des perspectives concernant sa poursuite.

Constats		Perspectives
		Il faut pouvoir alléger le dispositif… Comment ?
Le travail demande un investissement de départ important afin de maîtriser en profondeur les outils et ensuite pouvoir les utiliser de manière efficace, en particulier pour l'observation	➔	Une désignation et formation de nouveaux observateurs seront nécessaires pour continuer l'évaluation durant les années 2006-2007 et 2007-2008.
Les outils élaborés longuement durant le début de l'année et au cours de leurs utilisations sont maintenant définitifs	➔	Un temps précieux est gagné et les travaux pourront reprendre plus rapidement durant 2006-2007
Les outils permettent *a posteriori* une structuration importante au départ d'un flou imposant	➔	La satisfaction de découverte renforce positivement les efforts consentis tout au long de l'année
La démarche de recherche mise en place dans le groupe d'enseignants-chercheurs organisant la formation est adoptée dans ce processus	➔	Le travail est prolongé dans cette visée d'isomorphisme
Les étudiants interrogés afin de participer ont tous répondu positivement à une exception près	➔	La facilité de contact témoigne d'un intérêt et est un adjuvant pour le travail de l'année suivante.
Les étudiantes de l'Université ont mis à profit cette expérience pour développer et exercer leurs compétences de recherche, avec satisfaction	➔	Cette expérience renforce les liens entre ces institutions de formations
L'expérience séduit différents acteurs de formation de maître en Belgique et en-dehors	➔	Le réseau de travailleurs s'intensifie

La place de cette communication

Comme le présentait Marthe Hurteau, lors de ce 19e colloque de l'ADMEE, la fonction première de l'évaluation de programme est de poser un jugement de valeur. Les démarches et outils méthodologiques sont des éléments soutenant cet objectif. C'est dans cette perspective que nous avons développé ou adaptés les outils. Notre travail de récolte n'en est encore qu'à ses débuts mais l'approche complexe nous permet de croire en des résultats solides. Dans deux ans, nous proposerons un support reprenant en détails la méthodologie et les outils, mais aussi les résultats complets et détaillés. De ces résultats nous tirerons, d'une part, les jugements attendus et, d'autre part, des perspectives pour la formation des maîtres.

Cette présentation se voulait une contribution à la recherche en formation initiale des enseignants ainsi qu'en évaluation de programmes. Convaincus que le partage

d'expériences est nécessaire afin de résoudre des problèmes internationalement partagés concernant ces deux problématiques, nous demeurons intéressés par les avis et autres manifestations quant à cette communication. Pour ce faire, n'hésitez pas à prendre contact avec nous via l'adresse Internet en tête de document.

Références

Administration Générale de l'Enseignement et de la Recherche Scientifique [AGERS], (2001). *Devenir enseignant. Le métier change, la formation aussi.* Bruxelles : CFWB.

Hurteau, M. (2006). *L'évaluation de programme: notions, modèles et méthodologies.* Présentation au 19e colloque de l'ADMEE Europe, Luxembourg.

Quivy, R. & van Campenhoudt, L. (1995). *Manuel de recherche en sciences sociales.* Paris:Dunod

Service de Pédagogie Universitaire. (2001). Le projet personnel de l'étudiant : un facteur de réussite ? *RESEAU n°49.* Namur : FUNDP.

TenterPlus (1999). *Recherche Sociologique sur la participation étudiante à l'Isell Ste-Croix*, Rapport Final (non diffusé).

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

ET SI ON PARLAIT DE LA PERTINENCE DE L'EVALUATION ?

Carette Vincent, Université Libre de Bruxelles

Mots-clés : pertinence de l'évaluation, compétence, effet-maître.

Résumé:
Cette communication décrit les résultats d'une recherche menée en Communauté française de Belgique sur l'évaluation des compétences. Sur la base d'épreuves d'évaluations de compétences (modèle en 3 phases), nous avons été amenés à mettre en évidence des classes plus performantes que d'autres à des épreuves de compétences. L'analyse, à partir d'un questionnaire et d'entretiens, a permis de dégager des caractéristiques des pratiques de classe d'enseignants de classe susceptibles d'expliquer les performances de leurs élèves.

Ces caractéristiques diffèrent considérablement de celles que l'on trouve dans la littérature lorsque l'on analyse les pratiques des classes en relation avec des performances d'élèves obtenues avec des épreuves plus classiques inspirées par les principes de la pédagogie par objectifs.

Sachant le poids donné actuellement au poids accordé aux épreuves d'évaluation nationale ou internationale, cette communication interroge la notion de pertinence en évaluation qui est d'après nous peu débattue. Si lorsque l'on modifie la forme de l'évaluation, les constats que l'on déduit à partir des résultats obtenus par les élèves diffèrent, il nous apparaît que la question de la pertinence est un enjeu central de l'évaluation de demain.

Problématique

En quelques décennies, la recherche de l'efficacité, de l'équité et de la qualité est devenue un objet incontournable des politiques d'éducation et de formation au niveau international. Si les pays anglo-saxons ont été les premiers à mettre en place des instruments de mesure de l'efficacité, ils ont été suivis par les grandes organisations internationales qui ont ensuite incité d'autres Etats à rejoindre ce mouvement. *« Aujourd'hui, la fixation d'objectifs à atteindre, la mise en place de standards, le développement des tests sont présentés comme le préalable nécessaire au pilotage des systèmes et des organisations, les acteurs de l'éducation et de la formation devant rendre compte de leur capacité à accompagner l'entrée dans la société de la connaissance. »* (Derouet, 2005).

L'actualité pédagogique nous invite régulièrement à des débats sur la qualité de l'enseignement, sur l'efficacité des pratiques enseignantes à travers l'analyse, entre autres, de résultats à des épreuves internationales. Celles-ci organisées sous l'égide d'organisations internationales (OCDE, Banque mondiale, Unesco, etc.) permettent la diffusion de résultats obtenus par des élèves du même âge et de pays différents à des épreuves de lecture, de mathématiques, de sciences, etc. Ces résultats entraînent des

comparaisons entre les différents pays desquelles on déduit un ensemble d'indicateurs qui présumeraient de l'efficacité ou non des systèmes scolaires.

Le principe méthodologique d'évaluer les acquis des élèves et de comparer les résultats obtenus avec d'autres variables ne se limitent pas à l'analyse des systèmes éducatifs dans leur globalité mais est également un principe régulièrement utilisé pour évaluer l'efficacité d'un établissement ou l'efficacité des pratiques enseignantes. Même si ce principe n'est pas la seule approche que l'on rencontre dans la littérature sur ce que l'on appelle dans le jargon pédagogique « l'effet-établissement » ou « l'effet-maître », il apparaît néanmoins que les recherches sur l'efficacité des pratiques enseignantes ou de l'établissement en fonction des résultats obtenus par les élèves à des épreuves d'évaluation ont une réelle audience dans les sphères politiques et éducatives.

Basée sur des méthodologies rigoureuses qui font appel à des traitements statistiques sophistiqués, ces recherches inscrites dans ce que l'on appelle le paradigme « processus-produits » présentent un ensemble de résultats sur les caractéristiques dont devrait faire preuve un enseignant « efficace ».

En synthèse, de manière schématique et en se référant, entre autres, aux travaux de Bressoux (1994 et 2002), l'enseignant efficace dans une approche « processus-produits » serait un enseignant « dirigiste » qui propose des activités « très » structurées. Il travaille à petit pas. Il arrive à gérer son programme avec dextérité en insistant suffisamment longtemps sur les matières importantes. Il propose des évaluations qui correspondent à ce qui a été effectivement enseigné. Il pose de nombreuses questions et pratique un feedback positif (des louanges rares, mais distribuées à bon escient). Il est expert, s'intéresse aux apprentissages fondamentaux et se démarque de l'enseignant « animateur ». L'enseignant efficace peut travailler de manière « frontale » et négliger la différenciation des apprentissages et les travaux de groupes. Il arrive à impliquer les élèves dans les exercices individuels. C'est une personne rigoureuse.

Tableau 1 - L'enseignant efficace issu des recherches « processus-produits »

gestion de classe et des apprentissages	Enseignant dirigiste opposé à un enseignant « animateur ». Enseignement systématique et découpé (importance de la notion d'objectifs). Travaux de groupes peu pertinents Enseignement centré sur les apprentissages fondamentaux. L'enseignant peut négliger la différenciation. Evaluation régulière sur ce qui a été effectivement appris. Importance de la gestion du temps. Importance des louanges adressées aux élèves (à bon escient) par l'enseignant.
Rôle de l'enseignant dans l'établissement	Importance de la cohérence des enseignants entre-eux. Intérêt de la mise en place de projets. Importance d'une volonté d'amener les élèves vers les meilleures performances.
Valeurs	Une certaine productivité
Réflexivité	Intérêt « soupçonné » de la réflexivité de l'enseignant.

Cette synthèse a de quoi surprendre bon nombre de pédagogues habitués à d'autres discours basés sur des propositions pédagogiques qui s'éloignent fortement de celles présentées par ces chercheurs. Ainsi, l'enseignant « facilitateur d'apprentissage », qui fait travailler ses élèves en groupes, qui évite un enseignement « frontal », qui propose des projets à ses élèves, etc., ne serait pas, au regard de ces résultats de recherche, efficace.

Tableau 2 - L'enseignant efficace dans les discours pédagogiques contemporains

	L'enseignant efficace dans les discours pédagogiques contemporains.
Gestion de classe et des apprentissages	Enseignant qui gère et anime des séquences d'apprentissages. Séquences d'apprentissages basées sur la résolution de problèmes. Pertinence des travaux de groupes. Pertinence et importance des pratiques de différenciation. Evaluation formative. Importance de la mise en place de conseil de classe.
Rôle de l'enseignant dans l'établissement	Importance du travail en équipe. Importance de la mise en place de projets. L'entretien de bonnes relations avec les partenaires éducatifs.
Valeurs	Justice, solidarité, non violence.
Réflexivité	Importance de la réflexivité de l'enseignant.

Même si ces propos demandent à être nuancés, la distorsion entre les principes pédagogiques habituellement défendus dans la formation initiale ou continue des enseignants basés, du moins on l'espère, sur la recherche d'une réelle efficacité et ceux énoncés au terme des différentes recherches inscrites dans le paradigme « processus-produits » nous interpellent. Les pédagogues formateurs d'enseignants ont-ils tendance à défendre le contraire de ce qu'il faudrait présenter ? Baignent-ils dans un climat idéologique qui les amènerait, en dépit de résultats scientifiques probants, à enseigner des principes pédagogiques inefficaces ? Ou bien, ces recherches inscrites dans une méthodologie basée sur l'étude d'une efficacité mesurée à partir d'épreuves d'évaluation, présentent-elles des insuffisances, des manquements qui les conduiraient à des conclusions abusives ? Ces recherches ne seraient-elles pas idéologiquement orientées vers une vision limitée de l'efficacité ?

Les réponses à ces questions représentent, d'après nous, un réel enjeu pour l'école de demain. La place de plus en plus importante donnée par les responsables politiques, encouragée d'ailleurs par les médias, aux enquêtes internationales doit nous faire réfléchir aux conséquences des décisions politiques et pédagogiques qui pourraient être prises sur la base de leurs résultats. La proximité entre les méthodologies utilisées dans les enquêtes internationales et celles utilisées dans les recherches « processus-produits » sur « l'effet-établissement » ou « l'effet-maître» pourrait rapidement conduire à des conclusions « scientifiques » sur les caractéristiques « universelles » d'un système éducatif efficace ou du moins efficient, voire d'un établissement ou d'un enseignant efficaces.

S'interroger sur l'efficacité de l'enseignant est pour le pédagogue une question centrale parce que dans la problématique de l'école, la pratique enseignante est une des rares variables sur laquelle on peut agir pour améliorer l'apprentissage des élèves. De fait, l'enjeu de la description des pratiques enseignantes efficaces est loin d'être anodin. Les discours actuels diffusés dans la presse le confirment. Entre les volontés d'une école pour tous ouverte sur le monde et les nostalgiques d'une école du « vrai savoir », le débat fait rage. A partir des enquêtes internationales comme PISA, le débat prend une ampleur politique où on n'hésite plus à affirmer que l'enseignant doit maintenant se concentrer sur ce que l'on appelle en Belgique « les apprentissages de base », laissant supposer dans la foulée qu'il les a abandonné. Les pédagogues sont souvent considérés comme les responsables de réformes qui détruisent les fondements de l'école. Tout se met en place progressivement pour un retour au « vrai enseignement » qui trouverait sa légitimité dans l'analyse des résultats de recherches scientifiques internationales. L'entretien de Luc Ferry, philosophe et ancien ministre français de l'éducation, dans un quotidien francophone (La Libre Belgique) daté du 1 septembre 2005 (jour de la rentrée des classes dans l'enseignement primaire en Belgique) est édifiant. Luc Ferry, dans cet article intitulé *« Retourner aux fondements de l'école »* fait le procès des pédagogies actives qui seraient en partie responsable de ce qu'il appelle la « débâcle » de l'école. *« …La réalité est que depuis les années 60 et la rénovation pédagogique, on a surdéveloppé les méthodes actives en partant de l'idéologie de l'auto-construction. L'on fit croire que les enfants n'apprenaient bien que ce qu'ils construisaient eux-mêmes. Il y a eu donc une multiplication à l'infini de méthodes suscitant la spontanéité des élèves. En les amenant à mettre la main à la pâte, en introduisant du ludique partout et en excluant tout par cœur. »*

Dans ce contexte, la question de la définition des caractéristiques de l'enseignant efficace représente un enjeu dont nous ne mesurons peut-être pas toujours l'importance. Elle nous conduit à défendre l'idée qu'il est intéressant de s'interroger sur la distorsion entre les discours pédagogiques habituellement diffusés dans les formations continues et initiales et les résultats sur les caractéristiques d'un enseignant efficace issus des recherches « processus-produits ».

Une manière d'aborder ce problème est de questionner les méthodologies utilisées qui conduisent à se prononcer sur l'efficacité ou non des pratiques enseignantes.

Si la discussion peut amener à s'interroger sur le bien-fondé des méthodologies qui lient l'efficacité des pratiques enseignantes aux performances des élèves, nous nous sommes plutôt orientés dans cette recherche vers un débat sur le choix des variables dépendantes et indépendantes habituellement utilisées dans les recherches « processus-produits ». En d'autres termes, le caractère « surprenant » des caractéristiques de l'enseignant efficace présenté dans les recherches « processus-produits » pourrait trouver son origine dans la conception des épreuves d'évaluation utilisées pour déterminer le niveau des acquis des élèves et également dans le choix des variables susceptibles d'expliquer les résultats.

Cette hypothèse trouve sa source dans des travaux que nous menons en Communauté française de Belgique depuis plusieurs années sur la notion de compétence, et de manière plus précise sur l'évaluation des compétences.

A l'instar de systèmes éducatifs comme le Québec ou le canton de Genève, la Communauté française de Belgique s'est engagée dans une réforme de ses référentiels. Rédigés en termes de compétences, les nouveaux référentiels définissent ce que les élèves doivent maîtriser à différents moments du cursus scolaire

En publiant ces documents définissant les exigences attendues en termes de compétences, les responsables politiques indiquent, si l'on se place dans une logique « processus-produits », que l'efficacité de l'enseignement devra être mesurée, entre autres, selon des compétences acquises par les élèves. Par suite, si nous poursuivons notre raisonnement, des pratiques enseignantes seront considérées comme efficaces si elles permettent d'amener les élèves à devenir compétents.

Et si on évaluait des compétences, les caractéristiques des enseignants efficaces seraient-elles les mêmes que celles défendues par des auteurs inscrits dans les recherches évaluatives « classiques » ? Telle est la question centrale de notre recherche.

L'introduction de la notion de compétence dans des systèmes scolaires tel que celui de la Communauté française de Belgique présentait une réelle opportunité d'essayer de répondre à cette question. En effet, en introduisant cette notion par l'intermédiaire de référentiels rédigés en termes de compétences, les responsables politiques de la Communauté française se sont engagés dans une redéfinition de ce que l'on attend des élèves. Cette redéfinition de ce que l'on pourrait appeler « l'attendu » interroge l'apprentissage, mais surtout l'évaluation. Par suite, si le niveau des études se définit maintenant en termes de compétences, il a lieu de s'interroger sur la manière dont on va pouvoir le contrôler.

La notion de compétence induirait une nouvelle manière d'évaluer les apprentissages. Définie comme la capacité qu'a un élève à pouvoir utiliser à bon escient ce qu'il a effectivement appris, cette notion met l'accent sur ce que l'on appelle « la mobilisation des acquis ».

De fait, la notion de compétence introduirait un nouvel « attendu » qui se traduirait concrètement par le fait que l'on doit maintenant évaluer les élèves sur leur capacité à mobiliser leurs acquis lors de la résolution de tâches nouvelles plutôt que sur des « *exercices mécaniques* » ou « *de la simple restitution de matières* » (*Socles de Compétences*,1999).

Nos hypothèses de travail sont que si l'on évalue les élèves en référence à ce nouvel « attendu », les caractéristiques des enseignants efficaces seraient différentes de celles décrites par les auteurs des recherches inscrites dans le paradigme « processus-produits » et se rapprocheraient de celles défendues dans les discours pédagogiques contemporains diffusés dans la formation initiale et continue des enseignants.

En d'autres termes, la question posée dans cette recherche est la question de la pertinence de l'évaluation définie par De Ketele (1989) « *comme le caractère plus ou moins approprié de l'épreuve, selon qu'elle s'inscrit dans la ligne des objectifs visés* ».

Méthodologie et principaux résultats

Ces hypothèses reposent sur un ensemble de constats réalisés lors de recherchee menéee par des équipes du Service des Sciences de l'Education (Rey, Carette, Defrance, Kahn, 2001 ; Rey, Carette, Defrance et Van Lint, 2004 ; Carette, 2006) de l'Université Libre de Bruxelles entre 1999 et 2001.

Dans le cadre de nos travaux, nous avons été amenés à construire un modèle d'évaluation de compétences. La plupart des auteurs (Roegiers, 2002 ; Beckers, 2002 ; Scallon, 2004) qui s'intéressent à l'évaluation des compétences s'accordent, avec des nuances, sur les conditions nécessaires d'une telle forme d'évaluation.

La première condition est de proposer aux élèves des situations complexes. La seconde condition est que ces situations complexes doivent être inédites. La troisième condition est que les situations complexes qui doivent être résolues par les élèves nécessitent la mobilisation de procédures normalement apprises. A ces trois conditions, nous en ajoutons une quatrième : toute évaluation de compétence doit avoir un caractère diagnostique pour permettre de gérer les paradoxes engendrés par l'introduction de cette notion. En effet, la notion de compétence en insistant sur la mobilisation des acquis modifie ce que l'on attend des élèves et augmente le niveau d'exigence. C'est le premier paradoxe. Le second paradoxe est que dans une approche par compétence, on demande aux enseignants d'amener les élèves à pouvoir mobiliser leurs acquis alors que notre connaissance sur cette activité cognitive reste bien mystérieuse.

De fait, pour ne pas ériger « la complexité inédite en norme » (Crahay,2006) et pour permettre aux enseignants d'analyser la mobilisation des élèves, nous avons construit un modèle d'évaluation de compétences composées de trois phases distinctes.

- La première phase confronte les élèves à une tâche complexe,
- La seconde phase reprend la tâche proposée dans la première phase, mais celle-ci est décomposée en items. De fait, dans cette seconde phase, on impose la démarche aux élèves,
- Enfin, la troisième phase propose des items décontextualisés qui mesures l'acquisition par les élèves de procédures nécessaires à la réalisation de la tâche proposée.

Six épreuves construites selon ce modèle ont été administrées auprès de 1257 élèves suivant le dispositif expérimental suivant :

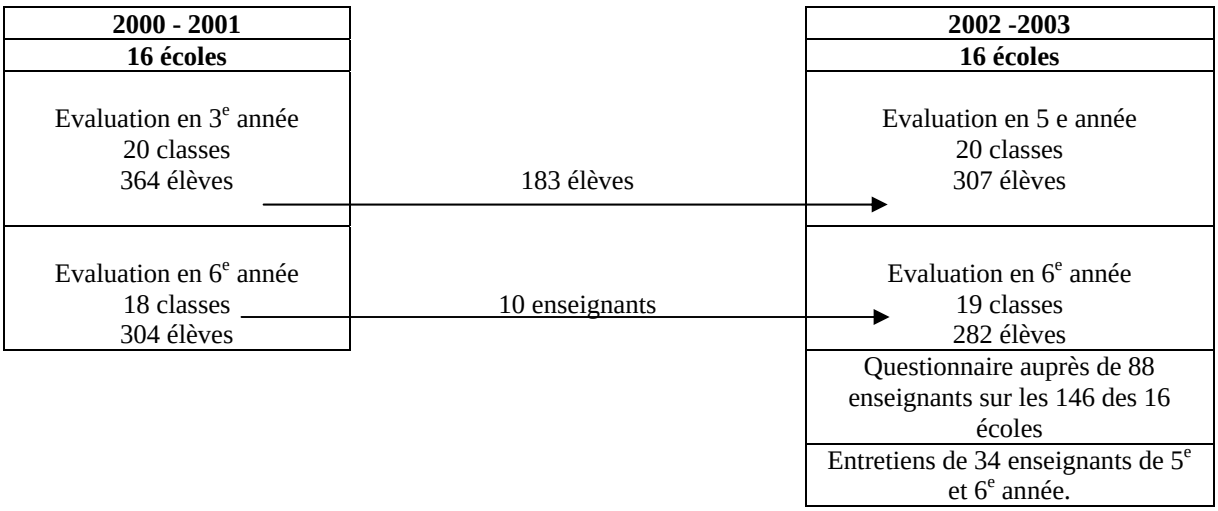


Figure 1 – dispositif expérimental

Lors de l'année scolaire 2000-2001, nous nous sommes rendus dans 16 écoles différentes de Wallonie. Nous avons évalué 364 élèves de 3e primaire et 304 élèves de 6e primaire avec des épreuves construites selon le principe du modèle en trois phases. Deux années plus tard, nous sommes retournés dans les mêmes écoles. Cette fois, nous avons évalué 307 élèves de 5e année primaire et 282 élèves de 6e année primaire. 183 élèves de 5e année avaient déjà été évalués deux années auparavant lors de la première expérimentation. Nous sommes retournés dans les classes de 10 enseignants de 6e primaire qui avaient participé à la première expérimentation.

Pour tenter de déterminer les caractéristiques des pratiques des enseignants qui ont participé à notre expérimentation, nous avons construit un questionnaire que nous avons administré à tous les enseignants des 16 écoles primaires concernées par notre expérimentation. Nous avons également réalisé des entretiens semi-dirigés des enseignants de 5e et de 6e primaire dont nous interrogions les élèves.

Ce dispositif de recherche nous a permis de réaliser plusieurs constats importants :
- De nombreux élèves sont en difficulté lorsqu'on leur présente des tâches complexes,
- Il existe bien une hiérarchie entre les trois phases. Les élèves réussissent moins bien les tâches complexes (Phase 1) et mieux les exercices qui évaluent les procédures nécessaires à la résolution des tâches complexes,
- L'analyse des liens entre les trois phases indique que la maîtrise des procédures (Phase 3) est une condition nécessaire à la résolution des tâches complexes, mais pas suffisantes. En effet, de nombreux élèves qui maîtrisent les procédures sont incapables de les utiliser dans de nouvelles circonstances,

- On observe des différences importantes entre des classes de niveau social équivalent.

Sur la base de ce dernier constat, nous avons essayé de déterminer les caractéristiques des enseignants de notre échantillon à partir d'un questionnaire adressé à l'ensemble des instituteurs des écoles de notre échantillon et d'entretiens administrés auprès de 34 instituteurs dont nous avions évalué les élèves.

L'analyse des questionnaires traités suivant le principe statistique des « Nuées Dynamiques » et des entretiens nous ont conduit à dresser le portrait de l'enseignant des classes performantes à des épreuves de compétences.

Les enseignants « efficaces » évalués à partir des résultats à des épreuves d'évaluation de compétences sont des enseignants qui défendent fortement le principe d'éducabilité. De fait, ils défendent l'idée que tous les élèves peuvent apprendre malgré des difficultés liées entre autres à l'influence du milieu social. Ils défendent plutôt l'idée qu' »il n'y a pas des élèves intelligents et d'autres qui ne le sont pas » et estiment plutôt que l'école contribue à la progression de l'intelligence des élèves.

Ils défendent plutôt une école orientée vers une préparation à la vie et ont pour la plupart d'entre eux une vision assez optimiste des élèves d'aujourd'hui. Ils sont partagés sur la question de la pratique du redoublement et sur l'intérêt des réformes mises en place en Communauté française de Belgique.

Par rapport à la gestion de la classe, ils défendent l'idée que l'apprentissage doit se réaliser dans l'action. Ils défendent plutôt un enseignement « contraignant », c'est-à-dire un enseignement où leur rôle essentiel est d'inciter les élèves à apprendre. Ils travaillent régulièrement avec leurs collègues dans le cadre, entre autres, du travail en cycles.

Ces enseignants proposent régulièrement des problèmes aux élèves que ce soit sous la forme de défis ou de projets qui donnent du sens et de la cohérence aux apprentissages. Ces projets doivent, toutefois, être orientés vers l'apprentissage plutôt que vers la production. Ils ne négligent pas l'automatisation des procédures, car ils défendent l'idée que les élèves sans cet outillage risquent d'être complètement démunis face aux problèmes. Ils insistent sur le principe que les élèves doivent régulièrement confronter ce qu'ils pensent aux conceptions des autres élèves et avec celles de l'enseignant Pour favoriser cette confrontation, un climat de classe où les élèves se sentent bien est nécessaire. Ce climat doit permettre d'établir une bonne communication.

Au-delà de ces pratiques, les enseignants efficaces présentent auprès de leurs classes « des intentions cognitives », et plus précisément pour la majorité d'entre eux des « intentions de mobilisation » qui favorisent l'adoption « d'un cadrage instruit » par leurs élèves.

Observer et écouter les élèves sont des qualités importantes. Ces enseignants se remettent souvent en question et abordent les questions portant sur leurs pratiques en les « problématisant ».

Toutefois, ils ont des avis partagés sur l'intérêt des concertations obligatoires imposées dans le travail en cycles. Enfin, environ la moitié des enseignants concentrent leurs activités autour du français et des mathématiques et l'autre pas.

Par rapport à l'autorité, l'enseignant efficace a plutôt tendance à penser que son exercice n'est pas pervers et ne maintient pas l'élève dans la dépendance. Il défend plutôt l'idée que l'exercice de l'autorité permet de structurer la personnalité des élèves.

Et si on parlait de la pertinence de l'évaluation

L'analyse des discours des enseignants des classes performantes à partir d'un questionnaire et d'entretiens semi-dirigés nous a conduit à dégager les principales caractéristiques de leurs pratiques. Celles-ci se démarquent clairement de celles énoncées habituellement dans les recherches qui lient l'efficacité des pratiques aux performances des élèves mesurées à partir d'épreuves d'évaluation plus « classiques », c'est-à-dire des évaluations qui s'inscrivent plutôt dans une pédagogie par objectifs. Elles se rapprochent des caractéristiques énoncées dans les discours pédagogiques contemporains.

Ce constat interroge directement les principes sur lesquels se fondent la plupart des épreuves d'évaluation (nationales ou internationales). Notre recherche nous a amené à mettre en évidence qu'une modification des formes d'évaluation conduisait à des constats diamétralement opposés sur la manière d'envisager l'efficacité de certaines pratiques de classe. Elle a montré que lorsque l'on focalise l'évaluation sur le « savoir mobiliser » des élèves tel que le demande l'approche par compétences, les caractéristiques des enseignants efficaces sont proches de celles énoncées habituellement dans les discours pédagogiques.

Dans le tableau ci-après, nous avons essayé de comparer différentes caractéristiques que devraient présenter une épreuve de compétences avec celles que l'on peut observer dans des épreuves d'évaluation construites selon les principes de la pédagogie par objectifs.

Tableau 3 – comparaison de l'évaluation

	Approche par Compétences	**Approche par Objectifs**
Optique de l'évaluation	**Compétences** : évaluer consiste à proposer une ou des situations complexes, qui nécessitera (ont) de la part de l'élève une production elle-même complexe pour résoudre la situation.	**Objectifs :** évaluer consiste à prendre un échantillon représentatif d'objectifs spécifiques et opérationnels et à générer un échantillon de questions qui traduisent au mieux cet échantillon d'objectifs
Type de questions	Tâches complexes et inédites	Nombreux items choisis qui traduisent un échantillon représentatifs d'objectifs.
Ce que l'on évalue	La mobilisation des acquis	La maîtrise de procédures. La maîtrise de contenus.
Validité et fiabilité	Privilégie la pertinence à la validité et la fiabilité Nécessité d'inventer de nouvelles techniques. Validité de contenu.	Etude de la fiabilité et de la validité suivant la théorie des scores
Le rôle de la situation	La situation fait partie intégrante de la résolution	Pas de situation ou situation « prétexte » ou « habillage »
Formes de questions	Questions ouvertes longues (page blanche)	Questions fermées ou ouvertes à réponses courtes

Si les épreuves d'évaluation nationales ou internationales ont comme but d'aider les responsables politiques à piloter le système éducatif, il apparaît que nos résultats ne sont pas anodins. Ils interrogent directement la pertinence des évaluations.

Dans notre recherche, il nous est apparu que la question de la pertinence de l'évaluation était peu débattue. Pour des raisons de validité et de fiabilité, de nombreuses épreuves qui affirment évaluer des compétences restent enfermées dans des carcans statistiques construits dans le contexte de la pédagogie par objectifs et qui se révèlent être peu pertinents dans une approche par compétence.

L'approche par compétence en modifiant ce que l'on attend des élèves interroge directement l'évolution des formes d'évaluation. Dans les pays qui se sont engagés dans une approche par compétence, peut-on garder des principes de construction d'épreuves issues en grande partie de la pédagogie par objectifs pour évaluer les compétences ? Les épreuves d'évaluation internationale qui affirment évaluer des compétences sont-elles réellement pertinentes ? Telles sont les questions auxquelles il nous semble important d'apporter rapidement des réponses.

Références

Attali, A. et Bressoux, P. (2002). *L'évaluation des pratiques éducatives dans les premiers et seconds degrés.* Paris : rapport établi à la demande du Haut Conseil de l'Evaluation de l'Ecole.

Beckers, J. (2001). Aider les élèves à développer des compétences à l'école : évolution ou continuité ? in Puzzle. *Centre interfacultaire de formation des enseignants*, n° 10.

Beckers, J. (2002). Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité. Bruxelles. Labor : 2002.

Bosman, Ch., Gérard, F.M., Roegiers X. (Eds) (2000). *Quel avenir pour les compétences ?* Bruxelles : De Boeck Université

Bressoux, P. (1994). Les recherches sur les effets-écoles et les effets maîtres. *Revue Française de Pédagogie*, n°108

Bronckart, JP & Dolz, J. (2002). La notion de compétence : quelle pertinence pour l'étude de l'apprentissage des actions langagières ? In Dolz, J et Ollagnier, E. (EDS) *L'énigme de la compétence en éducation*. Bruxelles : De Boeck Université

Carette, V, Defrance, A, Kahn, S., Rey, B. (2002)**.** *Les compétences à l'école fondamentale : une révolution qui suscite des questions***.** Acte du colloque du 2ème congrès des chercheurs en Education : Bruxelles.

Carette, V. (2005). *L'échec scolaire en Communauté française : un problème de culture. Le journal de l'Alpha,* n° *148.*

Carette,V., Defrance, A., Kahn, S., Rey, B. (2002). *Et si on évaluait des compétences ?* Actes du 15ème colloque international de l'ADMEE-Europe : Lausanne.

Crahay, M (Eds). (1994). *Evaluation et analyse des établissements de formation. Problématique et méthodologie*. Bruxelles : De Boeck.

Crahay, M. (2006). *Dangers, incertitudes et incomplétude de la logique de compétence*. Liège : Les Cahiers du service de Pédagogie expérimentale.

D'Hainaut, L. (1983). *Des fins aux objectifs.* Bruxelles: Labor.

De Ketele, J.M. (2000). En guise de synthèse : Convergences autour des compétences. In Bosmean, C., Gérard, F.M., Roeigiers, X. (Eds). *Quel avenir pour les compétences?* Bruxelles : De Boeck.

Felouzis, G. (1997). *L'efficacité des enseignants*. Paris, PUF.

De Ketele, J.M. et Gérard, F.M. (2004). La validation des épreuves d'évaluation selon l'approche par compétences. *Mesure et évaluation en éducation*. (à paraître).

Gauthier, C. (Ed.) (1997). *Pour une théorie de la pédagogie. Recherches contemporaines sur le savoir enseignant*. Bruxelles : De Boeck.

Hameline, D. (1979). *Les objectifs pédagogiques en formation initiale et continue.* Paris : ESF.

Lafontaine, D. (2002). Au-delà des performances des jeunes de 15 ans, Un système éducatif se profile… *Le point sur la recherche en Education*, n° 24

Martineau, S., Gauthier, C., & Desbiens, J. F. (1999). La gestion de classe au coeur de l'effet enseignant. *Revue des Sciences de l'Éducation*, 25

Ministère de la Communauté française (1999). *Socles de compétences Enseignement fondamental et premier degré de l'enseignement secondaire.*

Bruxelles : Administration générale de l'Enseignement et de la Recherche scientifique.

PISA 2003 (2004). *Evaluation de la culture mathématique des jeunes de 15 ans.* Bruxelles : Ministère de la Communauté Française. Service Général du Pilotage du Système Educatif.

Rey, B. (1996). *Les compétences transversales en question*. Paris : ESF

Rey, B., Carette, V., Defrance, A., Kahn, S. (2002). *Les compétences à l'école. Apprentissage et évaluation.* Bruxelles : De Boeck.

Roegiers, X. (2001). *Une pédagogie de l'intégration, Compétences et intégration des acquis dans l'enseignement*. Bruxelles : De Boeck-Université.

Roegiers, X. (2004). L'école et l'évaluation. Des situations pour évaluer les compétences des élèves. Bruxelles : De Boeck

Scallon, G. (2004). L'évaluation des apprentissages dans une approche par compétences. Bruxelles : De Boeck

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

DEUX DECADES POUR BROUILLER LES FRONTIERES ENTRE L'EVALUATION FORMATIVE ET L'EVALUATION SOMMATIVE

Aldina Silveira Lobo, FCPE Portugal

La question

Au début tout était simple. Le mot *évaluation* n'existait pas, même si le concept, lié à la classification, le remettait hors processus enseignement/apprentissage. L'expression *évaluation formative* (EF), née de Scriven et adoptée par les behavioristes Bloom, Hastings et Madaus (1971), s'est intégrée dans une autre famille. Les cognitivistes et constructivistes l'ont adoptée, eux aussi. Maintenant, elle habite les classes avec l'évaluation sommative (ES) et on a du mal à comprendre cette relation: qui est le responsable de quoi? Quel espace occupent l'une et l'autre? Quelles fonctions, outils, procédures alimentent chacune? Leurs frontières sont brouillées! C'est vrai qu'on a toujours dissocié les deux, mais avec l'élargissement progressif du glossaire évaluatif, la reconnaissance des pressions des examens sur l'évaluation interne et la diversification du rôle du maître, il devient évident que, les deux évaluations, elles cohabitent, en tension, l'espace classe.

Voyons le sens du mot *formative* appliqué à l'évaluation des apprentissages. Allal (1991) fait une très bonne comparaison entre l'EF des neo-behavioristes et celle des cognitivistes. Pour les premiers l'enseignement était organisé dans des petites unités qui surgissaient par des étapes séquentielles. A la fin de chacune on faisait une EF de régulation rétroactive, pour qu'on pût prendre des décisions dans l'avenir. Le présent était complètement comprimé entre le passé et le futur. Le modèle défendu par Allal était un modèle de régulation interactive, un modèle du présent qui donne au professeur et à l'élève un rôle actif, de gestion des apprentissages. Dans cette mesure, il est encore aujourd'hui très vivant.

Dans les années 80, on soulignait le caractère formative de l'évaluation ; elle devrait être continue, descriptive, qualitative… les chiffres n'appartenaient qu'au monde sommatif. De ceci, on parlait très peu, car il était déjà connu et identifié avec l'évaluation hors classe. Les professeurs essayaient de mettre en place la nouvelle

évaluation, la méconnue. C'était très difficile de le faire ; en général, ils ne réussissaient pas. Voilà pourquoi d'autres expressions sont parues. Les coordonnées de l'*apprentissage assisté par l'évaluation* (Hadji, 1990) ou de *l'évaluation formatrice* (Nunziati, 1990) révèlent des réflexions concernant l'application de l'EF. L'évaluation formatrice stressait encore plus vivement le rôle de l'élève, comme le fait Earl (2003), qui, plus récemment, a présenté son concept d'*évaluation comme apprentissage*.

Cependant, on s'aperçoit, dans cette dernière décennie, que la spécification ce fait par rapport à l'ES, soit-elle interne ou externe. Je parle de l'*évaluation alternative* (Gipps & Stobart, 2003) – alternative à la sommative –, de l'*évaluation pour les apprentissages* (Black & Wiliam, 1998a) – par opposition à celle des apprentissages –, de l'*évaluation authentique* (Wiggins, 1989), de *la contextuelle* (Berlak et al., 1992), de *la formative alternative* (Fernandes, 2005) – par opposition à celle faite exprès, à un moment donné.

Le problème, c'est qu'en regardant les classes et en écoutant les professeurs qui sont sur le terrain, on se rend compte qu'ils doivent accomplir plusieurs fonctions : former, aider à comprendre, réguler, informer, certifier, sélectionner. C'est-à-dire, il y a des fonctions qui se lient soit à l'EF soit à l'ES. Ainsi, surtout dans ce siècle, on reconnaît l'inévitabilité de la cohabitation de ces deux systèmes, même si quelques uns veulent donner plus d'importance à l'EF, car son développement permet de réussir les deux évaluations (Biggs, 1998), (Black, 2003), (Brookhart, 2001), (Fernandes, 2005), (Gipps, 2004), (Harlen, 2006), (Perrenoud, 1999), (Roegiers, 2004), (Shepard, 2001), (Stiggins, 2002), (Wiliam, 2000).

Alors, la question qu'on se pose aujourd'hui, c'est comment se rapportent les deux évaluations ? En classe, où termine l'une et où commence l'autre ? En réalité, ce qui intéresse, c'est d'avoir des réponses pour quelques questions : pour accomplir les différentes fonctions, peut-on s'en servir des mêmes outils ? des mêmes techniques ? des mêmes données ? Comment sait-on si on travaille l'une ou l'autre ?

Définition de termes

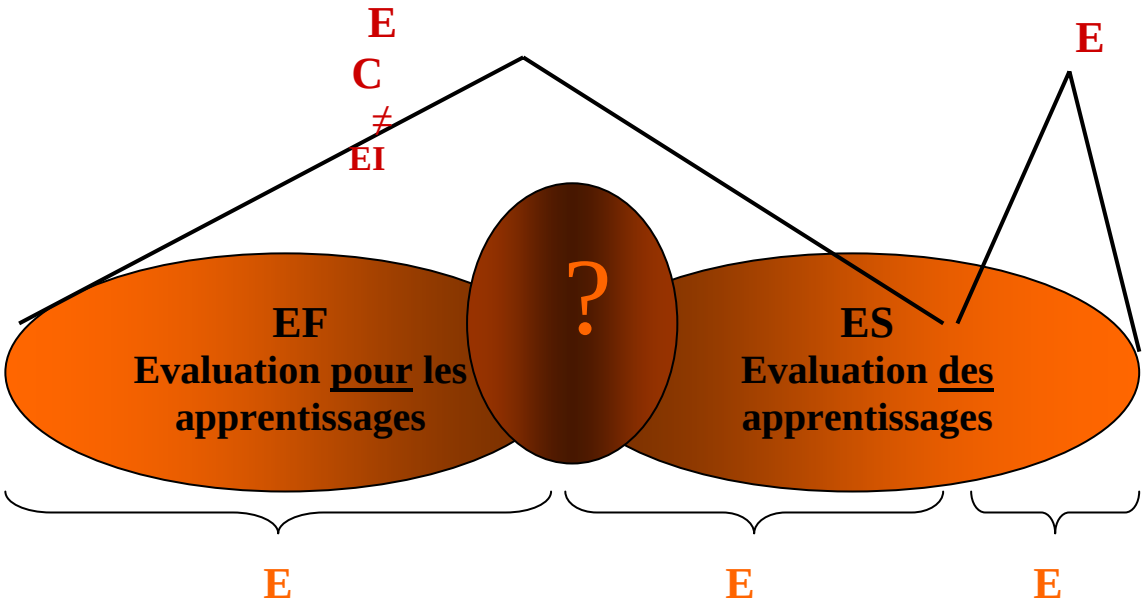
Commençons par l'éclaircissement des concepts. Si on prend les définitions données par plusieurs auteurs, on obtient ce qui suit :

- EF – c'est « un type d'évaluation menée par l'enseignant, qui, en cours d'apprentissage, vise à détecter les difficultés de l'élève en vue d'y remédier » (Roegiers, 2004, p.58). Elle inclut toutes les activités menées par professeurs et élèves qui fournissent des donnés utilisés comme *feedback*, avec le but de modifier les activités d'enseignement et d'apprentissage de façon à raccourcir le *gap* entre la situation actuelle de l'élève et les niveaux de référence désirables (Black & Wiliam, 1998a).
- ES – celle qui « arrive à la fin d'une période ou d'un cours et qui est usée pour donner de l'information sur combien les élèves ont appris et sur la qualité du fonctionnement du cours » (Gipps, 2004, p.vii). « C'est une réponse à des pressions externes et au besoin d'*accountability* » (Brookhart, 2001, p.157).
- Evaluation interne (EI) – « celle qui est de la responsabilité totale des écoles et des professeurs » (Fernandes, 2005, p.20).
- Evaluation externe (EE) – « celle qui est de la responsabilité et initiative totale d'une entité extérieur à l'école » (Fernandes, 2005, p.21).
- Evaluation en classe (EC) – « celle soutenue par des activités appartenant aux programmes éducatifs suivis par les élèves. Ces activités peuvent être développées soit dans soit hors la salle de classe » (Crooks, 1988, p.439). On y inclut des tâches pour résoudre des tests formels, des questionnaires oraux, etc. On y inclut aussi l'évaluation des variables attitudinales et motivationnelles des apprentissages.

Cela veut dire que la correspondance classique n'est plus valable : l'EF n'est pas synonyme d'EC car celle-ci correspond à l'EF plus une partie de l'ES ; et l'ES n'est pas synonyme d'EE car celle-ci ne correspond qu'à une partie de l'ES. Même l'EI n'est pas forcément pareil à l'EC, une fois que les décisions venant des conseils

de gestion ou des conseils pédagogiques sont exclues de l'EC. La figure 1 peut éclaicir ce raisonnement.

La question, donc, c'est de savoir ce qui se passe en classe. Maintenant, c'est clair qu'il y en a deux sortes d'évaluations : l'EF et l'ESI (évaluation sommative interne), qui sont prises en tension. Cependant, l'EE exerce aussi sa pression sur les deux. Plusieurs études l'assurent (Aitken, 2000), (Atkin, Black, & Coffey, 2001), (Black, 2003), (Borko & Elliott, 1998), (Clinciu, Cocorada, Luca, & Pavalache-Ilie, 2005), (Gipps & Stobart, 2003), (McMillan, Myran, & Workman, 1999), (Wiliam, 2000)... Mais comment jouent-elles ? Y en a-t-il des frontières bien définis ? Y en a-t-il des outils adéquats à chacune ? Y en a-t-il des moments express à chacune ?



EC – évaluation en classe
EF – évaluation formative
EI – évaluation interne
EE – évaluation externe
ESE – évaluation sommative externe
ESI – évaluation sommative interne

Figure 1- Relation entre les différents concepts

Cohabitation inévitable

Traditionnellement, le contrôle des apprentissages se faisait après le processus enseignement, mais surtout dès l'existence des constructivistes que l'importance de l'évaluation en classe est stressée. Ils ont fait anticiper ce contrôle en le poussant vers le moment des apprentissages. L'évaluation est même devenue le soutien de l'apprentissage (Crooks, 1988), (Hadji, 1990). Et pourquoi ce changement ? Parce qu'il faut profiter l'EF pour augmenter la qualité des acquis.

En accord avec Natriello (1987), l'évaluation des acquis des élèves a quatre fonctions générales : certification, sélection, direction et motivation. La première sert à identifier le niveau atteint par l'élève ; la deuxième pour le faire accéder à des parcours scolaires et professionnels ; la troisième pour que l'évaluateur communique ses désirs à l'évalué et refasse son diagnostique de façon à mieux planifier ; et, la motivation implique les évalués dans la tâche.

Cardinet (1994) aussi, il trouve quatre fonctions principales de l'évaluation pédagogique : a) améliorer les décisions sur l'apprentissage de chaque élève ; b) informer l'élève et ses parents de ses progrès ; c) passer des certificats réclamés par l'élève et la société ; d) améliorer la qualité générale de l'enseignement. Comme l'auteur précédent, deux des fonctions exigent une EF et deux autres une ES. Cependant, chez Cardinet la participation des parents et de la société à l'école est plus évidente.

Task Group on Assessment and Testing, dans un rapport présenté en 1988 (Harlen, 2006), a mis en évidence quatre fonctions de l'évaluation : a) la formative, qui facilite l'apprentissage pour permettre de préparer les pas suivants ; b) la diagnostique, qui identifie les difficultés d'apprentissage ; c) la sommative, qui offre des synthèses des réussites à chaque moment ; d) l'évaluative, qui utilise l'évaluation des élèves pour des différents rapports de classe, d'école ou du système éducatif. Cette conceptualisation de l'évaluation a été reprise par plusieurs auteurs dont Wiliam (2000) et Harlen (2006) sont exemple.

Black (2003) relève trois fonctions de l'enseignement : la formative, pour aider à l'apprentissage ; la sommative interne pour certifier et permettre les déplacements ; c) la sommative pour *accountability*. Il croit que les trois interagissent et que, même si l'existence des tensions est inévitable, leur cohabitation peut apporter des synergies.

Earl (2003) aussi les a prises en identifiant l'évaluation : a) à des fins sommatives (liée à l'évaluation des apprentissages) pour certifier et informer les parents et les élèves des progrès, en signalant sa position relative par comparaison à ses copains ; à des fins formatives (l'évaluation pour les apprentissages) pour faire des descriptions qui permettent de définir les pas suivants ; c) à des fins encore plus formatives (évaluation comme apprentissage) pour inciter l'apprenant à sentir sa perspective critique face à l'évaluation et à l'apprentissage. Cette *évaluation comme apprentissage* de Earl me fait rappeler l'évaluation formatrice que Scallon et Nunziati (Nunziati, 1990) ont trouvé pour relever le rôle de l'élève dans le processus.

En dépit de l'intérêt de l'équilibre de la pyramide de Earl, j'interprète le concept *évaluation pour les apprentissages* comme ayant forcément la participation active de l'élève. Voilà pourquoi je préfère la distinction qu'on a commencé à généraliser à partir des travaux de Black et Wiliam (1998a), (1998b) et des membres de Assessment Reform Group (ARG, 1999), (ARG, 2002), (ARG, 2002) : l'*évaluation pour les apprentissages* et l'*évaluation des apprentissages*.

En bref, le professeur doit aider l'élève à comprendre les informations recueillies à travers d'un *feedback* qui lui permet de mieux gérer ses acquis. Face à des diagnostiques fréquents il est possible de mieux définir les compétences à développer, sélectionner les outils adéquats aux circonstances, délimiter des contenus curriculaires, évaluer des stratégies d'enseignement, motiver, contrôler des frustrations… Ainsi, on détache trois phases cruciales : la cueillette des évidences d'apprentissage (il faut bien choisir les méthodes et techniques), l'interprétation de ces données (il faut trouver les critères de référence par rapport au groupe, aux critères définis ou à la personne) et la dévolution de l'information traitée (il faut faire attention à la communication du *feedback*). Généralement, le professeur doit signaler la différence entre l'accomplissement de l'élève et le désirable, attribuer le succès à son effort et expliciter à nouveau le concept à apprendre. Comme atout, on relève les

moments de réflexion critique, de construction de portfolios, des tâches qui exigent des réponses longues, des opinions etc. Malheureusement, les recherches rendent compte de pratiques peu approfondies et routinières où prédomine la fonction classificatrice de l'évaluation.

Du côté de l'évaluation des apprentissages, il faut retenir la pensée de Harlen, Gipps, Broadfoot et Nuttal qui la partage en deux : l'ES cumulative (*summing-up*) et l'ES de contrôle (*checking-up*). La première offre une image du succès de l'évalué venant de l'addition de différents données tout au long du chemin parcouru à des fins formatives ; la seconde prend place avec la résolution des tests ou d'autres tâches avec le but spécifique d'enregistrer les acquis à un moment donné (Gipps, 2004). L'ES peut, donc, être appliquée à l'aide de l'EF.

Le problème c'est que dans la plupart des Etats, l'EE (qui est sommative de contrôle) a un effet très lourd sur la décision de certification et de sélection, même au Portugal où elle ne pèse que 30%. Il est facile de se laisser influencer par la pression de l'entraînement à l'examen, avec les mêmes types de questions et de buts. Les professeurs se plaignent de la manque de temps pour tout faire et optent, en général, par privilégier la préparation à l'EE en dépit des tâches plus formatives comme le développement des apprentissages approfondies. La tension vécue par les professeurs réside sur le fardeau de l'option EF/ES comme plusieurs recherches l'attestent (Aitken, 2000), (McMillan et al., 1999)... La présence de l'ES de contrôle en classe reste incontestable.

Avec cette évolution du concept de l'ES, les frontières entre l'EF et l'ES (I) sont devenues diffuses. L'ES continue à être ponctuelle mais plus fréquente. Elle augmente ses fonctions pour assister l'apprentissage. Par contre, sa fonction d'*accountability*, chaque fois plus exigeante, appuyé sur des tests ou des examens externes, s'isole en gardant ses caractéristiques traditionnelles.

On conclue, donc, que l'EF et l'ES cohabitent inévitablement en classe.

Relation intime

En reconnaissant l'existence, en classe, des deux évaluations, certains auteurs les considèrent des systèmes indépendants, parallèles, mais la plupart d'entre eux, surtout depuis le virement du siècle, défendent son interaction où l'EF peut même prendre le devant (Biggs, 1998), (Black, 2003), (Gipps, 2004), (Harlen, 2006), (Stiggins, 2002), (Wiliam, 2000).

Un des modèles de cohabitation est celui de Harlen (2006). Elle considère qu'une même activité de l'élève conduit à des évidences qui peuvent être traitées différemment selon le but désiré: formatif ou sommatif. Si seul le professeur utilise les évidences en les comparant avec les critères de classe définis préalablement et il s'y arrête en synthétisant son appréciation dans un numéro ou une lettre, il travaille sur l'ES. Si le professeur les utilise en les comparant soit avec ces mêmes critères de classe soit avec les critères spécifiques de l'élève et les deux définissent les pas à suivre, il est dans le formatif. Dans ce cas, il est souhaitable que l'élève et son professeur soient complices dans tout le parcours. Le danger c'est de changer de l'un à l'autre sans refaire la lecture des données. Il faut en faire attention !

Mais ce n'est pas si simple. Parfois on les mélanges et il est difficile à dire où on est. En effet, Harlen croit que l'EF et l'ES évoluent dans une dimension qui va de l'EF informelle à l'ES formelle, en passant par l'EF formelle et l'ES informelle. Le problème c'est d'identifier les frontières entre ces deux dernières. Si on regarde le tableau 1, on trouve beaucoup de cases pareilles. Le professeur prend part dans toutes les situations, mais, pour ces deux, c'est un seul jugé (le professeur) qui organise son travail de façon à recueillir des *évidences naturelles*, c'est-à-dire, intégrées dans les activités d'enseignement/apprentissage. Il s'en sert aussi pour refaire ses plans. Ce sont surtout le focus et le but qui changent, en accord avec ce qu'on a déjà vu sur l'EF: l'EF formelle concerne des décisions sur le futur, l'ES informelle sur les acquis, sur le passé, mais les deux opèrent au présent, pour informer et réguler. Ici elles se mélangent : l'information appartient à l'EF ; la régulation, normalement attribuée à l'EF, surgit dans le domaine de l'ES. Cependant, elles ne s'appliquent pas directement

aux apprentissages mais aux plans de l'enseignement, ayant un but formatif. Il semble que la traditionnelle action prévue chez l'EF, surtout quand on parle d'interaction et de régulation, s'est perdue, en cédant de l'espace à l'ES. De quelque façon, il y a toujours (ou presque) un *feedback* fondamental.

Tableau 1 : Une dimension possible des buts et pratiques évaluatives (Harlen, 2006)

	EF			ES
	Formative informelle	Formative formelle	Sommative informelle	Sommative formelle
Focus	Quels sont les pas suivants.?		Qu'est-ce qui a été achevé jusqu'à maintenant?	
Objectif	Informer sur les pas d'apprentissage qui suivent	Informer sur les pas d'enseignement qui suivent	Réguler le processus en fonction des plans	Enregistrer les réussites des personnes
Evidence recueillie	Appartenant au travail habituel de classe	Introduite dans le travail habituel de classe	Introduite dans le travail habituel de classe	Tâche ou test à part
Bases/reference du jugement	Par référence à l'élève	Par référence à l'élève et à des critères	Par référence à des critères	Par référence à des critères
Jugé par	Élève et professeur	Professeur	Professeur	Professeur ou évaluateur externe
Action prise	*Feedback* aux élèves e au professeur	*Feedback* pour les plans du professeur	*Feedback* pour les plans du professeur	Rapport aux élèves, parents, autres profs...
Epithète	evaluation pour les apprentissages	Combinaison (*Matching*)	Plongée ? (*Dip stick*)	Evaluation des apprentissages

Ainsi, à partir des mêmes données, on peut faire appliquer soit une EF soit une ES dépendant des critères de comparaison et de l'usage qu'on donne à son interprétation. Selon Harlen, cela dépend aussi du fait d'impliquer (ou pas) l'élève dans les décisions pendant le processus.

Le brouillard s'est installé

En résumé, l'ancien axe d'analyse temporel n'est plus déterminant. En gros, la question ne se pose pas sur quand fait-on l'évaluation, mais à quoi servent les données obtenus, car une partie de l'ES s'est déplacée vers la salle de classe, en créant ainsi, dans beaucoup d'Etats (où il y a des examens), trois dimensions de l'évaluation : la formative, la sommative interne et la sommative externe. L'ESI peut être formelle ou informelle. L'informelle envahit l'espace de l'EF, en lui prêtant même ses verbes identificateurs d'action.

Il est curieux de voir que l'envers était autrefois valable. Au Portugal, par exemple, vers la fin du siècle XIX et le début du XXème, les enfants étaient obligés de passer des examens mais, pour la plupart des niveaux scolaires, la classification finale devrait surgir de la moyenne des notes des épreuves écrites, des oraux et aussi des notes de comportement et de l'avancement dans les études enregistrées pendant l'année scolaire.

La relation EF/ES qui, dans le plan théorique, était nette il y a 20 ans, reste aujourd'hui à l'intérieur d'un brouillard épais. Il se peut que, en réalité, rien n'aie changé, mais, en théorie, et en ayant la préoccupation de la transparence en évaluation, l'EC est devenue beaucoup plus complexe. A nous de pénétrer les classes et d'en déceler les frontières.

Références

Aitken, R. (2000). Teacher perceptions of the use and value of formative assessment in secondary English programmes. 15-20.

Allal, L. (1991). Stratégies d'Evaluation Formative: Conceptions Psicopedagogiques et modalités d'aplication. L. Allal, J. Cardinet, & P. Perrenoud L'évaluation formative dans un enseignement différencié - Actes du colloque ŕ l'Université de Genčve, mars 1978 (6.a. ed.,). Berna: Peter Lang SA.

ARG, A. R. G. (1999). Assessment for Learning - Beyond the black box. Cambridge : University of Cambridge Faculty of Education.

ARG, A. R. G. (2002). Assessment for Learning: 10 principles. 2. ARG, Assessment Reform Group.

Atkin, J. M., Black, P., & Coffey, J. E. (2001). The Relationship Between Formative and Summative Assessment - In the Classroom and Beyond. Classroom assessment and the National Science Education Standards, 1-14. National

Research Council.

Berlak , N., Newmann, F., Adams, E., Archbald, D., Burgess, T., Raven, J., & Romberg, T. E. (1992). Toward a new science of educational testing and assessment. New York: State University of New York Press.

Biggs, J. (1998). Assessment and Classroom Learning: a role for summative assessment? Assessment in Education, 5(1), 103-110.

Black, P. (2003). Testing: Friend or Foe? Routledge Falmer.

Black, P., & Wiliam, D. (1998a). Assessment and Classroom Learning. Assessment in Education, 5(1), 7-74.

Black, P., & Wiliam, D. (1998b). Inside the Black Box: Raising Standards Through Classroom Assessment . Londres: nferNelson.

Bloom, B., Hastings, J., & Madaus, G. (1971). Handbook of Formative and Sumative Evaluation of Student Learning. E.U.A.: McGraw-Hill.

Borko, H., & Elliott, R. (1998). Tensions between Competing Pedagogical and Accountability Commitments for Exemplary Teachers of Mathematics in Kentucky. 1-13. National Center for Research on Evaluation, Standards, and Student Testing (CRESST) - UCLA.

Brookhart, S. M. (2001). Successful Students' Formative and Summative Uses of Assessment Information. Assessment in Education, 8(2), 58-65.

Cardinet, J. (1994). Pour apprécier le travail des élèves. Bruxelas: DeBoeck Université.

Clinciu, I., Cocorada, E., Luca, M., & Pavalache-Ilie, M. (2005). Effets pervers des évaluations sommatives. Comment évaluer? Outils, dispositifs et acteurs - actes du 18čme Colloque International de l'ADMEE-Europe Reims: ADMEE-Europe//IUFM Champagne Ardenne.

Crooks, T. (1988). The Impact of Classroom Evaluation Practices on Students. Review of Educational Research, 58(4), 438-481.

Earl, L. (2003). Assessment as Learning: Using Classroom Assessment to Maximise Student Learning. Thousand Oaks, CA: Corwin Press.

Fernandes, D. (2005). Avaliação das Aprendizagens: desafios as teorias, práticas e políticas. Cacém: Texto Editores.

Gipps, C., & Stobart, G. (2003). Alternative Assessment. T. Kellaghan, & D. E.

Stufflebeam International handbook of educational evaluation, 549-576. Dordrecht: Kluwer.

Gipps, C. (2004). Beyond Testing: Towards a Theory of Educational Assessment. Oxon: Routledge Falmer.

Hadji, C. (1990). L'apprentissage assisté par l'évaluation (A.A.E.) mythe ou réalité. Cahiers Pédagogiques , 281 (Fev.), 20-23.

Harlen, W. (2006). On the Relationship Between Assessment for Formative and Summative Purposes. J. Ed. Gardner Assessment and Learning, 103-117. Londres: Sage Publ.

McMillan, J., Myran, S., & Workman, D. (1999). The Impact of Mandated Statewide Testing on Teacher's Classroom Assessment and Instructional Practices.

Natriello, G. (1987). The Impact of Evaluation Processes on Students. Educational Psychologist, 22(2), 155-175.

Nunziati, G. (1990). Pour construire un dispositif d'évaluation formatrice. Cahiers Pédagogiques, 280, 47-64.

Perrenoud, P. (1999). L'évaluation des élèves. De la fabrication de l'excellence à la régulation des apprentissages. Entre deux logiques. Bruxelas: DeBoeck Université.

Roegiers, X. (2004). L'école et l'évaluation, des situations pour évaluer les compétences des élčves. Bruxelas: De Boeck.

Shepard, L. A. (2001). The Role of Classroom Assessment in Teaching and Learning . V. Ed. Richardson The Handbook of Research on Teaching (4ª. ed., 1066-1101). Washington, DC: American Educational Research Association.

Stiggins, R. (2002). Assessment Crisis: The Absence Of Assessment FOR Learning. Phi Delta Kappan, 1-11.

Wiliam, D. (2000). Integrating formative and summative functions of assessment. 1-13.
Notes: Texto sapresentado ao Grupo de Trabalho 10 do Congresso Internacional sobre Educação Matemática, Makuhari, Tokyo, Agosto 2000

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

L'ÉVALUATION FORMATIVE : VERS UNE CLARIFICATION CONCEPTUELLE

Julie Desjardins, Colette Deaudelin, Catherine Allen, Université de Sherbrooke

Contexte

Depuis une dizaine d'années, de nombreux pays ont entrepris de réformer leur système d'éducation. Influencés par les récents courants de recherche sur l'apprentissage, la plupart des pays ont choisi d'inscrire la formation des jeunes dans un paradigme de développement de compétences, paradigme qui commande d'accroître le rapport entre les savoirs et leur mobilisation en contexte. Dans ce contexte de renouvellement pédagogique l'évaluation des apprentissages est revisitée et des changements sont prescrits. Les objets, les formes et les fonctions, de même que la place de l'évaluation dans le processus d'enseignement-apprentissage sont considérés sous un angle nouveau. L'importance de l'évaluation formative est réaffirmée. Une étude de l'OCDE (2005) en indique d'ailleurs le caractère essentiel dans le contexte des réformes. Au Québec, les récentes orientations soulignent à grands traits la nécessité d'inscrire les pratiques d'évaluation dans une démarche de soutien à l'apprentissage. Les vertus de l'évaluation formative sont clamées haut et fort et les enseignants sont invités à remodeler leurs pratiques d'évaluation pour les aligner sur le soutien à l'apprentissage.

Problématique et objectifs de recherche

L'idée de pratiquer une démarche d'évaluation à visée formative n'est toutefois pas récente (Desjardins, 2005). Plusieurs auteurs ont contribué à la diffusion de cette perspective depuis au moins une vingtaine d'années même si en pratique, on observe aujourd'hui que cette démarche d'évaluation est rarement ou peu intégrée aux pratiques des enseignants (Conseil supérieur de l'éducation, 1992 ; OCDE, 2005 ; Shephard, 2000)

Dans le cadre d'une recherche menée entre 2004 et 2006 et visant à décrire des pratiques d'évaluation formative chez des enseignants du primaire, nous avons aussi constaté que, derrière un apparent consensus sur la signification du concept d'évaluation formative, il s'agissait d'un vaste champ conceptuel comportant de nombreuses variations sémantiques. Il nous est donc apparu nécessaire de procéder à une démarche d'analyse conceptuelle visant à baliser le champ des significations de ce concept.

Méthodologie

Dans cet objectif de clarification conceptuelle, nous avons procédé à une recension des écrits sur la notion d'évaluation formative. Outre ce concept, le repérage des écrits s'est fait à partir d'une liste de termes complémentaires étroitement associés au concept ciblé. Ces termes sont : rétroaction, régulation, autorégulation, autoévaluation et métacognition.

Une première opération a permis de repérer, dans les banques de données électroniques, des textes parus entre 1990 et 2005 et comportant une forte composante conceptuelle à propos des concepts ciblés. Par la suite, une démarche en cascade a été utilisée : les listes de références des textes repérés lors de la première opération ont conduit à la recension d'autres travaux pertinents à l'étude. Quelques textes parus avant 1990 ont été retenus lors de cette opération lorsqu'il s'agissait de textes cités par de nombreux auteurs. Au terme de cette démarche 99 textes ont été retenus (38 en français et 61 en anglais) dont la provenance est variée (Canada, États-Unis, Suisse, France, Belgique, Royaume Uni, Nouvelle-Zélande, Norvège, Hollande, Corée du Sud, Italie, Espagne, Australie et Liban). L'échantillonnage ayant permis la constitution du corpus s'appuie sur le principe de «saturation de l'information» (Van der Maren, 1995). L'analyse des textes a été réalisée à partir d'une grille permettant de repérer : les définitions des processus évoqués, les objets ciblés, les acteurs impliqués, les éléments de contexte à considérer, ainsi que les cadres théoriques sous-jacents. Une attention a aussi été apportée à la présence de concepts complémentaires auxquels ont recours les auteurs lorsqu'ils référent à l'évaluation formative.

Résultats

Dans le cadre de la présentation, nous ferons état du vaste champ conceptuel dans lequel s'inscrit l'évaluation formative. Nous distinguerons notamment les acceptions différentes de l'évaluation formative et des concepts complémentaires selon divers cadres théoriques de l'apprentissage. Par exemple, au regard des objets ciblés par l'évaluation, l'émergence des théories cognitivistes a contribué à une centration sur les processus de l'élève, voire sur les microprocesssus cognitifs et à considérer la régulation comme un processus interne à tout système (Torrance et Pryor, 1998; Behrens, 1997; Gagné et Thouin, 1991, Cizek, 1997). Le cadre béhavioriste suggère plutôt une focalisation sur la réponse de l'élève. Quant au cadre constructiviste, il a donné lieu à la notion de viabilité. La présentation permettra de signaler les nuances, les complémentarités, les oppositions, ainsi que les zones d'ombre qui nous paraissent nécessiter des éclaircissements théoriques.

En discussion, nous signalerons les limites de certaines conceptualisations, notamment lorsque les auteurs omettent de les inscrire dans une perspective théorique. Nous examinerons comment cette omission conduit à des conceptualisations soit composites ou approximatives, qui nuisent à la clarté des concepts en jeu.

Nous terminerons en soutenant la nécessité d'une clarification conceptuelle pour assurer une plus grande pénétration de l'évaluation formative dans les pratiques enseignantes. Une description fine et nuancée du phénomène et des processus sous-jacents fournit des outils conceptuels indispensables aux formateurs et aux praticiens pour assurer le développement de pratiques d'évaluation efficaces et véritablement centrées sur l'apprentissage.

Références

Behrens, M. (1997). Le portfolio, un moyen d'améliorer le dialogue entre formateur et apprenant. *Mesure et évaluation en éducation,* 20(3), 55-83.

Cizek, G. J. (1997). Learning, achievement, and assessment. In G. D. Phye (dir.), Handbook of classroom assessment : learning, adjustment, and achievement (p. 1-32). États-Unis: Academic Press.

Conseil supérieur de l'éducation (1992). Évaluer les apprentissages au primaire : un équilibre à trouver. Avis au ministre de l'éducation. Bibliothèque nationale du Québec.

Desjardins, J. (2005). La réforme des pratiques d'évaluation des apprentissages : virage ou recommencement? In Y. Lenoir, F. Larose et C. Lessard (dir.), Le curriculum de l'enseignement primaire : regards critiques sur ses fondements et ses lignes directrices (p. 297-318). Sherbrooke : Éditions du CRP.

Gagné, F., et Thouin, M. (1991). L'évaluation formative des apprentissages en orthographe et l'attitude des élèves à l'égard de l'évaluation. *Mesure et évaluation en éducation,* 14(1), 5-16.

Organisation de coopération et de développement économiques (OCDE) (2005). L'évaluation formative. Pour un meilleur apprentissage dans les classes du secondaire. Éditions OCDE.

Shepard, L. A. (2000). The role of classroom assessment in teaching and learning. (CSE technical report 517). The regents of the university of California.

Torrance, H., et Pryor, J. (1998). *Investigating formative assessment : teaching, learning and assessment in the classroom*. Buckingham et Philadelphie: Open university press.

Van der Maren, J.-M. (1995). *Méthodes de recherche pour l'éducation*. Montréal : Presses de l'Université de Montréal.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

RÉDUCTION DE LA TAILLE DES CLASSES ET IMPLICATION DES ÉLÈVES DANS LES TÂCHES SCOLAIRES

Pascal Bressoux, Lima Laurent, Université Pierre Mendès France de Grenoble

Résumé :

La taille des classes et ses effets sur les acquis des élèves ne cesse de faire l'objet de débats, tant sociaux que scientifiques. De nombreuses études ont été consacrées à cette question, de qualité variée et conduisant à des résultats souvent mitigés. L'étude STAR aux USA, la seule de grande envergure ayant été totalement randomisée, a incontestablement apporté un ensemble de résultats très riches, révélant des effets positifs substantiels dans le court terme et rémanents dans le long terme de la réduction de la taille des classes sur les acquis des élèves (Finn, Gerber, Achilles & Boyd-Zaharias, 2001 ;Wang & Finn, 2000). En France, il semblait admis après le rapport de Meuret (2001) pour le compte du Haut Comité d'Evaluation de l'Ecole, que l'effet de la réduction de la taille des classes ne pouvait être sensible que s'il touchait les premiers niveaux scolaires, pour les élèves défavorisés et avec une réduction très importante des effectifs (donc très coûteuse). Le débat a toutefois été récemment relancé par une étude Thomas Piketty (2004) qui utilise une méthode proposée par Angrist et Lavy (1999) permettant de corriger les biais d'endogénéité (les classes à faible taille ne se distribuent pas aléatoirement ; on constate en particulier que, en moyenne, les élèves les plus faibles sont dans des classes dont la taille est moins élevée). Piketty a montré que les effets bénéfiques de la réduction de la taille des classes, en particulier pour les élèves les plus défavorisés, pouvaient être importants, même avec une réduction modeste des effectifs (donc peu coûteuse) : par exemple, il a montré que réduire la taille des classes en zone d'éducation prioritaire (ZEP) à 18 élèves (ce qui n'aurait conduit qu'à une très faible augmentation des effectifs dans les classes hors ZEP) pouvait réduire de 40 % les écarts d'acquisitions entre élèves ZEP et élèves non-ZEP. Cependant, on sait encore bien peu de choses sur les processus qui sont générés en classe par la réduction des effectifs. C'est à cette question qu'a été consacrée l'étude qui sera présentée dans cette communication.

Le Ministère de l'Education Nationale français a lancé en 2002-2003 une expérimentation d'envergure visant à réduire la taille des classes de CP (1ère année élémentaire) à 10 élèves dans les zones défavorisées. Un groupe de 100 classes expérimentales a vu sa taille de classe réduite à 10 élèves. Un groupe de 100 classes témoins, à effectif « normal » (21 élèves en moyenne par classe) a été choisi pour contrôler les effets de l'expérimentation. Les acquis scolaires des élèves ont été testés en début et en fin d'année scolaire. De plus, des observations ont été conduites à deux reprises dans chaque classe (en mars et en juin). Les observateurs devaient, entre autres, observer l'implication des élèves dans les tâches scolaires en utilisant la méthode d'échantillonnage temporel initialement proposée dans le dispositif d'observation OGP (cf. Altet, Bru, Bressoux & Leconte-Lambert, 1994, 1996). Trois

élèves sont observés tour à tour toutes les 20 secondes durant 20 minutes. Au moment où il est observé, l'élève est codé « impliqué » s'il manifeste les signes d'attention à la tâche prescrite, « non impliqué » s'il manifeste de façon évidente des signes de non-attention à la tâche prescrite), ou « indécidable » si les signes sont ambigus. On obtient ainsi un score d'implication sur 20 pour chaque élève observé. Malgré la relative grossièreté d'une telle procédure, les travaux antérieurs ont montré la robustesse des mesures ainsi obtenues (Altet, Bru, Bressoux & Leconte-Lambert, 1994,1996).

535 élèves ont été observés. Afin de contrôler l'impact des variables parasites et de respecter la structure hiérarchisée des données (des élèves groupés dans des classes), les données ont été traitées à l'aide de modèles multiniveaux.

Nos résultats révèlent que, toutes choses égales par ailleurs, les élèves qui ont été scolarisés dans une classe à effectif réduit ont en moyenne des acquisitions scolaires plus élevées en fin d'année que ceux du groupe témoin ; l'effet est néanmoins d'ampleur assez limitée. De plus, il apparaît que la réduction des effectifs a favorisé l'implication des élèves dans les tâches scolaires. Cette implication est elle-même positivement liée aux acquisitions scolaires des élèves. Ainsi, il semblerait possible de conclure qu'une partie de l'effet bénéfique de la réduction la taille des classes transite par une meilleure implication des élèves. Cette variable jouerait alors un rôle partiellement médiateur dans la relation entre taille des classes et acquis des élèves.

Références

Altet, M., Bressoux, P., Bru, M., & Lambert, C. (1994). Etude exploratoire des pratiques d'enseignement en classe de CE2. *Les Dossiers d'Education et Formations*, Ministère de l'Education Nationale, DEP, *44*.

Altet, M., Bressoux, P., Bru, M., & Lambert, C. (1996). Etude exploratoire des pratiques d'enseignement en classe de CE2 : Deuxième phase. *Les Dossiers d'Education et Formations*, Ministère de l'Education Nationale, DEP, *70*.

Angrist, J. D., & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *Quaterly Journal of Economics, 114*(2), 533-575.

Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record, 103*(2), 145-183.

Meuret, D. (2001). Les *recherches sur la réduction de la taille des classes*. Rapport pour le Haut Comité d'évaluation de l'Ecole.

Piketty, T. (2004). *L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997*. Paris : EHESS Paris-Jourdan.

Wang, M. C., & Finn, J. D. (2000). *How small classes help teachers do their best*. Philadelphia: Temple University Center for Research in Human Development and Education.

L'EVALUATION D'UN DISPOSITIF DE PROMOTION DE LA REUSSITE EN PREMIERE ANNEE D'ETUDES UNIVERSITAIRES EN BELGIQUE FRANCOPHONE : DE NOUVELLES PISTES ENVISAGEES.

Alice Salcin, Université libre de Bruxelles
Dorothée Baillet, Université libre de Bruxelles
José-Luis Wolfs, Université libre de Bruxelles

Mots-clés : Réforme (évaluation dispensatoire et changement de seuil de réussite), effet sur le taux de réussite, première année d'étude universitaire.

Résumé :

Suite aux accords de Bologne, la Belgique francophone a réformé, à travers le décret du 31 mars 2004, son enseignement supérieur en vue de favoriser son intégration dans l'espace européen. Parmi les mesures prises, deux visent à promouvoir la réussite des étudiants : l'introduction d'une évaluation dispensatoire en BA 1 et le changement de seuil de réussite relatif à chaque cours au sein d'une année académique. Lors de notre participation au colloque de l'ADMEE 2005, nous nous étions intéressés à l'impact de celles-ci. Suite à cette étude, nous avions présenté une série de pistes qui permettraient de mieux comprendre les résultats observés. Ce sont certaines de ces pistes que nous proposons d'explorer.

Premièrement nous avons tenté de vérifier si les enseignants ont modifié leurs modalités d'évaluation. A cette fin, nous avons vérifié – pour les cours déjà étudiés - si les « 10 » et « 11 » sont proportionnellement plus nombreux en 2004-2005 que pour les années précédentes.

Deuxièmement, nous avons essayé d'expliquer les raisons pour lesquelles l'augmentation du taux de réussite des cours (en tenant compte des « 10 » et « 11 ») ne se répercute pas au niveau des Facultés. Ainsi, nous avons analysé les résultats des cours pour lesquels il n'y avait pas d'évaluation dispensatoire (donc évalués pour la 1ère fois en juin) afin de vérifier s'ils ont été moins bien réussis en 2004-2005 que les autres années.

Enfin, nous avons cherché à vérifier l'hypothèse selon laquelle la session de juin a pu être alourdie par rapport aux autres années (par les échecs de l'évaluation dispensatoire). Pour explorer cette question, nous avons analysé les données de l'année académique 2005-2006 à partir desquelles nous avons tenté de mesurer la proportion d'étudiants dont la session de juin a effectivement été plus importante.

Introduction

Le taux de réussite moyen au terme de la première année d'études universitaires en Belgique francophone (toutes universités et toutes facultés confondues) se situe globalement autour de 40% (D. Leclercq, 2002).

Ces chiffres doivent être tempérés car la Belgique est un des rares pays de l'Union européenne à avoir mis en place un régime de libre accès aux études universitaires. La seule condition d'accès est le diplôme de fin d'études secondaires délivré par l'établissement scolaire lui-même.

Ces chiffres sont cependant suffisamment préoccupants pour mobiliser l'attention des chercheurs (Droesbeke, Hecquet & Wattelaer, 2001 ; Galand, Neuville & Frenay, 2005 ; Parmentier, 1994 ; Romainville, 2000) et susciter des initiatives visant à promouvoir la réussite.

À l'échelle de l'ensemble de la Belgique francophone, l'initiative la plus récente prise en ce domaine s'inscrit dans la mise en place des articles 78[1] et 84[2] du décret dit de « Bologne » (31 mars 2004). Leurs contenus sont les suivants : l'introduction d'une évaluation dispensatoire en BA 1 et le changement de seuil de réussite pour un cours au sein d'une année académique (et non d'une année à l'autre).

Lors de la première étape de cette recherche (Wolfs, J.L., Salcin, A. & Baillet, D., 2005), nous avions tenté de mesurer un effet éventuel de deux mesures. Pour cela, nous avions analysé le taux de réussite des étudiants de BA 1 à 3 niveaux (du macro au micro) :

- Celui de l'Université libre de Bruxelles,
- Celui des Facultés et Instituts,
- Celui des cours concernés par l'évaluation dispensatoire.

Globalement, nous n'avons pas pu démontrer un impact généralisé et évident des mesures citées ci-dessus. Ainsi :

---

[1] **Art. 78.** *« Un étudiant ne doit plus se présenter aux épreuves et examens d'un enseignement pour lequel il a obtenu une note au moins égale au seuil de réussite (à savoir 10/20) au cours de la même année académique. Au sein d'un programme d'études, un étudiant ne doit plus se présenter aux épreuves et examens d'un enseignement pour lequel il a obtenu une note d'au moins 12/20 au cours des cinq années académiques précédentes, quel que soit l'établissement en Communauté française où il s'inscrit par la suite. La note ainsi obtenue fait l'objet d'un report. »* (Ministère de la Communauté française de Belgique. 2004).

[2] **Art. 84.** *« (…) Pour les étudiants de première année d'études, les évaluations à l'issue du premier quadrimestre sont dispensatoires : elles peuvent faire l'objet d'une valorisation de tout ou partie de l'épreuve au sens de l'article 78, mais n'entrent pas en compte en cas d'échec. »* (Ministère de la Communauté française de Belgique. 2004).

- A l'échelle de l'Université, le taux global de réussite en Ba 1 en 2004-2005 n'était pas supérieur à celui des deux années académiques précédentes,
- A l'échelle des 9 Facultés dont les résultats ont été analysés plus en profondeur, nous avons retenu 4 Facultés dont les résultats avaient été stables au cours des 2 années précédentes. Parmi celles-ci, le taux de réussite de 3 Facultés est resté stable en 2004-2005 et celui de la dernière est en baisse,
- Enfin, au niveau des cours, nous avons constaté que de nombreuses circonstances empêchaient une comparaison valide au cours des trois dernières années (ex : changement de titulaires, introduction de changements particuliers au niveau du cours ou de la Faculté etc.). Sur 89 cours considérés au départ, 13 ont finalement été retenus, parmi ces 13, seuls 6 présentaient des résultats stables au cours des deux années précédentes et parmi ces derniers, on observe une différence significative en 2004-2005 par rapport à l'année précédente dans 3 cas : pour 2 cas, il s'agit d'une diminution et pour le troisième d'une augmentation.

Problématique

Dans cette communication, nous proposons d'explorer deux des difficultés que nous avions rencontrées lors de notre recherche précédente (octobre 2005). Il s'agit :

- D'une part, pour les 3 niveaux étudiés: de la complexité de dissocier l'effet des deux mesures (évaluation dispensatoire et changement de seuil),
- D'autre part, au niveau des cours concernés par l'évaluation dispensatoire : de la difficulté de comparer les taux de réussite en 2004-2005 à ceux des années précédentes et ce, en raison du changement de seuil. En effet, si nous prenions en compte les « 10 » et les « 11 » (considérées comme des notes de réussite en 2004-2005), nous constations une augmentation du taux de réussite qui, pourtant, ne se répercutait pas au niveau des résultats globaux des Facultés.

A cette fin, nous avons exploré les questions et hypothèses suivantes :

- Les enseignants ont-ils modifié leurs modalités d'évaluation ? A cette fin, nous vérifierons – pour les cours déjà étudiés - si les « 10 » et « 11 » sont proportionnellement plus nombreux en 2004-2005 que pour les autres années,
- Pourquoi une augmentation du taux de réussite des cours concernés par l'évaluation dispensatoire, ne se répercute-t-elle pas au niveau du taux de réussite des Facultés ? Pour ce faire, nous analyserons les résultats des cours pour lesquels il n'y avait pas d'évaluation dispensatoire (donc qui étaient évalués pour la première fois lors de la session de juin) et ce, afin de vérifier s'ils ont été moins bien réussis en 2004-2005 que les autres années,
- La session de juin a-t-elle été alourdie par les échecs de l'évaluation dispensatoire ? Pour explorer cette question, nous analyserons les données de l'année académique 2005-206 à partir desquelles nous tenterons de mesurer la proportion d'étudiants dont la session de juin a effectivement été plus importante.

1. Les enseignants ont-ils modifié leurs modalités d'évaluation ?

Lors de notre recherche précédente (Wolfs, J.L., Salcin, A. & Baillet, D., 2005), nous avions mis en évidence une augmentation importante du pourcentage de « 10 » et « 11 » en seconde session durant l'année 2004-2005 (voir ci-dessous : tableau 1).

Tableau 1 – Pourcentage de 10 et 11 année par année (reports compris)

Cours	**2001-2002**		**2002-2003**		**2003-2004**		**Moyenne 2001>2004**		**2004-2005**		**Différence 01>04/04-05**	
	1s	**2s**	**1s**	**2s**	**1s**	**2s**	**1s**	**2s**	**1s**	**2s**	**1s**	**2s**
C6	6,7	8,8	10,8	10,4	10,1	9,2	9,2	9,5	11,9	18,3	2,7	**8,8**
C10	11,9	1,6	1,7	0,6	8	2,3	7,2	1,5	10,6	11,1	3,4	**9,6**
C1	12,2	9,3	12,4	10,8	10,7	6,1	11,7	8,7	12,9	15,4	1,2	**6,7**
C4	10,9	4,4	6,7	4,6	8,7	3,4	8,8	4,1	12,6	17,8	3,8	**13,7**
C7	10,9	3,6	8,6	3,8	9,4	2,9	9,6	3,4	14,5	16,7	4,9	**13,3**
C2	8,7	14,8	11,8	9,7	8,7	11,6	9,7	12	17,5	25,4	7,8	**13,4**
C9	9,4	3,8	3,4	3,1	8,1	1,1	7	2,7	14,3	17,2	7,3	**14,5**
C5	23,7	16,2	24,8	18,5	25,4	14,3	24,6	16,3	17,4	19	-7,2	**2,7**
C8	13,9	9,9	14,7	5,9	10,9	7,2	13,2	7,7	15,5	30,1	2,3	**22,4**
C3	12,5	4,4	14,7	4,8	8,4	4,2	11,9	4,5	14,3	21,8	2,4	**17,3**
H3	11,6	9	10,2	13,3	13,4	10,5	11,7	10,9	20,2	31,1	8,5	**20,2**
H2	20,3	7,8	16,7	3,4	9,6	4,8	15,5	5,3	15,9	23,4	0,4	**18,1**
H1	12	10,9	14,2	8,6	13,7	4,8	13,3	8,1	15,2	17,9	1,9	**9,8**

L'analyse de ce tableau met en évidence de manière particulièrement spectaculaire l'augmentation de « 10 » et « 11 » en seconde session en 2004-2005. En effet, si l'on compare le pourcentage moyen de « 10 » et « 11 » pour les années précédant la mise en application du décret dit de « Bologne » avec le pourcentage de « 10 » et « 11 » au cours de cette dernière, on remarque des augmentations de ces notes allant de 2,7% à 22,7%.

Ce constat nous avait donc amené à poser la question suivante : les enseignants ont-ils modifié leurs modalités d'évaluation ? Et plus particulièrement, ont-ils modifié leur manière de noter les résultats des examens ?

Pour répondre à cette question, nous avons d'abord comparé les proportions de « 10 » et de « 11 » par rapport aux autres notes des cours étudiés lors de notre première analyse (Wolfs, J.L., Salcin, A. & Baillet, D., 2005). Afin que les résultats des différentes années étudiées soient comparables, nous avons choisi de respecter les deux conditions suivantes :

- Critère de stabilité : vérification de la stabilité des proportions de « 10 » et « 11 » avant l'année de la réforme (2004-2005),
- Soustraction des notes reportées de la première à la seconde session pour toutes les années étudiées[3].

Ces deux étapes nous paraissaient indispensables afin de pouvoir donner du sens aux différences significatives hypothétiques avec les résultats de l'année 2004-2005.

[3] Cette soustraction explique que les pourcentages de « 10 » et « 11 » en seconde session soient différents dans les tableaux 1 et 2.

Tableau 2– Proportion de 10 et 11 année par année (sans les reports en seconde session)

Cours	2001-2002		2002-2003		2003-2004		Moyenne 2001>2004		2004-2005		2001>2004		2001>2005			
	1s	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s	2s	1s		2s	
	%	%	%	%	%	%	%	%	%	%	Chi	Chi	%	Chi	%	Chi
C9	12,5	11,6	14,7	8,6	8,4	13,9	11,8	11,4	14,3	5,3	NS	NS	2,5	NS	-6,1	NS
C1	12,2	10	12,4	11,9	10,7	6,7	11,7	9,5	12,9	3,6	NS	NS	1,2	NS	-5,9	S.01
C8	13,9	14,5	14,7	10,5	10,9	11,7	13,1	12,2	15,5	21,3	NS	NS	2,4	NS	9,1	S.05
H3	11,6	13,9	10,2	19,5	13,4	12,4	11,7	15,3	20,2	11,5	NS	NS	8,5	S.001	-3,8	NS
C4	10,9	10,6	6,7	11,4	8,7	11	8,7	11	12,6	9,2	NS	NS	3,9	S.05	-1,8	NS
C5	23,7	13	24,8	29,4	25,4	29,3	24,6	23,9	17,4	5,7	NS	NS	-7,2	S.05	**-18,2**	S.001
H1	12	18,3	14,2	13,8	13,7	8,1	13,3	13,4	15,2	15,7	NS	S.05	1,9	NS	2,3	NS
C7	10,9	6	8,6	19	9,4	12,2	9,6	12,4	14,5	14,3	NS	S.05	4,9	S.05	1,9	NS
C2	8,7	20,8	11,8	11,7	8,7	14,2	9,7	15,6	17,5	12	NS	S.05	7,8	S.001	3,6	S.05
C10	11,9	6,6	1,7	6,9	8	12,5	7,2	8,6	10,6	8,2	S.001	NS	3,4	S.001	-0,4	NS
H2	20,3	14	16,7	6,8	9,6	13,9	15,5	11,6	15,9	18,8	S.001	NS	0,4	S.001	7,2	NS
C9	9,4	14,1	3,4	14,1	8,1	3,3	6,9	10,5	14,3	9,6	S.001	NS	7,4	S.001	-0,9	S.05
C6	6,7	14,3	10,8	20,8	10,1	13,6	9,2	16,2	11,9	9,8	S.05	NS	2,7	S.01	-6,4	S.05

L'analyse de ce tableau nous permet de mettre en évidence que sur les 13 cours étudiés :

- 6 respectent le critère de stabilité. Parmi ceux-ci, nous constatons que :
  - Pour 1 cours, la proportion de « 10 » et « 11 » reste stable en 2004-2005 quelque soit la session,
  - Pour 2 cours, la proportion de « 10 » et « 11 » reste stable en 2004-2005 uniquement en première session,
  - Pour 2 cours, la proportion de « 10 » et « 11 » est significativement différente uniquement en 1ère session,
  - Enfin, pour le dernier cours, la proportion de « 10 » et « 11 » est significativement différente quelque soit la session.
- 3 respectent le critère de stabilité uniquement en première session. Parmi ceux-ci, nous constatons que :
  - Pour 1 cours, la proportion de « 10 » et « 11 » reste stable en 2004-2005 quelque soit la session,
  - Pour un autre cours, la proportion de « 10 » et « 11 » reste stable en 2004-2005 uniquement en seconde session,
  - Enfin, pour le troisième cours, la proportion de « 10 » et « 11 » est significativement différente quelque soit la session.
- 4 cours respectent le critère de stabilité uniquement en seconde session. Parmi ceux-ci, nous constatons que :
  - Pour 2 cours, la proportion de « 10 » et « 11 » reste stable en 2004-2005 uniquement en seconde session,

- Enfin, pour les 2 autres cours, la proportion de « 10 » et « 11 » est significativement différente quelque soit la session.

En conclusion, il apparaît que les résultats de trois cours (C4, C5 et H3) sur les 13 étudiés restent stables et montrent une différence significative avec les résultats de 2004-2005 pour la première session. Un seul de ceux-ci montre également une différence significative en seconde session. En conséquence de ceci, il semblerait que trois enseignants aient modifié leur manière d'attribuer des notes de « 10 » et « 11 » aux examens en première session : 2 à la hausse et 1 à la baisse. Alors que la proportion de « 10 » et « 11 » pour les cours à la hausse diminue en seconde session. Cet équilibre est-il volontaire ?

Cette faible proportion (3/13 cours) nous pousse à penser que les changements observés dans les pourcentages de « 10 » et « 11 » ne résultent pas forcément d'une volonté des enseignants de modifier leurs pratiques évaluatives suite à la mise en place du décret dit de « Bologne ». En effet, nous pensons que le changement du seuil de report (passage de 12/20 à 10/20) au cours d'une année académique combiné à une approche stratégique des examens par les étudiants pourraient suffire à expliquer l'augmentation du pourcentage de « 10 » et « 11 » en seconde session.

Pourquoi une augmentation du taux de réussite des cours (en tenant compte des « 10 » et des « 11 ») ne se répercute-t-elle pas au niveau du taux de réussite global des Facultés ?

Pour rappel, la première partie de cette recherche (Wolfs, J.L., Salcin, A. & Baillet, D., 2005) nous avait permis de mettre en évidence un paradoxe : alors que le taux de réussite des cours concernés par l'évaluation dispensatoire était significativement plus élevé en 2004-2005 (en tenant compte des « 10 » et « 11 ») que les années précédentes, les taux de réussite facultaires chutaient ou, au mieux, restaient stables. Ce paradoxe nous avait donc mené sur la piste de recherche suivante : les cours évalués pour la première fois en juin auraient-ils été moins bien réussis en 2004-2005 que les années précédentes ?

Pour procéder à l'analyse des résultats de ces cours, nous avons retenu les cours dont les conditions d'enseignement étaient restées stables dans le temps. Plus précisément, nous avons ciblé les cours ayant conservé le même intitulé et le même titulaire depuis 2001-2002. Pour cela nous avons effectué un premier tri en consultant les programmes de cours des quatre dernières années. Ensuite, nous avons contacté chaque titulaire concerné afin de confirmer cette stabilité mais aussi de vérifier si selon eux, cette comparaison était possible.

Après avoir écarté les cours ayant subi des modifications, nous avons effectué pour les deux cours correspondant à nos critères et dont les données nous avaient été transmises, les manipulations suivantes:

- Calcul des fréquences et pourcentages de réussite en y incluant ou pas les « 10 » et « 11 »,
- Vérification de la stabilité des taux de réussite d'une année à l'autre.

Tableau 3 – Cours H4

	2001-2002	2002-2003	2003-2004	Moyenne 2001>2004	2004-2005	2001>2004 / 04-05
	%	%	%	%	%	Chi
R 12+	***60,24***	59,42	***54,79***	58,15	***50,23***	**S.01**
R 10, 11 et 12+	63,78	65,38	61,91	63,69	63,64	**N.S.**

En ce qui concerne le cours H4, si l'on ne tient pas compte des « 10 » et des « 11 », il apparaît que le taux de réussite est significativement inférieur en 2004-2005 par rapport aux années précédentes. Notons que cette différence s'explique surtout par une chute de 10% du taux de réussite par rapport à l'année 2001-2002 mais qui n'est plus que de 5% si l'on considère les résultats de l'année 2003-2004.

Par contre, en incluant les « 10 » et « 11 », le taux de réussite en 2004-2005 est comparable à celui des années antérieures. Il semblerait donc que le changement du seuil de réussite (au sein d'une année académique) ait permis de maintenir le taux de réussite en équilibre pour ce cours particulier durant l'année 2004-2005.

Tableau 4 – Cours H5

	2001-2002	2002-2003	2003-2004	Moyenne 2001>2004	2004-2005	2001>2004 / 04-05
	%	%	%	%	%	Chi
R 12+	37,01	42,31	43,79	41,03	40,23	**N.S.**
R 10, 11 et 12+	39,57	46,73	47,66	44,65	47,73	**N.S.**

En ce qui concerne le cours H5, que l'on tienne compte ou pas des « 10 » et « 11 », les différences observées au niveau du taux de réussite entre l'année 2004-2005 et les deux années académiques précédentes ne sont pas significatives. Par contre, nous relevons que le taux de réussite enregistré en 2001-2002 est significativement différent de celui des autres années sans pour autant savoir l'expliquer.

Sur base de ces chiffres, il semblerait donc que le changement du seuil de réussite introduit par la mise en place du décret dit « de Bologne » n'ait pas eu d'impact sur le taux de réussite de ce cours.

En résumé, les taux de réussite en 2004-2005 sont :

- En comptabilisant les « 10 » et « 11 » : en équilibre pour les deux cours étudiés,
- Sans comptabiliser les « 10 » et « 11 » : significativement en baisse pour le cours H4 mais pas pour le cours H5.

Pour conclure, il semblerait donc que notre hypothèse (les cours évalués pour la première fois en juin ont été moins bien réussis en 2004-2005 que les années

précédentes) ne soit pas confirmée par l'analyse des taux de réussite des cours H4 et H5.

Voici quelques explications possibles :

- Les critères de sélection des cours (critères de stabilité : même titulaire, même intitulé, même volume d'enseignement au cours des 4 dernières années académiques) pourraient nous avoir conduits à exclure des cours dont le taux de réussite aurait chuté ou augmenté,
- Les réussites partielles (Art.79[4]) ne sont pas comptabilisées dans le taux de réussite global de la Faculté H. Or, les résultats des étudiants ayant bénéficié de cette mesure sont inclus dans les données que nous avons analysées pour chaque cours. Il est donc possible que les seuls résultats de ces étudiants (incluant un grand nombre de « 10 » et de « 11 ») expliquent une augmentation du taux de réussite au niveau des cours. Augmentation qui ne se répercuterait pas au niveau de la Faculté H puisque les réussites partielles ne sont pas comptabilisées dans le pourcentage global de réussite facultaire.

2. *La session de juin a-t-elle été alourdie par les échecs de l'évaluation dispensatoire ?*

Pour répondre à cette dernière question, nous avons comptabilisé le nombre d'examens supplémentaires que les étudiants ont repassé en juin et qui n'auraient pas eu la possibilité de représenter les années antérieures à la réforme. Les résultats de janvier 2005 ayant été « écrasés » par ceux de juin, nous avons dû analyser les résultats relatifs à l'année académique 2005-2006. On peut donc supposer que certains étudiants ont établi de meilleures stratégies que l'année de la réforme. Toujours concernant les données exploitées, nous avons utilisées celles de la Faculté H qui de par le temps imparti nous étaient le plus facilement accessibles. Nous avons constaté que sur les 485 étudiants inscrits en BA 1 en 2005-2006 à la Faculté H se répartissent de la manière suivante :

Tableau 5 – Distribution des examens représentés

Fréquences Pourcentages		**Nombre d'examens à représenter**							**Total**
		0	1	2	3	4	5	6	
Nombre d'examens effectivement représentés	0	57 11,7%	*18* *3,7%*	*7* *1,4%*	*8* *1,6%*	*9* *1,8%*	*16* *3,3%*	*46* *9,5%*	161 33%
	1		45 9,2%	24 4,9%	13 2,7%	13 2,7%	14 2,8%	12 2,5%	121 24,8%
	2			37 7,6%	20 4,1%	16 3,3%	10 2%	7 1,4%	90 18,4%
	3				33 6,8%	21 4,3%	5 1%	5 1%	64 13,1%

[4] **Art. 79.** *« Un jury peut prononcer la réussite d'une année d'études dès que l'étudiant y a acquis plus de 48 crédits. Dans ce cas, le solde des crédits doit être intégralement obtenu au cours de l'année d'études suivante. »* (Ministère de la Communauté française de Belgique. 2004)

4					22 4,5%	10 2%	1 0,2%	33 6,7%
5						8 1,6%	4 0,8%	12 2,4%
6							4 0,8%	4 0,8%
Total	57 11,7%	63 12,9%	68 13,9%	74 15,2%	81 16,6%	63 12,7%	79 16,2%	**485** **99,2%**

A la lecture de ces tableaux, nous constatons que 57 étudiants sur les 485 inscrits en BA 1 de la Faculté H, soit 11,7% ont réussi les 6 examens de janvier pour lesquels une évaluation supplémentaire était organisée en juin.

Nous constatons également que 104 étudiants (en *italique* dans le tableau), soit 21,4% ont choisi de ne pas du tout exploiter la possibilité de représenter les examens échoués en janvier. Par contre, 324 étudiants, soit 66,8% ont décidé de représenter au moins un examen échoué en janvier. Parmi ceux-ci, 149 étudiants soit 30,7% ont choisi de représenter l'ensemble des examens qu'ils ont échoués en janvier. Ces derniers se distribuent de la manière suivante :

- 45 étudiants soit 9,2% n'ont dû en représenter qu'un seul,
- 37 étudiants, soit 7,6% en ont représenté deux,
- 33 étudiants, soit 6,8% en ont représenté trois,
- 22 étudiants, soit 4,5% en ont représenté quatre,
- 8 étudiants, soit 1,6% en ont représenté cinq,
- 4 étudiants, soit 0,8% ont représenté les 6 examens.

Alors que 175 étudiants, soit 36% ont choisi de ne représenter qu'une partie des examens qu'ils ont échoués en janvier comme le montre le graphique suivant :

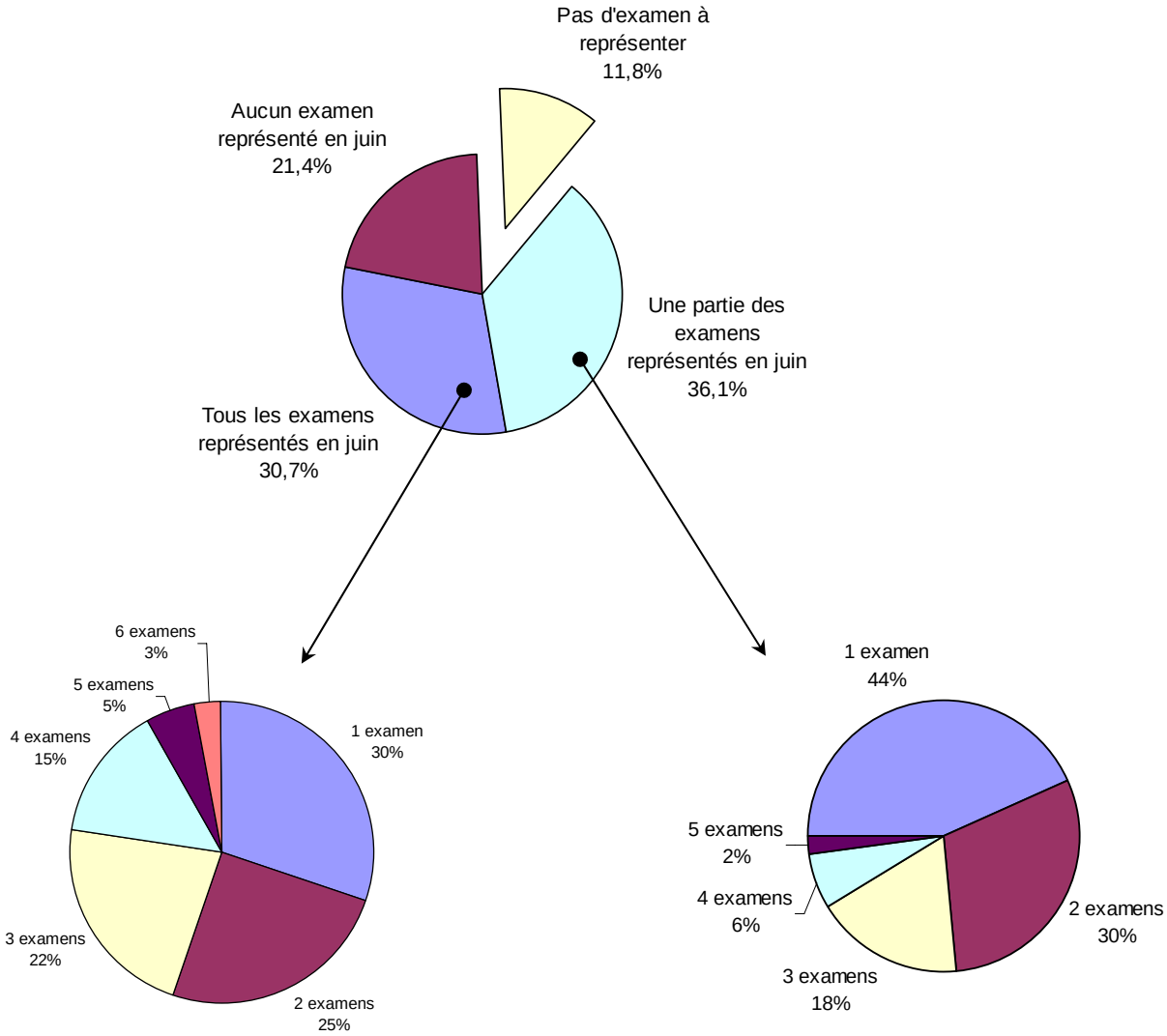


Figure 1 - Répartition des examens représentés en juin.

En conclusion, 66,8% des étudiants ont effectivement alourdi leur session de juin. La moitié d'entre eux (30,7%) a repassé l'ensemble des examens échoués en juin et l'autre moitié a effectué des choix stratégiques en ne représentant qu'une partie des examens échoués en janvier.

Conclusion et pistes de recherches

Suite à notre première recherche sur l'introduction d'une évaluation dispensatoire, nous avions constaté une augmentation du taux de réussite si on prenait en compte les « 10 » et les « 11 ». Celle-ci ne semblait pourtant pas influencer le taux de réussite global par Faculté.

Pour comprendre ce paradoxe, nous avons d'abord tenté de vérifier si les enseignants (dont nous avions étudié les cours lors de notre précédente recherche) ont attribué plus facilement les notes de « 10 » et de « 11 ». Nous avons pu démontrer des modifications dans l'attribution de ces notes pour 3 cours sur les 13 étudiés. Nous en concluons que l'augmentation de ces notes résulte principalement du fait qu'elles peuvent, depuis 2004-2005, être reportées en seconde session, ce qui n'était pas le cas les années précédentes.

Pour analyser le fait qu'une augmentation du taux de réussite des cours concernés par l'évaluation dispensatoire (en tenant compte des « 10 » et des « 11 ») ne se répercute pas sur le taux de réussite global des Facultés étudiées, nous avons émis l'hypothèse que les résultats des cours non concernés par l'évaluation dispensatoire (donc évalués pour la première fois en juin) avaient été moins bien réussis de par un alourdissement potentiel de la session de juin. Pour cela, nous avons analysé les résultats de deux cours de la Faculté H qui respectaient nos critères de stabilité (même enseignant, même intitulé de cours et approximativement même quantité d'heures pour les trois années précédant la réforme). Nous n'avons pas pu mettre en évidence une baisse du taux de réussite pour ces deux cours. Toutefois, nos critères de sélection sont très restrictifs et ne nous permettent pas d'analyser d'autres cours qui pourraient avoir été moins bien réussis.

Par ailleurs, nous posons également l'hypothèse que cette « non répercussion » au niveau facultaire peut également s'expliquer par les réussites partielles qui ne sont pas comptabilisées dans les réussites au niveau central.

Enfin, nous avons également tenté d'analyser dans quelle mesure la session de juin a été alourdie. Ce qui pourrait éventuellement expliquer l'absence d'impact de la réforme en ce qui concerne la réussite globale. Pour cela, nous n'avons pu utiliser que les données de l'année 2005-2006 d'une seule Faculté comptant (cette année académique) 485 étudiants en BA 1. Nous avons constaté que seulement 11,7% des étudiants ont réussi l'ensemble des examens de janvier, donc 88,3% voyaient leur session de juin potentiellement alourdie. Toutefois 21,4% d'entre eux ont choisi de ne représenter aucun de ces examens en juin. La session a donc été effectivement alourdie pour 66,8% mais seulement la moitié (30,7%) a repassé l'entièreté des examens échoués alors que l'autre moitié (36,1%) semble avoir réalisé des choix stratégiques en ne représentant qu'une partie des examens échoués en janvier.

En conclusion, à ce stade-ci de nos recherches, nous ne pouvons pas montrer un effet des deux mesures sur le taux de réussite. Par conséquent, nos analyses tendent à mettre évidence un hiatus entre l'intention de départ du décret visant à promouvoir la réussite et les effets réels de ces mesures qui sont, comme d'autres (guidances, tutorat, etc.) difficiles à mesurer. Ceci pourrait illustrer l'écart entre des initiatives qui

semblent à priori favorables à la réussite mais qui selon les modalités d'application (parfois très diverses selon l'interprétation des acteurs) peuvent induire des comportements paradoxaux par rapport aux intentions initiales. Comme par exemple un degré d'exigence préalablement noté par un 12 et actuellement noté par un 10. En bref, il est difficile d'évaluer l'effet de ce type de mesure car lorsqu'on cherche à contrôler toute une série de conditions de départ, le nombre de cours restants sur lequel on peut se prononcer devient très réduit, même si un résultat ressort très nettement : l'alourdissement de la session de juin.

Toutefois la plupart de nos analyses se basant sur des résultats de l'année même de l'introduction de la réforme, il serait pertinent d'observer l'évolution de ces résultats ces trois, cinq et dix prochaines années.

Une autre piste possible pour enrichir notre étude réside plus précisément dans l'analyse des pratiques enseignantes. En effet, l'introduction d'une évaluation supplémentaire augmentant le travail d'élaboration et surtout de correction des enseignants, il pourrait être opportun de vérifier si cette dernière n'a pas suscité de changements dans le type d'évaluation proposé, par exemple : la transformation d'une évaluation à questions ouvertes à une évaluation à questions fermées. Transformation qui pourrait impliquer la mesure de compétences différentes, et par conséquent, pourrait également influencer les stratégies et les résultats des étudiants.

Références :

Droesbeke, J.J., Hecquet, I. & Wattelaer, C. (2001). *La population étudiante: description, évolution, perspectives.* Bruxelles : Ed. de l'Université de Bruxelles & Ed. Ellipses.

Galand, B., Neuville & Frenay. (2005). L'échec à l'université en Communauté française de Belgique. *Les Cahiers de Recherche en Education et Formation,* 39.

Leclerq, D. (2002), *Rapport du projet « MOHICAN »*. Communauté française de Belgique : Ministère de l'Enseignement supérieur.

Ministère de la Communauté française de Belgique. (2004). Décret définissant l'enseignement supérieur, favorisant son intégration à l'espace européen de l'enseignement supérieur et refinançant les universités. *Moniteur belge, 18 juin 2004,* 45239 à 45275.

Parmentier, P. (1994). *La réussite des études universitaires : Facteurs structurels et processuels de la performance académique en première année en médecine.* Thèse de doctorat non publiée, Faculté de psychologie et des sciences de l'éducation, Louvain-la-Neuve : Université Catholique de Louvain.

Romainville, M. (2000). *L'échec dans l'université de masse*. Paris : L'Harmattan.

Wolfs, J.L., Salcin, A. & Baillet, D. (2005). L'évaluation d'un dispositif de promotion de la réussite en première année d'études universitaires en Communauté française de Belgique: Questions et Paradoxes. In Actes du 18e Colloque international de l'Association pour le Développement des Méthodologies d'Evaluation en Education. *Comment évaluer ? Outils, dispositifs et acteurs.* Reims : Colloque IUFM/ADMEE 2005.

DISPOSITIF DE FORMATION-EVALUATION SOCIOCONSTRUCTIVISTE CENTRE SUR LES COMPETENCES, EN FORMATION INITIALE D'INFIRMIER-IERES

Salomé Anderfuhren Fiaux, Brigitte Kampel, Haute Ecole Cantonale Vaudoise de la Santé, HecvSanté, Lausanne

Mots-clés : Pédagogie professionnelle socioconstructiviste, cohérence curriculum-évaluation, changement institutionnel

Résumé :

Depuis 2002, la formation initiale d'infirmier-ère se situe en Suisse Romande au niveau tertiaire HES. Un référentiel de compétences commun aux divers sites de formation a été créé, intégrant explicitement agir réflexif, recherche, encadrement et promotion du rôle professionnel. Pour concrétiser ces dispositions, la filière infirmier-ère de la HECVSanté s'est donné un nouveau cadre de référence institutionnel, basé sur une vision socioconstructiviste de la formation, puis traduit dans un nouveau curriculum dont les éléments-clés sont :

- Structuration par l'apprentissage par problèmes (PBL)
- Intégration des connaissances dans les compétences prioritaire sur le morcellement des apprentissages
- Les situations d'apprentissages visent dès le début de la formation à relever des défis professionnels ; elles contraignent les étudiants à développer leur réflexivité
- Travail des professeurs en équipe coopérative semestrielle
- L'évaluation est essentiellement de nature formative, orientée par et centrée sur les compétences, intégrée dans des situations professionnelles complexes nécessitant la mobilisation de connaissances de nature diverses pour être problématisées et résolues
- Limitation des outils d'évaluation à cinq, tout au long de la formation, créant une forte cohérence entre évaluation formative et sommative

La construction du curriculum a été possible grâce à l'engagement professoral, permettant l'essor de cultures pédagogique, institutionnelle et individuelle nouvelles, transformant l'organisation du travail, la didactique et l'évaluation. Pour les étudiants aussi, la formation implique une rupture se caractérisant par un grand engagement et beaucoup d'adaptation. Les résultats préliminaires du développement des compétences professionnelles sont bons ; les nouveaux professionnels ont acquis une posture réflexive, nécessaire pour agir dans les situations professionnelles complexes. Les directives de Bologne nous contraignent à adapter le programme; nous prévoyons de renforcer l'intégration des connaissances pour en favoriser la stabilité ainsi que certaines thématiques, et de développer des enseignements communs aux différentes filières présentes dans notre institution

---

Présentation

Les mutations socio-sanitaires suisses ont conduit la formation d'infirmier-ère, en Suisse Romande à s'effectuer en 4 ans dans une HES (Haute Ecole Spécialisée). Elle réunit 10 professions sanitaires et sociales qui ont défini des référentiels de profession de niveau « réflexif ». Ces changements paradigmatiques majeurs ont été l'occasion pour notre profession et notre site de formation de (se) re-penser.

La filière infirmier-ère de la HECVSanté (Haute Ecole Cantonale Vaudoise de la Santé) a alors choisi de se donner un nouveau cadre de référence. Accompagnée par Jaques Tardif (Université de Sherbrooke), une réflexion du corps enseignant déboucha sur la rédaction et l'adoption d'une vision socioconstructiviste, opérationnalisée par un curriculum entièrement structuré par l'apprentissage par problèmes (PBL). Le changement, tant aux plans institutionnel qu'individuel, était donc conceptuel, organisationnel, didactique, instrumental et identitaire. Nous souhaitons vous entretenir de l'articulation entre dispositifs didactique et d'évaluation.

Orientations et choix pédagogiques

Les éléments clés sont :

- vision socioconstructiviste de l'apprentissage, enrichie de psychologie cognitive,
- formation de niveau tertiaire, menant à un agir professionnel réflexif,
- référentiel de formation basé sur le référentiel de compétences métier
- alternance intégrative, (1/3 de formation pratique)
- enseignements par des professeurs, tous infirmiers-ères et formés en pédagogie,

Sur ces bases, nous avons construit le curriculum en veillant constamment que

- L'intégration des connaissances dans les compétences ait la priorité sur leur morcellement,
- Les situations d'apprentissages visent à relever des défis professionnels et contraignent les étudiants à développer leur réflexivité,
- L'organisation du cursus soit faite pour optimaliser l'apprentissage des étudiants et non pour faciliter le planning des professeurs.

Nous avons donc choisi l'Apprentissage Par Problèmes comme structure curriculaire, aboli les enseignements disciplinaires et constitué des équipes semestrielles coopératives.

Chacune a eu pour mandat de construire un semestre. Elle rédige les situations didactiques emblématiques de la pratique professionnelle, conçoit le dispositif didactique requis et sélectionne les ressources impliquées (connaissances – habiletés – attitudes). L'équipe est également chargée de penser l'intégration du stage professionnel semestriel, composant du développement des compétences (préparation du stage, consignes de travail pour le stage, ateliers réflexifs au retour) ; et assure l'évaluation formative et sommative.

Le dispositif d'évaluation

L'évaluation des compétences et des savoirs est indispensable à la formation, mais n'en constitue pas le moteur principal, même si elle en est un levier (Paquay, 2000).

Toutefois, lorsqu'elle y est intégrée, elle soutient la motivation, la régulation et l'autonomie, tant personnelle que professionnelle. Notre dispositif de formation-évaluation valorise ce postulat. Deux volets fondamentaux interdépendants la constituent : la régulation (formative) et la certification (sommative).

En cours d'année, toutes les activités des étudiants (APP, ateliers cliniques, travaux obligatoires), ont une visée formative. Durant le semestre, les étudiants reçoivent des évaluations in-formatives descriptives précises sur le degré de professionnalisation développé. Puis, lors d'une rencontre individuelle avec un formateur, l'étudiant présente son auto-évaluation et la confronte. Le degré de développement des compétences atteint est ensuite synthétisé dans un bulletin des apprentissages, lequel comporte un projet de développement, obligatoire si le niveau démontré est insuffisant. L'accent se porte donc sur la régulation.

Pour assurer la cohérence entre formation et pratique, nous avons limité les instruments formels d'évaluation à 5, réputés pertinents pour l'agir professionnel (Analyse et projet de soins, Carte conceptuelle ou réseau conceptuel, Base d'orientation ou carte d'étude, Récit et analyse de soin, Analyse de traces issues du portfolio)[1]. Les étudiants en réalisent 3 différents chaque année. Ils sont utilisés dans des productions faites à l'école, -mobilisation de connaissances mémorisées-, et dans des productions issues d'expérience réalisées lors des stages. Les indicateurs de compétences annuels progressifs fournissent les critères d'évaluation.

En fin d'année académique, deux épreuves sommatives, centrées sur les compétences « centre de gravité » de l'année, permettent l'acquisition des crédits ECTS. La 1ère, un « examen », mobilise une sélection des connaissances apprises pour problématiser et résoudre les problèmes posés par une situation professionnelle complexe. La 2ème est une analyse réflexive approfondie d'une expérience de stage, relatant les conditions et les exigences des actions réalisées, les ressources mobilisées pour agir, les apprentissages généralisables et les possibilités de transfert.

Ainsi, l'évaluation est systématiquement centrée sur les compétences : elle porte toujours sur des connaissances contextualisées, privilégie les situations mal définies,

[1] NB Les outils utilisés font l'objet d'une autre présentation par les mêmes auteures

met l'accent sur les changements autant que sur l'état final, et exige le recours à des critères de correction multiples. La cohérence entre les volets formatifs et sommatifs est un point fort et efficace du dispositif.

Conclusion : points forts et limites du dispositif

Construire et réaliser ce programme a été un processus de plusieurs années, nécessitant le développement d'une nouvelle culture pédagogique institutionnelle et individuelle.

Si nos pratiques antérieures étaient déjà portées par le paradigme socioconstructiviste, la centration permanente sur les apprentissages, leur intégration constante au développement des compétences plutôt que sur l'enseignement, et ce aussi bien dans le dispositif de formation que celui d'évaluation, ont représenté un défi pour les professeurs et pour l'organisation du travail.

En effet, chaque professeur doit avoir une connaissance approfondie du cadre socioconstructiviste, des compétences visées, de l'étalement des compétences et des indicateurs de développement sur 4 ans. De plus, travailler en équipe implique que chaque collègue soit conscient de son apport spécifique, de sa complémentarité avec les autres membres, ainsi que de l'ensemble du curriculum. Cela nécessite de nombreuses rencontres pour construire des pratiques communes et s'approprier des éléments aussi divers que la gestion d'un groupe d'apprentissage par problèmes, l'évaluation des compétences, de nouveaux outils d'apprentissage et d'évaluation. Cela a demandé un engagement fort de chaque professeur, car ce travail était bien sûr réalisé en parallèle avec la poursuite de l'ancien programme de formation. Les difficultés principales que nous avons rencontrées concernent l'évaluation : utiliser des outils communs, pratiquer en permanence l'évaluation formative, évaluer les compétences et non seulement les connaissances, déléguer l'évaluation sommative.

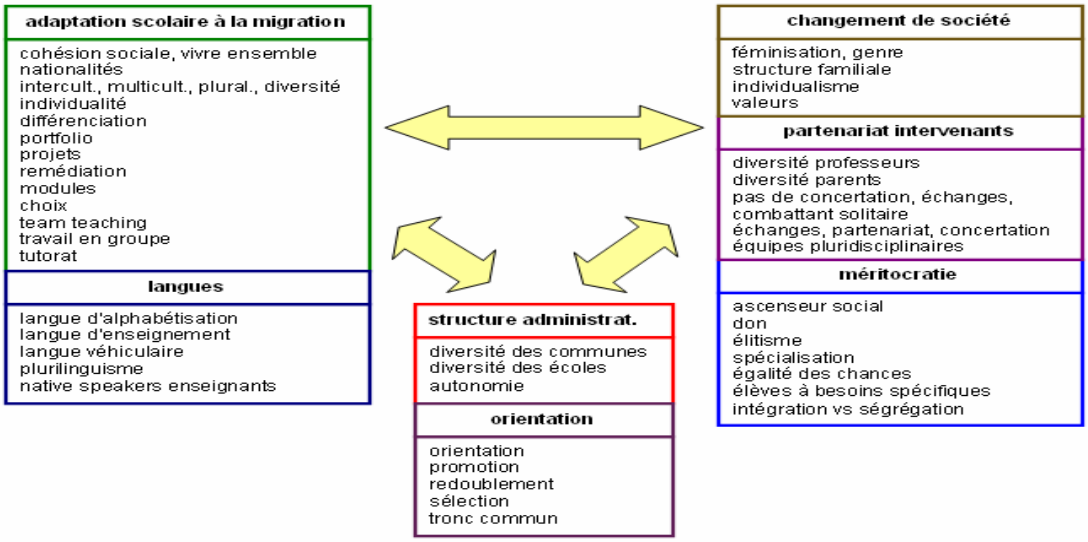
Les étudiants vivent aussi une rupture importante en entrant dans notre programme. Ils doivent à la fois comprendre les enjeux d'une formation professionnalisante, le développement des compétences et ses exigences cognitives et sociales. Ils ont aussi à développer des stratégies d'autogestion des apprentissages, donc avoir un haut degré d'autonomie comme apprenant. La première année de formation est un moment d'adaptation important, mais rares sont ceux qui quittent l'école en raison des

orientations pédagogiques. Le principal atout en est la construction du sens des apprentissages : nos étudiants se sentent soignants dès le premier jour de formation. Nous constatons aujourd'hui que nos choix pédagogiques portent leurs fruits. Les praticiens et les formateurs observent des étudiants en stage proactifs en analyse et résolution de problèmes, très responsables, affirmés et assertifs. A l'école également, les étudiants sont autonomes, capables de faire des autoévaluations pertinentes ; ils plébiscitent l'APP. Toutefois, nous constatons un manque de stabilité des connaissances chez nos étudiants, et un manque de temps d'intégration dans la formation. Des changements sont déjà en cours, afin d'adapter le programme aux directives de Bologne. Cet automne, une première volée débutera un cursus aboutissant à un Bachelor in Science in Nursing.

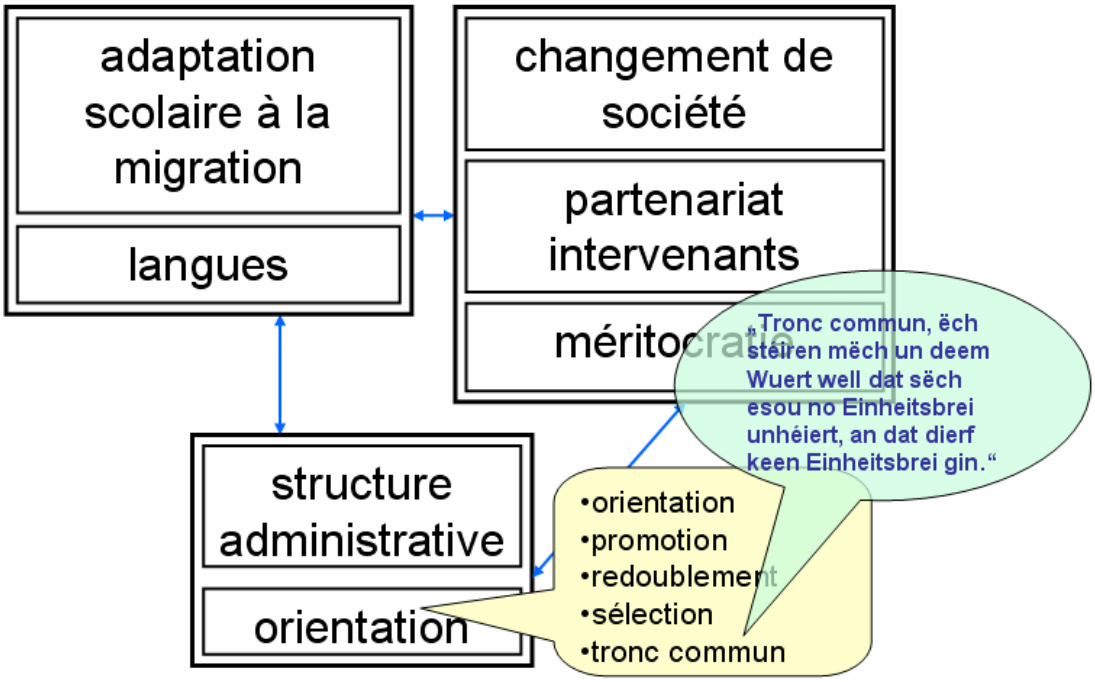
LES REPRÉSENTATIONS SOCIALES DES 36 INTERVENANTS DIRECTS ET INDIRECTS DU SYSTÈME SCOLAIRE SUR LA GESTION DE L'HÉTÉROGÉNÉITÉ DANS LE SYSTÈME SCOLAIRE LUXEMBOURGEOIS: ANALYSE DU CONTENU DE DONNÉES ISSUES D'UN DISPOSITIF DE FOCUS GROUP.

Christian Meyers, Georges Barthel, Danielle Zerbato, Romain Martin,
Université du Luxembourg

La présente communication se voit dans la ligne directe de la communication « Les représentations sociales de 56 intervenants directs et indirects du système scolaire sur la place de l'école dans la société luxembourgeoise de demain » et se fonde aussi sur des travaux de recherche effectués par des membres du groupe EMACS dans le cadre du projet de recherche « La place de l'école dans la société luxembourgeoise de demain » s'inscrivant dans le Programme pluriannuel de recherche « VIVRE » du Fonds National de Recherche. Elle peut puiser dans le même recueil de données qui a été constitué à partir d'un dispositif de focus groups ayant confronté 36 intervenants directs et indirects dans un premier tour aux questions sur les finalités de l'Ecole, au fonctionnement actuel du système scolaire luxembourgeois ainsi qu'aux réformes et changements potentiels de celui-ci et dans un deuxième temps par la transcription des discussions autour de douze propositions ou scénarios de développement potentiels du système scolaire luxembourgeois subdivisé en trois grands axes, à savoir : Axe 1 : Vers une gestion de l'hétérogénéité des élèves ; Axe 2 : Vers une redéfinition du rôle et de la tâche de l'enseignant ; Axe 3 : Vers un dispositif d'assurance-qualité de l'école. Mais le but de communication est de présenter au public les résultats de recherche quant à la gestion de l'hétérogénéité dans le système scolaire luxembourgeois, donc plus précisément les résultats de l'analyse de contenu des propos tenus par les intervenants par rapport à l'Axe 1. Pour ce faire, nous avons dégagé, à partir d'une catégorisation du corpus de données, les thèmes principaux reliés au concept d'hétérogénéité. Voici donc le tableau synthétique :



Nous avons dégagé sept grandes catégories - avec leurs sous-catégories respectives - qui ont été regroupées en trois grands domaines qui sont tous les trois en interaction. Dans chaque sous-catégorie, nous retrouvons les citations des intervenants, citations qui ont donc été rassemblées par thème. Le tableau ci-dessus sert à l'expliciter sous forme graphique :



Après avoir brièvement expliqué notre façon de procéder, venons-en maintenant à la présentation et à l'interprétation des données organisées comme nous l'avons montré ci-dessus. Pour chacune de ces sept catégories, nous présenterons une brève synthèse et montrerons les liens entre les trois domaines. En ce sens, il nous semble très important d'approfondir plus particulièrement les liens entre la gestion de l'hétérogénéité et l'orientation car les sous-catégories de redoublement (respectivement d'échec scolaire) et de tronc commun sont parmi les plus importants éléments du débat actuel sur les réformes du système scolaire luxembourgeois. En entrant ici brièvement dans les considérations des intervenants, nous pouvons préciser que, pour ceux-ci, il faut d'abord s'entendre sur la définition du tronc commun qui, pour des raisons historiques, est connoté négativement et implique, aux yeux du public, un nivellement vers le bas et une uniformisation (« une bouillie uniforme »). Tandis que les adversaires citent des sources britanniques qui considèrent le tronc commun comme compromettant les chances des élèves doués, les partisans du tronc commun y voient par contre une chance pour favoriser la différenciation et le

développement individuel des élèves, à condition de viser dès l'école primaire un socle de compétences intégrant des cycles d'apprentissage. Certains intervenants arrivent finalement à la conclusion que tronc commun, abolition du redoublement, lutte contre l'échec scolaire, sont étroitement liés et contribuent à une gestion optimale de l'hétérogénéité, tout en constituant un avantage financier important. Puis, des citations percutantes illustreront de manière vivante ces propos et bien d'autres que les intervenants ont tenu lors des focus groups. Ceci nous permettra finalement de dégager, selon les représentations sociales des intervenants, les principaux problèmes par rapport à la gestion de l'hétérogénéité auxquels le système scolaire luxembourgeois est confronté et de proposer des mesures concrètes afin de remédier à ces carences.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

GENERALISABILITE ET COMPETENCE SOCIALE : UNE ETUDE CHEZ DES ELEVES QUEBECOIS

Marie-Hélène Hébert, Université Laval

Martin Gendron, Université du Québec à Rimouski (Campus de Lévis)

Éric Frenette, Université du Québec à Rimouski

Égide Royer, Université Laval

Céline Morand, Centre de santé de Portneuf

Mots-clés : généralisabilité, évaluation des compétences, compétence sociale.

Résumé :

Les troubles du comportement chez les jeunes d'âge scolaire constituent une préoccupation de plus en plus importante pour les acteurs des systèmes éducatifs québécois et étrangers. Les recherches indiquent d'ailleurs qu'un nombre plus élevé d'élèves manifeste des troubles du comportement à l'école qu'auparavant (American Psychiatric Association, 1994 ; Gendron, Royer, Bertrand, & Potvin, 2005). Les facteurs de risque pouvant contribuer au développement de troubles du comportement chez les jeunes d'âge scolaire seraient nombreux (Walker, Colvin, & Ramsey, 1995). Parmi ceux-ci figureraient un ensemble d'habiletés liées à la compétence sociale (Bowen, Desbiens, Martin, & Hamel, 2001). La présente étude visait à éprouver, au moyen de la théorie de la généralisabilité, la fiabilité d'un dispositif d'évaluation conçu pour apprécier la compétence sociale d'adolescents en troubles du comportement selon différents indicateurs de compétence sociale (coopération, contrôle de soi, affirmation de soi), différentes sources d'information (adolescent lui-même, parent et enseignant) et différents temps de mesure (début d'année scolaire, mi-année scolaire, fin d'année scolaire). La compétence sociale des adolescents a été mesurée au moyen des versions « adolescent autorévélé », « parent » et « enseignant » du Questionnaire d'évaluation des habiletés sociales (Fortin, Royer, Marcotte, Potvin, & Joly, 2001 ; traduction du Social Skills Rating System de Gresham & Elliott, 1990). La collecte des données a eu lieu simultanément en novembre 2000, en mars 2001 et en juin 2001 dans trois écoles de la grande région de Québec. La compétence sociale des adolescents a été appréciée par trois sources : l'adolescent lui-même, un parent et un enseignant. Afin d'apprécier les qualités métrologiques du dispositif d'évaluation constitué, une étude de généralisabilité a été réalisée. Les coefficients de généralisabilité étant supérieurs à la limite jugée satisfaisante, le dispositif d'évaluation serait donc assez fiable pour assurer la différenciation des adolescents en regard de leur compétence sociale.

Problématique

Les troubles du comportement chez les jeunes d'âge scolaire constituent une préoccupation de plus en plus importante pour les acteurs des systèmes éducatifs québécois et étrangers. Les recherches indiquent d'ailleurs qu'un nombre plus élevé d'élèves manifeste des troubles du comportement à l'école qu'auparavant (American Psychiatric Association, 1994 ; Gendron, Royer, Bertrand, & Potvin, 2005). Tel que rapporté par Gendron et collaborateurs (2003), « Malgré que ces proportions tendent à augmenter d'année en année, ces jeunes sont sous-identifiés par les milieux scolaires (Fortin, Royer, Marcotte, & Potvin, 2000 ; Tremblay, 1998)… En effet, selon l'American Psychiatric Association (1994), la prévalence des troubles du comportement est estimée entre 6 à 16 % chez les garçons et entre 2 à 9 % chez les filles de moins de 18 ans. » (Gendron, Royer, Potvin, & Bertrand, 2003, p.351).

La théorie de l'apprentissage social de Bandura (1986) est utilisée par plusieurs chercheurs et praticiens pour tenter d'expliquer le processus et les déterminants pouvant influencer le développement de troubles du comportement chez les jeunes. Les facteurs de risque pouvant contribuer au développement de troubles du comportement chez les jeunes d'âge scolaire seraient nombreux (Gendron, Royer, Bertrand, Potvin, & Frenette, sous presse ; Walker, Colvin, & Ramsey, 1995). Parmi ceux-ci figureraient un ensemble d'habiletés sociales qui représentent une partie de la compétence sociale (Bowen, Desbiens, Gendron, & Bélanger, 2005 ; Bowen, Desbiens, Martin, & Hamel, 2001). Selon ces mêmes auteurs, il existe bon nombre de programmes d'intervention en milieu scolaire ayant pour but de prévenir ou d'intervenir auprès d'élèves en difficulté de comportement qui démontrent très souvent des lacunes importantes au plan des habiletés sociales.

Malgré leur popularité et leur potentiel certain, l'efficacité des programmes d'entraînement aux habiletés sociales demeure modeste (Gresham, 1998 ; Walker et al., 1995). Parmi les raisons évoquées pour tenter d'expliquer cette situation (Gresham, 1998), on peut ajouter à la liste l'identification imprécise des besoins des élèves en matière d'habiletés sociales et la rareté des outils de mesure spécifiques pour y arriver. Le Questionnaire d'évaluation des habiletés sociales (Fortin, Royer, Marcotte, Potvin, & Joly, 2001 ; traduction du Social Skills Rating System de Gresham & Elliott, 1990) est l'un des rares outils permettant de mesurer le niveau d'habiletés sociales et cela, avec la possibilité d'avoir recours à trois sources différentes d'information (élève, parent, intervenant). Nouvellement validées auprès d'un échantillon québécois, ces versions du questionnaire lorsque utilisées conjointement fournissent tant aux chercheurs qu'aux praticiens une information précieuse afin de pouvoir mieux cibler les moyens et les stratégies de l'intervention (programmes d'entraînement aux habiletés sociales) et d'en mesurer l'efficacité. Il apparaît donc important d'apprécier avec justesse le degré de compétence sociale (voire habiletés sociales), ainsi que la perception de chacun des acteurs (élève, parent, intervenant).

Théorie de la généralisabilité

Dans un cadre où les ressources seraient illimitées, il serait pensable d'avoir recours à tous les indicateurs de compétence sociale inimaginables et à toutes les sources d'information admissibles pour apprécier avec exactitude le degré réel de compétence sociale d'adolescents en troubles du comportement. Pour diverses raisons, cette approche ne peut être mise en pratique. Il faudra plutôt se contenter de déterminer, par le biais d'une étude de généralisabilité (Bain & Pini, 1996 ; Cardinet & Tourneur, 1985), si le résultat moyen observé par le biais d'un nombre limité d'indicateurs de compétence sociale et de sources d'information s'en trouve généralisable à tous les indicateurs de compétence sociale et sources d'information concevables. Dans l'affirmative, il sera alors possible de certifier que le résultat moyen observé, moyennant quelques indicateurs de compétence sociale et sources d'information sélectionnés, traduit le degré réel de compétence sociale d'adolescents en troubles du comportement.

Pour savoir si effectivement les résultats observés au moyen de quelques indicateurs de compétence sociale et de quelques sources d'information s'en trouvent généralisables à tous les indicateurs de compétence sociale et à toutes les sources d'information inimaginables, un coefficient de généralisabilité sera alors calculé. À chaque étude, selon l'intention de la mesure, un choix d'indice devra néanmoins s'imposer : le coefficient de généralisabilité dite « relative » s'emploiera pour hiérarchiser ou comparer entre eux les différents objets de mesure (p. ex. : quel est l'adolescent présentant le degré le plus élevé de compétence sociale ?), tandis que le coefficient de généralisabilité dite « absolue » s'utilisera pour comparer les scores observés à un standard quelconque (p. ex. : point de césure) (Bain & Pini, 1996).

Objectif de la recherche

La présente étude visait à éprouver, au moyen de la théorie de la généralisabilité, la fiabilité d'un dispositif d'évaluation conçu pour apprécier la compétence sociale d'adolescents en troubles du comportement (A) selon différents indicateurs de compétence sociale (I ; coopération, contrôle de soi, affirmation de soi), différentes sources d'information (S ; adolescent lui-même, parent et enseignant) et différents temps de mesure (T ; début d'année scolaire, mi-année scolaire, fin d'année scolaire).

Méthodologie

Le questionnaire

La compétence sociale d'adolescents en troubles du comportement a été appréciée au moyen des versions « adolescent autorévélé » (39 items), « parent » (52 items) et « enseignant » (42 items) du Questionnaire d'évaluation des habiletés sociales (Fortin et al., 2001 ; traduction du Social Skills Rating System de Gresham & Elliott, 1990). Ce questionnaire permet d'évaluer cinq sous-échelles d'habiletés sociales, soit (a) la coopération (ex.: « Je respecte les consignes de l'enseignante »), (b) l'affirmation de soi (ex.: « J'initie des conversations avec mes compagnons de classe »), (c) l'empathie (ex.: « J'écoute mes amis lorsqu'ils ont des problèmes »), (d) le contrôle

de soi (ex.: « Je termine calmement les chicanes avec mes parents ») et (e) la responsabilisation (ex.: « L'élève suit les règles lorsqu'il joue avec les autres »). La somme de ces sous-échelles constitue le score total représentant le degré d'habiletés sociales des adolescents en troubles du comportement. Les sous-échelles « empathie » et « responsabilisation » ne s'adressant qu'aux adolescents et aux parents respectivement, elles ont été retirées des analyses de la présente étude.

La collecte des données

La collecte des données a eu lieu en novembre 2000, en mars 2001 et en juin 2001 dans trois écoles de la grande région de Québec. La compétence sociale des adolescents en troubles du comportement a été appréciée par trois sources : l'adolescent lui-même ($\underline{n} = 49$), un parent ($\underline{n} = 39$) et un enseignant ($\underline{n} = 45$). La participation des adolescents, des parents et des enseignants s'est faite sur une base volontaire.

L'échantillon retenu

Le logiciel EduG (version 1.3) spécialisé en théorie de la généralisabilité ne permettant pas l'emploi de devis non balancés, l'étude repose sur les 23 adolescents en troubles du comportement pour lesquels chacun des indicateurs de compétence sociale (coopération, contrôle de soi, affirmation de soi) a été apprécié par chacune des sources (l'adolescent lui-même, un parent et un enseignant), et ce, à chacun des temps de mesure (début d'année scolaire, mi-année scolaire, fin d'année scolaire).

La démarche d'analyse

Afin d'évaluer les qualités métrologiques du dispositif d'évaluation à quatre facettes croisées ainsi constitué, les données ont été analysées selon la théorie de la généralisabilité au moyen du logiciel EduG (version 1.3). Les observations collectées ont été structurées suivant les plans d'observation, d'estimation et de mesure suivants :

Plan d'observation

Dans la situation de mesure à l'étude, quatre facettes peuvent être distinguées : les Adolescents en troubles du comportement (A), les Indicateurs de compétence sociale (I), les Sources d'information (S) et les Temps de mesure (T). On dénombrait 23 niveaux observés pour la facette Adolescents en troubles du comportement (23 adolescents en troubles du comportement pour lesquels chacun des indicateurs de compétence sociale a été apprécié par chacune des sources, et ce, à chacun des temps de mesure), trois pour la facette Indicateurs de compétence sociale (coopération, contrôle de soi, affirmation de soi), trois pour la facette Sources d'information (adolescent lui-même, parent, enseignant) et trois pour la facette Temps de mesure (début d'année scolaire, mi-année scolaire, fin d'année scolaire). Le plan d'observation d'une étude de généralisabilité demandant entre autres à ce que les facettes soient bien identifiées et les niveaux observés bien dénombrés, ces informations sont synthétisées dans le tableau 1 suivant :

Tableau 1 - Dispositif d'évaluation : plan d'observation.

Facettes	Niveaux observés	Nombre de niveaux
A	A_1 à A_{23}	23
I	coopération, contrôle de soi, affirmation de soi	3
S	adolescent lui-même, parent, enseignant	3
T	début d'année scolaire, mi-année scolaire, fin d'année scolaire	3

Au chapitre des relations entre les facettes, puisque la compétence sociale de chacun des adolescents a été appréciée par chacune des trois sources d'information, au moyen de chacun des trois indicateurs de compétence sociale, et ce, à chacun des trois temps de mesure, la facette Adolescents en troubles du comportement (A) est dite « croisée » avec les facettes Sources d'information (A x S), Indicateurs de compétence sociale (A x I) et Temps de mesure (A x T). Étant donné que chacun des trois indicateurs de compétence sociale a été apprécié par chacune des trois sources d'information, et ce, à chacun des trois temps de mesure, la facette Indicateurs de compétence sociale (I) s'en trouve, quant à elle, « croisée » avec les facettes Sources d'information (I x S) et Temps de mesure (I x T). Enfin, puisque chacune des trois sources d'information a été sollicitée à chacun des trois temps de mesure, la facette Sources d'information (S) se voit « croisée » avec la facette Temps de mesure (S x T). Les données exploitées s'en trouvaient alors structurées suivant le plan d'observation à quatre facettes croisées A x I x S x T, lequel se représente visuellement par le diagramme d'Euler-Venn suivant :

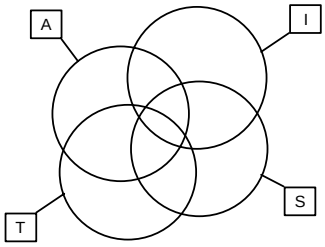


Figure 1 - Illustration du plan d'observation à quatre facettes croisées du dispositif d'évaluation A x I x S x T.

Plan d'estimation

En supposant que les 23 adolescents en troubles du comportement dont les résultats ont été retenus aux fins d'analyse aient été choisis parmi des milliers d'adolescents en troubles du comportement disponibles et les trois temps de mesure puisés à même des milliers de temps de mesure admissibles, les facettes A et T sont qualifiées d'« aléatoires infinies ». Les trois indicateurs de compétence sociale étudiés ayant été sélectionnés parmi les cinq indicateurs du Questionnaire d'évaluation des habiletés sociales (Fortin et al., 2001), il a été choisi de considérer la facette I « aléatoire finie ». Enfin, bien qu'il existe d'autres sources d'information aptes à juger de la compétence sociale d'adolescents en troubles du comportement, notamment un psychoéducateur ou un psychologue, il a été décidé de limiter l'analyse aux niveaux observés de la facette S (adolescent lui-même, parent, enseignant) puisque aucune

autre source potentielle d'information n'a été sollicitée par le questionnaire utilisé. La facette Sources d'information est par conséquent dite « fixée ». Le tableau 2 rapporte ces informations :

Tableau 2 - Dispositif d'évaluation : plan d'estimation.

Facettes	Niveaux observés	Niveaux admissibles	Type
A	23	Infini	Aléatoire infinie
I	3	5	Aléatoire finie
S	3	3	Fixée
T	3	Infini	Aléatoire infinie

Plan de mesure

Dans notre volonté de différencier les adolescents en troubles du comportement (A) quant à leur compétence sociale au moyen des indicateurs de compétence sociale (I), des sources d'information (S) et des temps de mesure (T), le plan de mesure A/IST a été retenu.

Résultats et discussion

Les tableaux 3 et 4 synthétisent, pour la situation de mesure à l'étude, les résultats de l'analyse de la variance et de l'analyse de la généralisabilité obtenus au moyen du logiciel EduG (version 1.3).

Tableau 3 - Analyse de la variance du dispositif d'évaluation A x I x S x T.

Sources de variation	Sommes des carrés	Degrés de liberté	Carrés moyens	Comp. aléatoires	Comp. mixtes	Espér. variance mixtes	Erreurs types	%
A	2568,2158	22	116,7371	2,5200	3,8045	3,8045	1,3050	21
T	171,0757	2	85,5378	0,2841	0,3523	0,3523	0,2991	2
S	551,9742	2	275,9871	-0,2787	0,5533	0,3553	1,2476	2
I	474,4477	2	237,2238	-0,3270	1,0258	0,8206	1,1494	5
AT	400,2576	44	9,0968	-0,0385	0,8416	0,8416	0,2703	5
AS	1954,9147	44	44,4299	3,0339	3,5656	2,3771	1,0585	13
AI	708,4412	44	16,1009	0,4800	1,3661	1,0929	0,4255	6
TS	73,5040	4	18,3760	0,1283	0,1315	0,0877	0,1572	0
TI	50,3929	4	12,5982	0,1220	0,1274	0,1019	0,1088	1
SI	1167,5523	4	291,8881	4,0583	4,0583	2,1644	2,4426	12
ATS	805,1626	88	9,1496	1,8790	2,5815	1,7210	0,4713	10
ATI	334,9404	88	3,8061	0,0978	1,2687	1,0150	0,2262	6
ASI	1010,8921	88	11,4874	2,6583	2,6583	1,4177	0,5842	8
TSI	31,1143	8	3,8893	0,0164	0,0164	0,0087	0,0773	0
ATSI	618,2190	176	3,5126	3,5126	3,5126	1,8734	0,3723	10

Tableau 4 - Analyse de la généralisabilité du plan de mesure A/IST.

Sources de variation	Variance vraie	Sources de variation	Variance d'erreur relative	Variance d'erreur absolue	%
A	3,8045				
		T		0,1174	15
		S		0,0000	0
		I		0,1368	17
		AT	0,2805	0,2805	36
		AS	0,0000	0,0000	0
		AI	0,1821	0,1821	23
		TS		0,0000	0
		TI		0,0057	0
		SI		0,0000	0
		ATS	0,0000	0,0000	0
		ATI	0,0564	0,0564	7
		ASI	0,0000	0,0000	0
		TSI		0,0000	0
		ATSI	0,0000	0,0000	0
Totaux	3,8045		0,5191	0,7789	
Coefficient de généralisabilité relative 0,8799 ; absolue 0,8301					

Les coefficients de généralisabilité relative (0,8799) et absolue (0,8301) étant supérieurs à la limite conventionnellement jugée acceptable (0,80), la précision de la mesure fournie par le dispositif d'évaluation est plutôt satisfaisante. Le dispositif d'évaluation serait donc assez fiable pour assurer la différenciation des adolescents en troubles du comportement en regard de leur compétence sociale sans égard au temps de mesure (T), à la source d'information (S) ou à l'indicateur de compétence sociale (I) privilégié.

Bien que les coefficients de généralisabilité soient de niveau satisfaisant, il a été décidé de procéder à un examen des sources de variation. Ce dernier indique que ce sont les interactions Adolescents en troubles du comportement x Temps de mesure (36 %) et Adolescents en troubles du comportement x Indicateurs de compétence sociale (23 %) qui concourent le plus à l'erreur de mesure dans l'appréciation de la compétence sociale des adolescents.

Une explication plausible à l'interaction Adolescents en troubles du comportement x Temps de mesure résiderait dans la mise en œuvre d'une intervention entre le premier (début d'année scolaire) et le second (mi-année scolaire) temps de mesure. Cette intervention a pris la forme d'une série de 20 ateliers d'entraînement aux habiletés sociales en classe sur une période de 20 semaines (Gendron, Royer, & Morand, 2005). Par ailleurs, d'autres hypothèses explicatives seront ultérieurement explorées, notamment en regard du fait que les adolescents en troubles du comportement ont tendance à se surestimer au niveau de leurs habiletés sociales préalablement à ce type d'intervention.

Fait à noter, en ce qui a trait à l'interaction Adolescents en troubles du comportement x Indicateurs de compétence sociale aucune explication n'a été trouvée. Considérant les résultats préliminaires obtenus, il y aurait lieu d'explorer l'hypothèse à l'effet que l'indicateur « coopération » serait à la base de cette variation.

Conclusion

Tout d'abord, il est à noter que la présente étude est basée sur l'utilisation d'un outil de mesure validé auprès d'un échantillon québécois. Largement utilisé, le Questionnaire d'évaluation des habiletés sociales (Fortin et al., 2001 ; traduction du Social Skills Rating System de Gresham & Elliott, 1990) a aussi la qualité de pouvoir recueillir la perception de différentes sources d'information (adolescent lui-même, parent, enseignant).

La présente étude a permis de confirmer la fiabilité de cet outil de mesure dans un dispositif d'évaluation visant à assurer la différenciation des adolescents en troubles du comportement en regard de leur compétence sociale sans égard au temps de mesure (T), à la source d'information (S) ou à l'indicateur de compétence sociale (I) privilégié.

Une recherche plus élaborée est en préparation afin d'évaluer la fiabilité de ce dispositif d'évaluation avec d'autres populations (niveaux primaire et secondaire, clientèle régulière et à risque, etc.).

Références

American Psychiatric Association (1994). *DSM-IV: Diagnostic and statistical manual of mental disorders*. Washington: APA.

Bain, D., & Pini, G. (1996). *La généralisabilité : mode d'emploi*. Genève : Centre de recherches psychopédagogiques.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bowen, F., Desbiens, N., Gendron, M., & Bélanger, J. (2005). L'acquisition et le développement des habiletés sociales. In L. Massé, N. Desbiens, & C. Lanaris (Eds.), *Les troubles de comportements à l'école: diagnostic et intervention*, 213-227. Montréal : Gaëtan Morin Éditeur Ltée.

Bowen, F., Desbiens, N., Martin, C., & Hamel, M. (2001). La compétence sociale. In Gouvernement du Québec (Ed.), *6 12 17 Nous serons bien mieux! Les déterminants de la santé et du bien-être des enfants d'âge scolaire*, 55-99. Québec : Les Publications du Québec.

Cardinet, J., & Tourneur, Y. (1985). *Assurer la mesure*. Berne : Peter Lang.

EduG (Version 1.3) [Logiciel informatique]. (2002). Lévis : EDUCAN.

Fortin, L., Royer, É., Marcotte, D., Potvin, P., & Joly, J. (2001). Épreuves de validité d'une mesure d'habiletés sociales auprès d'adolescents québécois à l'école secondaire. *Psychologie et Psychométrie, 22(1)*, 23-44.

Gendron, M., Royer, É., Bertrand, R., & Potvin, P. (2005). Les troubles du comportement, la compétence sociale et la pratique d'activités physiques chez les adolescents. *Revues des Sciences de l'Éducation, 31(1)*, 211-233.

Gendron, M., Royer, É., Bertrand, R., Potvin, P., & Frenette, É. (sous presse). Relation entre l'adaptation psychosociale, les habitudes de vie et les habiletés sociales chez les adolescents en troubles du comportement. *Revue de Psychoéducation, 35(2).*

Gendron, M., Royer, E., & Morand, C. (2005). *Programme PEC – Pratiquons Ensemble nos Compétences : Guide II – Volet scolaire.* Centre de santé de Portneuf.

Gendron, M., Royer, É., Potvin, P., & Bertrand, R. (2003). Troubles du comportement, compétence sociale et pratique d'activités physiques chez les adolescents : enjeux et perspectives d'intervention. *Revue de Psycho-Éducation, 32(2)*, 349-372.

Gresham, F. M. (1998). Social skills training: Should we raze, remodel, or rebuild? *Behavioral Disorders, 24(1)*, 19-25.

Gresham, F. M., & Elliott, S. N. (1990). *Social skills rating system manual*. Circle Pines: American Guidance Service.

Walker, H. M., Colvin, G., & Ramsey, E. (1995). *Antisocial behaviour in school: Strategies and best practices.* Toronto: Brooks/Cole.

COMPRÉHENSION DU BULLETIN SCOLAIRE DANS UN CONTEXTE D'IMPLANTATION D'UN NOUVEAU PROGRAMME DE FORMATION AU QUÉBEC: ANALYSE DES REPRÉSENTATIONS DES PARENTS DU 1ER DU 3E CYCLE DU PRIMAIRE D'UNE COMMISSION SCOLAIRE

Micheline-Joanne Durand, Université de Montréal
Nathalie Loye, Université de Montréal

Mots-clés : école primaire, bulletin, perception des parents

Résumé:
Cette présentation fait état d'une recherche exploratoire qui avait pour objectif de dresser le portrait des représentations des parents d'élèves du 1er et du 3e cycle du primaire d'une commission scolaire en regard du bulletin scolaire implanté à la suite du Renouveau pédagogique mis en place au Québec. En plus du curriculum et de l'intégration massive des élèves en difficulté, ce sont toutes les pratiques d'évaluation qui se sont transformées et qui ont mené à l'élaboration d'un bulletin descriptif à interprétation critériée. Des données empiriques recueillies à l'aide de questionnaires permettent de situer le niveau de compréhension 1) des enseignants qui ont contribué à l'élaboration du bulletin; 2) des parents de leurs élèves; 3) des enseignants qui ont utilisé le bulletin lors de la première année d'implantation du nouveau *Programme de formation*. Les résultats qui vous sont communiqués dans le présent document ont été compilés à partir des données recueillies auprès d'un groupe de parents en réponse à un premier questionnaire portant sur la compréhension du bulletin et du portfolio dans le contexte d'une rencontre tripartite. Ces données permettent de connaître le niveau de satisfaction des parents et font ressortir les aspects du bulletin qui leur posent problème et suscitent leur résistance face aux changements apportés dans la notation. Cette première recherche représente l'amorce d'une autre recherche collaborative, qui pourrait être effectuée avec divers milieux scolaires ou auprès de clientèles diversifiées dans d'autres contextes scolaires.

Problématique

Le *Renouveau pédagogique* amorcé depuis le début des années 2000 au Québec par l'implantation d'un nouveau programme de formation a amené des transformations dans les pratiques d'évaluation des apprentissages des élèves et dans la façon d'en communiquer les résultats aux parents qui ont ébranlé les représentations que ceux-ci s'en faisaient. Ainsi, les changements apportés à la démarche d'évaluation et particulièrement au bulletin scolaire font régulièrement la manchette de nombreux médias tant par le caractère descriptif du document, le langage et la notation qu'on y emploie que par les modalités d'évaluation mises en pratique.

Le contexte actuel pose un défi de taille au milieu scolaire. Celui-ci est invité à revoir non seulement les pratiques d'évaluation en cours, mais à proposer un bulletin fidèle aux nouvelles orientations qu'il a adoptées. Le bulletin ne doit plus, comme il le faisait jadis, se limiter à un inventaire des connaissances acquises, mais rendre compte

des compétences développées par l'élève et fournir un bilan de ses apprentissages. Afin d'instaurer ce changement, une commission scolaire de la région nord de Montréal, d'un commun accord avec les enseignants et les directions des 54 écoles primaires, a décidé de s'engager dans une démarche de recherche-formation sous la conduite d'un chercheur et d'un professionnel de recherche afin de créer de nouveaux bulletins uniformes pour tous les établissements scolaires et de moduler leur implantation de façon graduelle.

Plusieurs enseignants volontaires ont participé à l'un ou l'autre des groupes de développement spécifiques à chacun des cycles dans le but d'élaborer, d'expérimenter et de valider un modèle de bulletin scolaire simple, compréhensible et accessible aux parents. De plus, ils se sont engagés à utiliser un portfolio et à organiser une entrevue tripartite au moment de la remise du premier bulletin et à en assurer le suivi. Une attention particulière a été accordée aux réactions des parents face au bulletin. Les enseignants ont expliqué le bulletin lors de la première rencontre de parents puis, à la première remise du bulletin. Par la suite, ils ont recueilli l'opinion des parents, à l'aide d'un questionnaire, sur la rencontre tripartite de même que sur la compréhension des énoncés du bulletin et du bilan de fin de cycle. Pour la présente communication, les données recueillies auprès des groupes du 1er cycle et du 3e cycle ont été sélectionnées afin d'en dégager les principales tendances.

L'information transmise aux parents par le bulletin et le portfolio devait permettre aux parents de constater la progression de leur enfant dans le développement des différentes compétences et ce dans le contexte d'une rencontre tripartite où celui-ci est le principal acteur de cette transmission. Avant d'aller plus loin, définissons d'abord ces concepts afin de nous assurer d'une compréhension commune.

Cadre conceptuel

Informer les parents sur la progression des compétences de leur enfant fait partie de la tâche de l'enseignant. Plusieurs outils de communication (agenda, dossier, bulletin, bilan, etc.) et modalités d'évaluation (autoévaluation, grilles d'observation, coévaluation, etc.), présentés à des moments opportuns, permettent aux parents d'avoir un portrait d'ensemble des apprentissages réalisés en classe par leur enfant. Le bulletin (ou livret scolaire) est un document officiel, un référentiel des compétences scolaires évaluables et évaluées, une mémoire des cheminements de l'élève, un outil de communication avec les parents et un outil d'orientation (Jorro, 2000). Dans le contexte de l'évaluation formative, le bulletin présente un ensemble de jugements qui circonscrivent les points forts et les points faibles, qui résument une multitude d'observations effectuées en cours de cycle, qui illustrent la progression vers la maîtrise de la compétence (Scallon, 2004). «Avec les nouveaux programmes, la notation chiffrée ne tient plus; pour rendre compte de la progression des élèves dans le développement de leurs compétences, on devra de plus en plus recourir à des échelles de jugement […] des échelles uniformes […] à caractère universel. » (Scallon, 2004, p. 178)

Le bulletin s'adresse d'abord aux élèves et à leurs parents, puis aux différents intervenants qui travaillent auprès d'eux. « Une des règles de l'art de la communication efficace est que le moyen et le langage utilisés doivent être adaptés

aux destinataires. » (Laurier, 2005, p.147) C'est dire qu'il faut se servir d'un vocabulaire clair et précis pour que l'élève puisse comprendre son bulletin et le parent y retrouver les informations attendues quant à la progression de son enfant. Le bulletin élaboré par le groupe de recherche tient compte de ces deux aspects : emploi d'une échelle d'appréciation qualitative (voir Tableau 1) et d'une formulation claire des énoncés illustrant les forces (+) et les défis (✓) de l'élève pour une compétence donnée.

Tableau 1 : Illustration de l'échelle d'appréciation 1er cycle et 3e cycle

Si votre enfant obtient	**Il ou elle**
1	chemine très facilement
2	chemine facilement
3	chemine avec quelques difficultés
4	chemine avec beaucoup de difficulté
	non évalué
Appréciation des indicateurs	
+	progresse bien
✓	à travailler
	non vérifié

Le portfolio est un outil de communication entre l'école et les parents; il a pour but d'informer ces derniers, de façon continue, des apprentissages réalisés en classe par leur enfant. Le portfolio ou dossier d'apprentissage est une collection de travaux inachevés ou de productions terminées, qui permet à l'élève de rendre compte de ses réalisations, de ses progrès et de l'évaluation qu'il en fait, soit son autoévaluation. Pour Simon et al. (1994) c'est un « recueil cumulatif et continu d'indicateurs du cheminement de l'élève dans ses apprentissages, sélectionnés et commentés par l'élève et l'enseignant […] à des fins d'évaluation». (p. 29) «Consigner les travaux des élèves dans un même endroit, les commenter ; amener les élèves à réfléchir sur leurs stratégies et la démarche effectuée ; informer les parents des progrès de leur enfant, voilà quelques-unes des multiples utilités du portfolio» (Durand, 2005. p.26) lequel est employé en complémentarité avec le bulletin. Il permet d'illustrer le jugement posé par l'enseignant.

L'entretien ou la rencontre tripartite est une tendance qui reflète un mouvement général en matière d'évaluation authentique et qui s'oriente vers la mise en place de situations d'évaluation signifiantes, plus proches des conditions d'utilisation des savoirs et des savoir-faire dans la vie. (Allal et coll., 1998) «La rencontre tripartite au premier bulletin est très importante car elle permet d'expliquer aux parents comment

comprendre le contenu du portfolio et du bulletin. Elle permet surtout de donner confiance à l'élève pour qu'il puisse présenter lui-même ce qu'il fait en classe.» (Durand, 2005, p.27)

Peu de recherches se sont appliquées à vérifier les représentations que se font les parents du bulletin de leur enfant et la recherche d'écrits portant sur une combinaison de deux des trois aspects traités s'est avérée peu féconde.

La recherche dans le domaine

Le but des bulletins est d'informer les parents des progrès de l'enfant. Cela fournit l'occasion aux enseignants de donner des rétroactions positives et de s'exprimer sur les forces et faiblesses des enfants afin que ceux-ci puissent s'améliorer et que les parents puissent les soutenir adéquatement. (Newbury 2005) Selon certaines études américaines, les parents sont souvent insatisfaits de l'information que contiennent les bulletins. Ceux-ci ne seraient pas suffisamment descriptifs et ne rendraient pas compte des apprentissages et des faiblesses de leur enfant. (Deborah, 2005; Simon, 2006) L'instauration de bulletins descriptifs vise donc à améliorer la communication des acquis et des faiblesses de l'élève pour pallier aux inconvénients des bulletins traditionnels. (Deborah, 2005) L'appui des parents, dans le contexte de la réforme scolaire, est donc essentiel à la réussite des mesures qu'elle impose. (Xue, Y., Meisels, S.J., Bickel, D.D. & al. 2002) Toutefois, le changement dans les modes d'évaluation est rarement bien reçu par les parents, bien souvent parce qu'ils n'ont pas été suffisamment informés des buts de la réforme. (Norman & Cartwright 1998; Newbury 2005)

Les parents fondent leurs opinions sur leur propre expérience scolaire et sur les évaluations qui avaient cours lorsqu'ils étaient eux-mêmes des élèves. Ce faisant, ils sont plus à l'aise avec les types d'évaluation se servant de lettres ou de chiffres qu'avec des indicateurs plus complexes. (Greene, 1991; Norman & Cartwright 1998; Xue & al., 2002; Newbury 2005) Les parents veulent s'assurer des bonnes performances de leur enfant. En prenant connaissance de bulletins où les progrès de l'enfant sont indiqués par des lettres ou des chiffres, ils mesurent plus aisément les réussites en fonction de celles des autres élèves de la classe. (Xue & al, 2002)

Selon ces auteurs, 81% des parents sont satisfaits des bulletins descriptifs. Ils sont persuadés de pouvoir y lire les forces, les faiblesses et les progrès de leur enfant. Les parents qui ont un enfant plus faible tendent à favoriser le bulletin descriptif, croyant qu'il met en valeur les différents talents de l'enfant. Au contraire, les parents dont l'enfant est plus doué croient que le bulletin descriptif empêche l'enfant d'obtenir des notes plus élevées, ce qui risque de le démotiver. (Xue & al., 2002; Munk & Bursuck, 2001) Une autre tendance se dessine chez les parents, au fur et à mesure que l'enfant vieillit, selon laquelle plus les élèves évoluent dans leur cheminement scolaire, c'est-à-dire vers la deuxième ou la troisième année, plus les parents demandent que les bulletins descriptifs soient remplacés par des bulletins avec des lettres ou des chiffres, de manière à pouvoir comparer le rendement de leur enfant avec celui des autres élèves de la classe. (Greene, 1991; Xue & al., 2002; Newbury, 2005)

En ce qui concerne les rencontres, les parents désirent des séances d'information, non seulement pour mieux comprendre le bulletin de leur enfant, mais également pour entrer en communication avec le professeur et discuter des apprentissages de l'élève. (Greene, 1991; Norman & Cartwright 1998) Par ailleurs, plus l'élève vieillit, plus les

parents manifestent le désir de rencontrer l'enseignant pendant une période plus longue que celle qu'on leur propose généralement. (Greene, 1991) Ils expriment également le besoin d'obtenir des descripteurs spécifiques dans le bulletin et davantage de commentaires écrits par l'enseignant. (Greene, 1991 : Newbury, 2005)

Plus près de nous, une enquête du Centre de Recherche et d'Intervention pour la Réussite Scolaire CRIRÈS (Déniger et al., 2004), menée, entre autres, auprès des parents de l'ensemble des commissions scolaires du Québec, a permis de constater que les parents ne sont favorables qu'à un seul aspect du nouveau mode d'évaluation, soit celui qui leur permet de reconnaître les forces de l'élève et ce qu'il y a à améliorer. Cependant, les problèmes qu'ils relèvent concernent principalement l'implantation trop rapide du programme (23% des cas), l'appropriation du programme de formation (19% des cas) et l'évaluation des compétences (9% des cas). Ils affirment que l'évaluation est trop vague et trop difficile à comprendre. De plus, ils trouvent difficile de suivre la progression de l'enfant et préconisent l'établissement d'un mode d'évaluation standard de façon que le bulletin soit plus compréhensible.

Le présent projet va donc tenter de situer le niveau de satisfaction et de compréhension du bulletin scolaire comme outil de communication auprès des parents des élèves du 1er et du 3e cycle d'une commission scolaire et de dresser le portrait de ce qui pose problème. Il propose également quelques pistes pour expliquer les tendances identifiées.

Méthodologie

La recherche-formation en lignée directe avec la recherche collaborative

Pour faciliter l'appropriation progressive de ce changement, certaines modalités ont été mises en place par le service des ressources éducatives d'une commission scolaire de la région nord de Montréal et par moi-même, en tant que chercheure dans le cadre d'une recherche – formation de type collaboratif. Cette recherche se distingue rapport qu'elle veut établir avec le milieu scolaire et n'a pas pour but de témoigner sur les acteurs du milieu scolaire (enseignants, parents, conseillers pédagogiques), ceux-ci étant considérés comme des partenaires de l'investigation et contribuant à la recherche dans une réflexivité conjointe au développement de la pratique. (Desgagné, Bednarz, Couture, Lebuis et Poirier, 2001)

Plusieurs enseignants (13 au 1er cycle et 45 au 3e cycle), en compagnie de conseillers pédagogiques, ont participé à cette recherche-formation dont le but principal était d'élaborer un bulletin commun pour toutes les écoles primaires de la commission scolaire. Voici les étapes qui ont été suivies : appropriation de la démarche d'évaluation selon l'approche par compétences, conception, expérimentation et validation de situations d'apprentissage et d'évaluation, mise en place d'un portfolio d'apprentissage et réalisation d'une rencontre tripartite en vue de la remise du premier bulletin. Un plan de consultation a aussi été établi afin de sonder, à l'aide de questionnaires, les principaux acteurs dont les parents, à la suite de la rencontre tripartite. Ce sont les résultats d'un de ces questionnaires qui ont servi à l'analyse et à la présentation des présents résultats.

Les participants

L'échantillon des répondants était constitué des parents d'élèves venant des classes d'enseignants qui s'étaient portés volontaires pour élaborer, expérimenter et valider la démarche d'évaluation devant mener à l'utilisation d'un bulletin de type descriptif, par compétences. La première catégorie comprend des parents provenant de 13 classes du 1er cycle dont 8 de la première année et 5 de la deuxième année du cycle tandis que la deuxième catégorie correspond aux parents des élèves des enseignants du 3e cycle dont 34 groupes de 5e année et 11 groupes de 6e année. Ces parents sont issus de milieux socio-économiques différents : 29 groupes viennent de milieux dont l'indice de défavorisation est de 1,2 ou 3, soit favorisé, que nous avons regroupé dans *fav 1*; 15 groupes appartiennent aux milieux dont l'indice de défavorisation est de 4,5 ou 6, soit moyennement favorisé, que nous avons regroupé dans *fav 2;* le dernier groupe est formé de parents issus de milieux dont l'indice de défavorisation s'apparente aux 7,8, et 9, soit défavorisé que nous avons regroupés dans *fav 3.* Tous ces parents ont répondu à un premier questionnaire lors de l'organisation de la rencontre tripartite au cours de laquelle le bulletin par compétences à été remis pour la première fois. Parmi ces parents, un certain nombre ont inscrit des commentaires ou proposé des suggestions. Le tableau 2 illustre cette répartition. Au total, ce sont 1286 questionnaires et 832 commentaires qui ont été compilés.

Tableau 2 : Répartition des parents du 1er et du 3e cycle ayant répondu aux questionnaires selon leur niveau scolaire et de défavorisation.

Nb de groupes selon le niveau scolaire	Nb de groupes selon le niveau de défavorisation			Nb de parents ayant répondu aux questions	Nb de parents ayant inscrit des commentaires
	1	2	3		
2001-2002 : 1er cycle	13 groupes			266	210
1re année /8 groupes	3	5	0	169	133
2e année /5 groupes	2	0	2	97	77
2003-2004 : 3e cycle	45 groupes			1020	622
5e année /34 groupes	18	9	7	752	458
6e année/ 11 groupes	6	1	4	268	164
Total : 58 groupes	29	15	13	1286 questionnaires	832 commentaires

Le questionnaire

Le questionnaire comprenait des questions fermées incluant une échelle d'appréciation uniforme qualitative ainsi que la possibilité de laisser des commentaires. Douze questions se rapportant aux thématiques suivantes ont été soumises aux parents : 1) la compréhension générale du bulletin (3 questions), 2) la compréhension de l'information consignée en lecture, écriture et mathématique (3 questions), 3) la participation de l'élève à la présentation de son portfolio et de son bulletin (3 questions) et 4) la rencontre tripartite (3 questions). Les parents devaient

répondre en donnant leur niveau de satisfaction pour les questions 1 à 8, 10 et 11 selon 3 échelons : oui, très bien; en partie, plus ou moins bien; non, pas du tout. La question 9 proposait 4 réponses: très satisfaisante, satisfaisante, plus ou moins satisfaisante, non satisfaisante et la question 12 présentait un choix dichotomique : oui ou non.

Les analyses

Les données accessibles étaient compilées pour chacun des groupes-classes, ceux-ci servant d'unité d'analyse de base. Une première analyse descriptive réalisée à l'aide du logiciel Excel 2003 a permis de traiter les questions fermées en données compilées par groupes d'élèves afin de pouvoir réaliser des comparaisons entre les cycles (les niveaux scolaires), les milieux de défavorisation et selon les quatre thématiques, et de dégager les tendances illustrées par des graphiques. Les graphiques produits servent seulement à identifier les tendances, mais ces dernières ne sont pas appuyées par des tests statistiques. Une deuxième analyse est qualitative et plus approfondie Elle est réalisée avec le logiciel Alceste et porte sur les commentaires et suggestions des groupes de parents. Les catégories identifiées servent à expliciter et à illustrer les graphiques. Les commentaires des parents ont été ensuite classés selon que leur opinion était favorable ou défavorable au nouveau bulletin et aux modalités de la rencontre tripartite.

Résultats

Ce premier questionnaire s'adressait aux parents et visait à explorer leur taux de satisfaction relativement à la rencontre tripartite, au cours de laquelle l'élève leur présentait son dossier d'apprentissage et son bulletin. L'enseignant accompagnait l'élève dans cette présentation et répondait aux questions des parents selon la demande. Pour la plupart – parents, élèves et enseignant –, il s'agissait d'une première expérience. Douze questions fermées étaient posées aux parents qui pouvaient également ajouter leurs commentaires et opinions.

Pour chacune des quatre thématiques (voir tableau 3), une analyse descriptive des résultats a été faite en tenant compte des réponses aux questions fermées et une analyse qualitative a été réalisée à partir des commentaires recueillis auprès des parents du 1er et du 3e cycle.

Tableau 3 : Thématiques visées par les questions fermées du questionnaire

Thématiques	Questions
1. La compréhension de l'information consignée sur le bulletin	Q7, Q3, Q8
2. La compréhension spécifique de l'information consignée sur le bulletin	Q3, Q4, Q5
3. La participation des élèves à la présentation de leur portfolio	Q1, Q2, Q11
4. La rencontre tripartite	Q9, Q10, Q12

La compréhension générale de l'information consignée sur le bulletin

Trois questions concernent la compréhension générale de l'information consignée sur le bulletin. Le graphique 1 montre que les parents du 1er cycle, et en particulier de la première année, ont une meilleure compréhension générale du bulletin que les parents du 3e cycle. Toutefois, tous les parents comprennent le bulletin (Q7), estiment être bien informés sur ce qu'apprend leur enfant (Q3) et avoir un bon portrait des apprentissages qu'il réalise (Q8).

Niveau 6 /Fav 1 : «Ce bulletin est différent des autres mais je le trouve plus précis. Il est comme un autoportrait.»

En ce qui concerne l'influence du milieu de favorisation, aucune différence n'est décelable au 1er cycle alors que les parents du milieu de favorisation 2 tendent à être légèrement plus satisfaits que ceux des milieux de favorisation 1 et 3.

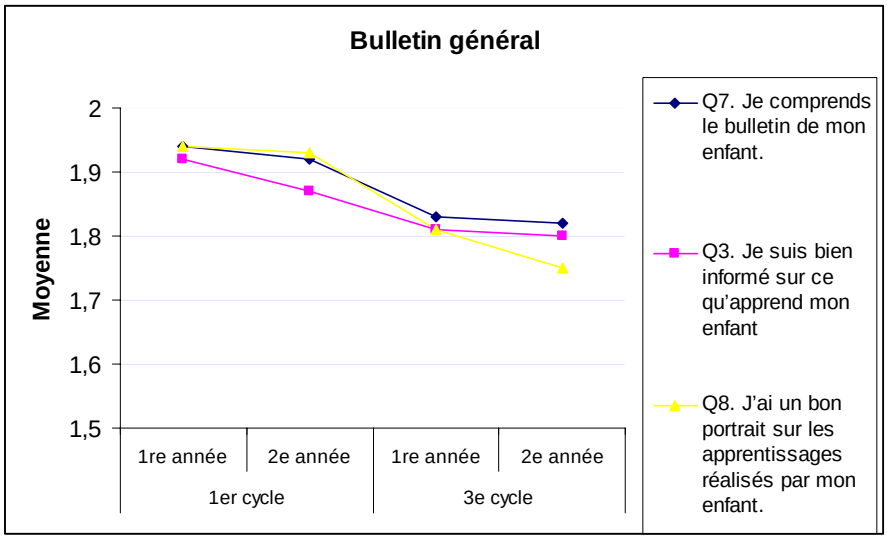


Figure 1 : La compréhension générale des parents quant à l'information consignée sur le bulletin selon le niveau scolaire du 1er et du 3e cycle du primaire.

Dans son ensemble, le degré de compréhension des parents des élèves du 3e cycle est légèrement inférieur à celui du 1er cycle et les commentaires des parents aident à expliquer ce phénomène. En effet, les parents du 3e cycle ont réagi plus fortement aux changements occasionnés par l'implantation d'un nouveau programme à la suite de quatre ou cinq années de communications sous la forme d'un bulletin plus traditionnel.

Niveau 6/ Fav 3 : «C'est parce qu'il y a un changement et quand il y a un changement, il faut s'adapter. Pas de notes mais plus d'explications claires sur la compréhension est un point positif.»

Toutefois, on trouve aux deux cycles un bon nombre de parents qui préfèrent le bulletin traditionnel avec des notes sous forme de pourcentages et des indications en rapport avec la moyenne du groupe.

Niveau 1/fav 2 : « Je n'aime pas ce type de bulletin avec des « + » et des « ✓». Ça ne donne pas une idée claire du niveau. Et même de 1 à 3, on ne connaît pas de façon précise l'endroit où notre enfant se situe. »

Niveau 6 /Fav 1 : «Le bulletin me permet de voir les forces et les faiblesses de mon enfant. La seule chose que je trouve dommage, c'est que nous ne sommes pas en mesure d'évaluer notre enfant VS la classe.»

Enfin, certains parents auraient apprécié qu'on inscrive des indications sur le bulletin concernant le comportement de leur enfant.

Niveau 6 /Fav 1 : «Nous aimerions une section décrivant le comportement de notre garçon en classe au bulletin.»

La compréhension spécifique de l'information consignée sur le bulletin.

La compréhension du bulletin est précisée à l'aide de trois questions qui portent spécifiquement sur la lecture, l'écriture et les mathématiques. Les parents s'estiment en général bien informés sur les apprentissages de leurs enfants en lecture (Q3), en écriture (Q4) et en mathématiques (Q5). Dans son ensemble (voir figure 2), le degré de satisfaction des parents des élèves du 3e cycle est légèrement inférieur à celui du 1er cycle sauf en mathématiques. En effet les parents du 1er cycle se disent moins bien informés sur ce qui concerne les mathématiques en première année, ce qui peut s'expliquer par le fait que beaucoup de temps d'apprentissage est monopolisé en début d'année pour l'initiation à la lecture et à l'écriture. Toutefois cette tendance est également présente en deuxième année.

Niveau 1 /Fav 1 : «En lecture, j'ai apprécié de voir l'évolution en deux semaines sur la reconnaissance des mots.»

Le milieu de favorisation ne semble pas avoir une influence particulière sur la compréhension des parents dans les deux cycles.

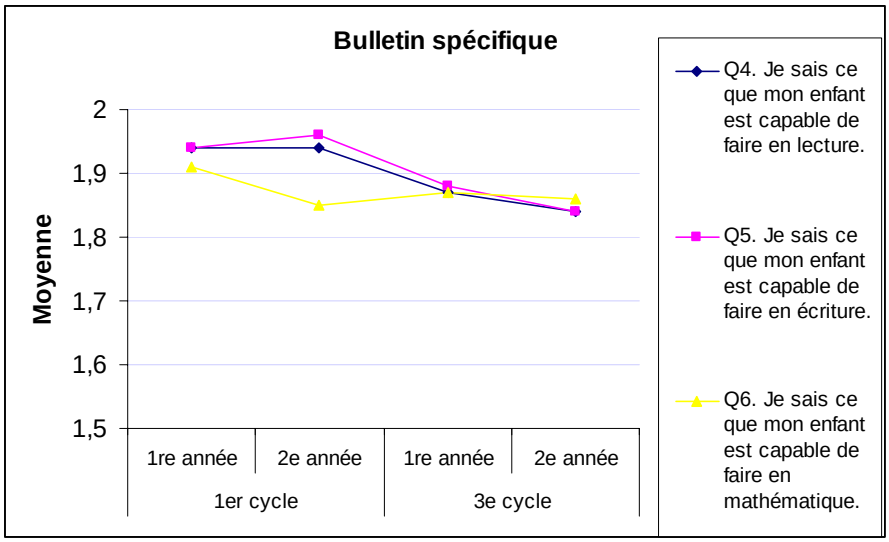


Figure 2- La compréhension des parents de l'information consignée en lecture, écriture et mathématique selon le niveau scolaire du 1er et du 3e cycle du primaire.

Les données recueillies auprès de l'ensemble des parents permettent d'affirmer que les parents ont bien compris le bulletin et qu'ils se considèrent comme bien informés sur les apprentissages réalisés par leur enfant. Ce sont là les principales fonctions du bulletin. Les résultats du 3e cycle sont également légèrement inférieurs à ceux du 1er cycle, mais ce phénomène s'explique par le fait que ces parents ont réagi plus fortement à la notation exprimée de façon qualitative dans ces matières, évaluées auparavant à l'aide de pourcentage, tandis que d'autres en ont constaté les avantages.

> *Niveau 5/ Fav 1 :« Je vois plus en détail le cheminement de ma fille, là où elle a de la difficulté; par exemple, avant c'était une note en français écriture. Maintenant, je sais que c'est avec ses accords qu'elle doit travailler. Je ne suis pas à l'aise totalement avec ce nouveau bulletin, car j'ai encore la notion des notes pour évaluer ses connaissances mais je suis positive à ce changement.»*

> *Niveau 2/fav 1 : « Par contre, nous aimions beaucoup le bulletin avec note surtout pour le français et les mathématiques. »*

Un certain nombre de parents des deux cycles réitèrent leur préférence pour la moyenne du groupe afin de pouvoir situer leur enfant par rapport aux autres

> *Niveau 1/fav 1 : « Avoir une moyenne du groupe pour être capable de situer l'enfant au niveau du groupe.»*

La participation des élèves à la présentation de leur portfolio

Trois autres questions portent sur la participation des élèves lors de la présentation de leur portfolio (Q1), de leur bulletin (Q2) ainsi que sur la fierté qu'ils ressentent (Q11). Le bulletin était examiné par les parents après la présentation par les élèves de leur portfolio et après les commentaires de leur enseignant. Cette

présentation a certainement aidé à la compréhension du bulletin, lequel résumait en quelque sorte ces informations.

Les parents des élèves du 1er cycle et du 3e cycle ont jugé que leur enfant avait été capable de présenter son portfolio (Q1) et qu'il était fier de le faire (Q11). Par contre la présentation du bulletin (Q2) a été plus difficile pour les enfants de 1re année (voir figure 3), ce qui est normal compte tenu du fait qu'elle avait lieu au début (octobre, novembre) de l'année scolaire. Très peu de commentaires des parents illustrent leur insatisfaction face à cette difficulté. Ici encore, le degré de satisfaction des parents ne varie pas en fonction du milieu de défavorisation.

Dans son ensemble, la présentation du portfolio a suscité de nombreux commentaires tant de la part des parents du 1er cycle que par ceux du 3e cycle. En effet, plus de la moitié des commentaires qu'ils ont écrits portent sur la source de motivation, de responsabilisation et de fierté de leur enfant et sur ses capacités à s'autoévaluer et à porter un jugement critique.

> *Niveau 1/fav 2 : « L'idée de demander à l'enfant de présenter son portfolio est très bonne et permet beaucoup d'échanges et de dialogues avec l'enfant sur son travail. Nous croyons que le portfolio nous en apprend beaucoup plus que le bulletin et est très utile pour valoriser l'enfant. Jessica était fière de nous le présenter.»*

C'est là un élément qui mérite d'être souligné car il s'inscrit dans les orientations de la réforme : viser la responsabilisation des enfants dans leur apprentissage et leur évaluation.

> *Niveau 6 /Fav 1 : «Au début, j'étais incertaine, mais en voyant les yeux de mon enfant lorsqu'il m'a présenté son portfolio, j'ai réalisé qu'il était fier de ce qu'il avait accompli. Finalement, j'aime beaucoup cette façon de fonctionner, mon enfant aussi et son estime de soi vient de grimper en flèche.»*

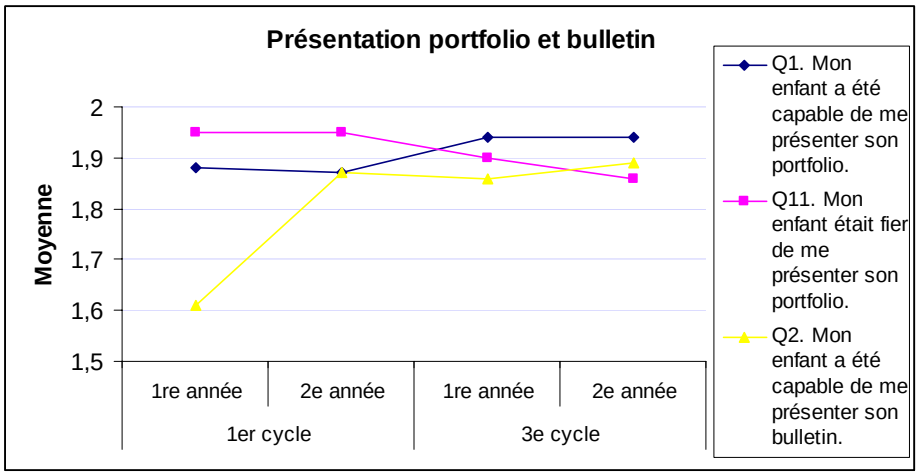


Figure 3 - La satisfaction des parents en regard de la participation de l'élève à la présentation de son portfolio et de son bulletin selon le niveau scolaire du 1er et du 3e cycle du primaire

La rencontre tripartite

Enfin les trois dernières questions portaient sur le rencontre tripartite. À l'occasion de cette rencontre, les enfants présentaient leur portfolio et leur bulletin à leurs parents et les enseignants circulaient pour répondre aux questions ou donner de plus amples informations. Quatre ou cinq enfants regroupés selon le même horaire, disposaient d'environ 30 minutes pour leur présentation. Ces trois questions portent sur le degré de satisfaction général (Q9), la participation des parents à cette rencontre (Q10) et le désir de répéter cette expérience (Q12). Contrairement aux autres questions pour lesquelles on proposait trois choix de réponses, la question 9 proposait quatre réponses possibles et la question 12 était dichotomique (oui ou non).

Les parents des deux cycles ont dit avoir aimé participer à cette rencontre (Q10), qu'ils satisfaisante ou très satisfaisante (Q9). De plus, ils pensent qu'une telle rencontre devrait avoir lieu chaque année (Q12). Encore une fois, le milieu de défavorisation ne semble pas jouer de rôle ici.

> *Niveau 1/fav 1 : « J'aime beaucoup ces rencontres et je pense que c'est très important que les enfants puissent avoir du temps dans son milieu scolaire où il peut recevoir ses parents. Présenter son bulletin est toujours un moment où notre enfant est très fier.»*
>
> *Niveau 2/fav 3 : « Je trouve qu'avec cette rencontre, je peux mieux voir et comprendre ce que mon enfant fait en classe et en plus, elle peut elle-même s'autoévaluer donc elle voit où elle doit s'améliorer.»*

Dans son ensemble, c'est la thématique qui présente le moins de différences entre les deux cycles. Toutefois, plusieurs commentaires sont à l'effet que les parents auraient souhaité pouvoir rencontrer l'enseignante individuellement, En particulier pour les parents d'élèves de la première et de la sixième année.

Niveau 1/fav 2 : « J'ai aimé la rencontre mais j'aurais aimé mieux une rencontre individuelle avec son professeur qu'en groupe.»

Niveau 6/Fav 3 : «Je désire que les rencontres se fassent individuellement avec le professeur. J'aimerais que la présentation se fasse l'élève et le prof pour nous expliquer le développement du bulletin avec les commentaires du professeur et de l'élève en même temps. Connaître les choses à travailler.»

Dans plusieurs commentaires, tant au 1er qu'au 3e cycle, le temps accordé à la rencontre était remis en question. Certains parents ont trouvé la rencontre trop courte alors que d'autres souhaitaient que la formule permette plus de temps pour communiquer. Quelques parents ont souhaité que leur enfant termine la présentation de leurs travaux à la maison. D'autres ont suggéré que l'enfant présente son portfolio et son bulletin à la maison, ce qui leur donnerait la possibilité de rencontrer l'enseignant individuellement par la suite.

Niveau 5 /Fav 2 : «Avoir minimum 30 minutes complet avec l'enfant. Le temps a manqué quelque peu.»

Niveau 1 /Fav 2 : «Le concept de la rencontre multiple est long, mais par contre nous avons le temps de bien voir ce que fait Mathias en classe.»

Niveau 6 /Fav 3 : «Nous avons plus de temps pour parler avec le professeur qu'en rencontre individuelle.»

Conclusion

Plusieurs facteurs, qu'il est impossible de tous mentionner ici, ont certainement influencé les parents dans leur perception du bulletin. Insistons en premier lieu sur la présentation par l'élève de son nouveau bulletin ainsi que de son portfolio qui avait lieu au cours d'une rencontre tripartite. Le contexte a été jugé favorable, il a satisfait les parents et les a aidés à comprendre les nouveaux enjeux de l'évaluation Dans un deuxième temps, mentionnons «l'effet enseignants». C'est à eux que revenait la tâche de bien informer les parents sur les changements apportés à la démarche d'évaluation et sur la façon de noter les résultats au bulletin. Les enseignants participants au groupe de recherche étaient tous des volontaires qui appuyaient ce changement. Il paraît évident que leur attitude positive et l'expertise qu'ils ont su développer lors des rencontres de la recherche-formation ont joué un rôle dans le degré de satisfaction exprimé par les parents.

Des parents volontaires ont rempli un questionnaire plus détaillé qui leur permettait de se prononcer sur la clarté des compétences et des indicateurs. Nous avons tenu compte de leurs commentaires, toutes les informations consignées au bulletin ont été analysées par le chercheur et l'équipe des conseillers pédagogiques. Aussi, le libellé des compétences et les indicateurs ont été revus : des ajouts et des modifications ont été apportés dans toutes les disciplines.

Voyant ces résultats encourageants, il sera intéressant d'approfondir l'analyse dans une prochaine recherche, afin de comprendre les doutes qu'entretient une majorité de parents à l'endroit de ce nouveau bulletin malgré le taux élevé de satisfaction envers la présentation du bulletin et le bon niveau de compréhension qu'illustrent les réponses au questionnaire. En effet, l'explication de la réticence des parents aux changements et leurs croyances bien enracinées sur des pratiques

évaluatives qui privilégient la comparaison (moyenne de groupe) et la notation en pourcentage devraient être davantage documentés.

Liste des références bibliographiques

Allal, L., Wegmuller, E., Bonaiti-Dugerdil, S. & Cochet Kaeser, F. (1998).Le portfolio dans la dynamique de l'entretien tripartite. *Mesure et évaluation en éducation*, 20(3), 5-31.

Deborah, L. (2005) A case study of parents' and teachers' perceptions of the usefulness of a newly developed elementary school report card in the Appoquinimink School District Panchisin. *DAI-A, 65*/12, 4469

Deniger, M.-A.& Kamanzi, C. (2004). *L'évaluation du nouveau programme de formation de l'école québécoise : la qualité de sa mise en oeuvre et ses effets perçus à ce jour, (*Rapport de recherche du CRIRES), http://www.ulaval.ca/crires/html/pubgra.html

Desgagné, Bednarz, Couture, Lebuis et Poirier, 2001, L'approche collaborative de recherche en éducation : un rapport nouveau à établir entre recherche et formation. *Revue des sciences de l'éducation*, XXVII(1), 33-64

Durand, M-J., (2005) Les nouveaux bulletins scolaires : un souci de rigueur. *Le bulletin du CRIFPE*, 11 (1), 23-27

Greene, J.L. (1991) *Parent Opinions about Student evaluations in elementary Schools.*

Jorro, Anne (2000) *L'enseignant et l'évaluation : des gestes évaluatifs en question* Bruxelles : De Boeck Université, (1re éd.), 184 pages

Laurier, Michel D., Tousignant, R et Morissette, D. (2005*). Les principes de la mesure et de l'évaluation des apprentissages*, Éditions Gaétan Morin, Québec, (3e éd.,), 176 pages.

Louis, R. (1999). *L'évaluation des apprentissages en classe, théorie et pratique*. Québec: Éd.Études Vivantes, 212 pages.

Norman, K., Cartwright, M. (1999) Investigation of the use of portfolios as an assessment procedure.*DAI-A, 60*/02, 341

Newbury, W.D. (2006) Reporting approaches that help foster understanding between parents and educators. *MAI, 44*-03.

Scallon, G. (2004) *L'évaluation des apprentissages dans une approche par compétences*, Québec :ERPI :, 342 pages.

Simon, M. & Forgette-Giroux, R. (1994). Vers une utilisation rationnelle du dossier d'apprentissage. *Mesure et évaluation en éducation*, 16 (3 et 4), 27-40.

Simon, M. (2006) *Linking Classroom and Large-scale Student-level Assessment Reporting Mechanisms*. Article non publié, Université d'Ottawa, Ottawa, Canada

Xue, Y., Meisels, S.J., Bickel, D.D. & al. (2000, April) *An Analysis of parents' attitudes towards authentic performance assessment.* Paper presented at the Annual Meeting of the American educational Research Association.New Orleans, LA

Journal of **E**ducational **M**easurement and **A**pplied **C**ognitive **S**ciences
http://www.jemacs.uni.lu/ 2007 volume 1, No 1

PRATIQUES D'ÉVALUATION DANS LE CADRE D'UN FORUM : LE RETOUR DE L'ENSEIGNEMENT MUTUEL ?

Loïc Clavier, Maître de Conférences en Sciences de L'Éducation, IUFM des Pays de Loire. France.

Mots-clés : évaluation, forum, mutualisation

Résumé :

Cette contribution fait état d'une recherche exploratoire à l'IUFM des Pays de Loire (France) à propos d'un dispositif de Formation Ouverte et A Distance (FOAD) utilisant les TIC (Technologies de l'Information et de la Communication).

Ce travail scrute plus particulièrement la relation entre les pratiques d'évaluation telles qu'elles sont prévues par l'outil de formation à distance utilisé (une plateforme Web) et leurs influences sur les stratégies d'apprentissage développées par les étudiants. En effet, si les pratiques d'évaluation mises en œuvre par les formateurs sont dans le cadre de l'utilisation des TIC, assez communes, il n'en demeure pas moins qu'appliquées à un espace précis, de la formation (le forum) elles génèrent une activité étudiante atypique qui semble glisser du travail collaboratif à un travail de mutualisation prenant les aspects d'un enseignement mutuel.

C'est ce glissement que nous comptons investiguer en interrogeant la relation entre évaluation et formation.

Peut-on postuler que certaines formes d'évaluation utilisées dans des dispositifs de formation (FOAD) influencent, sous certaines conditions d'organisation, les pratiques d'enseignements ?

Quels sont dès lors les enjeux dans la pratique d'enseignement si l'on considère qu'évaluation et formation sont étroitement liées ?

L'utilisation des TIC dans une FOAD a généré chez les formateurs le souhait d'utiliser l'évaluation comme un facteur d'émancipation c'est-à-dire comme une opportunité pour les étudiants de produire du savoir. Il apparaît au terme de ce travail exploratoire que certains étudiants sont allés au-delà de ce qui était envisagé, investissant une forme d'enseignement oublié en France.

1) Problématique

Dans un dispositif de préparation au concours externe de Conseiller Principal d'Education en Formation Ouverte et A Distance, quelle peut être l'influence d'une évaluation dans les différents forums de l'activité des étudiants sur leurs stratégies d'apprentissage ?

Notre contribution participe à une réflexion sur les nouvelles pratiques de l'évaluation dans le sens où elles sont menées dans un cadre nouveau… Pour autant si le contexte évolue (FOAD), il convient aussi de regarder attentivement ce qui a changé dans la relation évaluation-formation…

2) Première approche de la recherche

2.1 L'évaluation

Parce que le dispositif envisagé est très spécifique (préparation à un concours de la fonction publique), il nous semble pertinent d'éclairer immédiatement les enjeux de l'évaluation au sein de cette préparation ainsi que les attentes des étudiants.

Tous les étudiants demandent à être évalués dans les conditions du concours.

Leur objectif : à chaque production être capable d'étalonner leur performance en fonction des exigences du jury. Ce qui semble déterminer une posture hypocoristique chez l'étudiant.

Cette attente qui s'exprime sous la forme d'une injonction va totalement contingenter la relation de l'étudiant à l'évaluation et à la formation. Du reste l'institution accorde à cette forme de préparation une importance toute particulière puisqu'elle répond favorablement à la demande des étudiants non seulement en favorisant cette culture de l'évaluation mais aussi en sollicitant des correcteurs (anciens ou actuels) du concours CPE pour corriger les productions des étudiants.

Bien au-delà d'un entraînement, le contrôle produit ici… du sens. Ce qui nous amène à envisager l'évaluation différemment qu'au travers du désormais classique distinguo entre le contrôle (attestant d'une conformité) et l'évaluation (produisant du sens) (Ardoino, 1976). Nous élargirons avec Jean Aubégny (2004) les couples paradigmatiques en question dans l'évaluation :

« - Evaluation formative / évaluation sommative,
- Evaluation centrée sur les processus / évaluation centrée sur les produits,
- Evaluation formaliste / évaluation naturaliste,
- Evaluation structurelle / évaluation conjoncturelle,
- Evaluation dite objective / évaluation dite subjective. »

Ainsi, nous ouvrons des pistes très variées quant aux enjeux de l'évaluation, aux attentes des étudiants et aux pratiques des enseignants. Cependant, la conformation par le contrôle ne satisfait pas pleinement les enseignants notamment pour sélectionner les candidats au redoublement.

Une précision s'impose ici. Lorsqu'il s'agit de postuler à un concours, l'étudiant n'envisage pas le redoublement de sa préparation. Dans le cas précis de la formation qui est utilisée comme terrain d'investigation, le contexte prend une importance capitale et permet de comprendre comment un étudiant peut être amené à s'investir sur deux ans (voire trois) pour enfin être lauréat.

Des chiffres tout d'abord : taux de réussite au niveau national 2,5 %. Taux de réussite dans cet IUFM 25 %. Taux de réussite des redoublants : 80 %.

Ces chiffres parlent d'eux-mêmes. Les étudiants le confirment d'ailleurs. Il y a très peu de chance de réussir le concours lors de la première présentation. La demande de redoublement est donc très importante.

Voilà donc la raison pour laquelle les enseignants ne souhaitent pas uniquement s'appuyer sur les notes obtenues aux épreuves « blanches » lors de la préparation pour laisser accéder tel ou tel candidat au redoublement. Ils souhaitent s'informer sur la capacité de progression des candidats.

Jusqu'à l'année 2004 la préparation se déroulait sur un modèle classique, simultané en face à face pédagogique. Depuis la rentrée 2005, la formation utilise une autre organisation : la formation ouverte et à distance.

Le dispositif s'organise en temps de formation en « présentiel » et en temps de formation à distance. Ces temps sont obligatoirement liés les uns aux autres et font l'objet d'un travail d'articulation pensé avant le démarrage de la formation par les enseignants. L'évaluation du travail à distance s'élabore grâce à un outil de suivi intégré à la plateforme web (DOKEOS). L'évaluation ainsi produite est d'abord un contrôle de l'activité de l'étudiant. On y référence son assiduité (nombre de documents consultés, téléchargés, nombre de liens activés pour suivre la « navigation documentaire », nombre de travaux rendus…) mais assez peu la qualité de son investissement.

Les formateurs ont donc ouvert dans leurs cours respectifs un forum dont l'objectif est d'amener les étudiants à échanger, à débattre autour des questions traitées en cours ou à distance. Pour évaluer ce travail, il a été convenu avec les étudiants que seules les contributions argumentées apportant des éléments mutualisables au sein de la préparation seraient prises en compte. Donc, chaque formateur au moment de clore le forum fait l'état des lieux des productions et attribue à chaque étudiant-producteur une évaluation.

Cette évaluation vise à cerner la capacité de l'étudiant à s'autonomiser, à se prendre en charge par rapport aux attentes du jury. Ainsi en faisant référence aux cinq couples de paradigmes évaluatifs cités plus haut et à des travaux précédemment menés (Clavier, 2006), on peut envisager que ce qui traverse chaque couple paradigme repose sur la capacité émancipatrice de l'évaluation.

Dans notre cas ce qui ressort des premières explorations tend vers :

- une recherche de conformité de la part de l'étudiant ;
- une recherche d'autonomie pour l'étudiant de la part de l'enseignant.

Comment dépasser alors ce qui pourrait être considéré comme des apories inhérentes à ces postures ?

L'évaluation est utilisée comme indicateur de réussite en regard de ces objectifs.

Cet écart de visée entre étudiants et formateurs va se cristalliser notamment au travers des forum avec un enjeu fort. Parce que les étudiants susceptibles de redoubler selon les critères classiques (notes obtenues aux épreuves « blanches », admissibilité au concours) sont plus nombreux que le nombre de places qui leur est réservé, ces derniers vont devoir se départager dans

l'évaluation des forums. Comment cet enjeu influence alors le travail des étudiants ?

2 .2 Collaboration ou mutualisation

Jusqu'au changement de dispositif intervenu en 2005, la préparation permettait aux étudiants, sous l'égide des formateurs, de proposer des exposés thématiques en lien avec les sujets du concours de sorte à partager avec l'ensemble du groupe soit des informations, soit des réflexions. Les enseignants encourageaient aussi le travail de groupe en dehors des temps de formation pour amener les étudiants à confronter leurs points de vue dans des débats argumentés. Il s'agissait donc de collecter des éléments, d'en assurer la mémoire mais aussi de les organiser d'un point de vue réflexif au travers de confrontations… Une sorte de formation par les pairs.

La mise en place de la FOAD et des outils inhérents à la plateforme ainsi que le souhait des enseignants d'évaluer l'activité en forum va modifier implicitement la nature du travail des étudiants.

Jusque-là les étudiants collaboraient entre eux… sous le regard de l'enseignant. Placer cette collaboration sous le regard de l'enseignant dans un forum peut changer le type de relation entre étudiants. Certains vont s'engager dans la production de savoir, faire valider par l'enseignant le travail réalisé, le mettre en ligne et s'investir dans la communication avec leurs condisciples. Au travers du forum, ils répondent en expert aux questions des autres étudiants. Parmi les documents mis en ligne, on trouve des études de cas qu'ils proposent à la sagacité de leurs collègues en assurant la « correction » de l'exercice.

On retrouve ici certaines caractéristiques du « lernen durch lehren » décrit par Martin (2004). Le forum permet une situation de communication authentique. Parmi les documents produits, on trouve des études de cas donnant lieu à une réflexion didactique pour la construction, de cet exercice. Il y a une construction collective de savoir, l'enseignant se portant garant à l'issue des débats des éléments échangés.

Cependant, il ne s'agit pas d'une démarche étendue à tous les étudiants, les enseignants ont passé commande de productions documentaires (synthèse, réflexion, étude de cas…) et n'ont pas délégué une quelconque activité didactique aux étudiants. De même n'ont-ils pas « entraîné » les étudiants à prendre en charge des activités didactiques.

On peut postuler à ce niveau de la réflexion que la visibilité des échanges sur le forum ainsi que l'enjeu évaluatif associé aux productions sollicités ont produit des situations dépassant le niveau de la collaboration entre étudiants pour tendre vers une mutualisation.

3) Méthodologie

Nous avons mené ce travail dans une optique exploratoire au vu de la nouveauté du dispositif et de sa structure fluctuante. C'est pourquoi, dans un premier temps, nous avons utilisé une enquête par questionnaire afin de déterminer :

- ce que les étudiants attendaient de l'évaluation menée à propos des épreuves « blanches » ;
- ce que le travail collaboratif à distance leur a apporté ;
- comment ils ont utilisé les forums et quel était le moteur de leur investissement ?

Ces questions sont principalement ouvertes. Le questionnaire a ensuite été complété par des entretiens avec les étudiants les plus investis dans l'animation des forum et avec des étudiants n'ayant pas utilisé cette fonctionnalité de la plateforme. Quinze entretiens ont été menés en fin d'année.

4) Résultats

Les étudiants sont unanimes (comme nous l'avions précédemment signalé) pour insister sur le caractère incontournable de l'évaluation des épreuves blanches dans les conditions du concours. Ils se sont inscrits à la préparation à l'IUFM pour cela. Ils souhaitent pouvoir étalonner sur l'ensemble des épreuves (12 dissertations et 12 études de cas) leurs performances en fonction des attentes du jury.

Interrogés à ce propos en entretien, les étudiants confirment qu'ils cherchent à identifier le profil du candidat « parfait » et à le reproduire ! La notation est utilisée dans une démarche syncrétique. Cette logique fait craindre que ces étudiants se fourvoient dans une certaine cavillation lors des épreuves écrites.

Les réponses sur le travail collaboratif sont beaucoup plus hétérogènes. Tout d'abord la grande majorité des étudiants précise avoir eu des difficultés à produire des travaux et surtout des études de cas. Face à ces difficultés ils ont soit abandonné les projets de production sur lesquels ils s'étaient engagés, soit travailler en commun pour optimiser leur investissement. Une fois les productions bouclées, ils identifient clairement les gains réalisés. Lors des entretiens ils ont insisté (pour ceux ayant travaillé ainsi) sur les bénéfices réalisés en termes de compréhension des thèmes abordés par le travail collaboratif.

Lorsque nous interrogeons les étudiants, il devient clair que ceux qui ont respecté les consignes, concernant la production de documents liés à des thèmes ou problématiques du concours, l'ont fait dans une double perspective : approfondir leurs connaissances et se ménager en cas d'échec au concours l'opportunité d'un redoublement.

Lorsqu'on passe à l'utilisation du forum, les réponses du questionnaire montrent une grande difficulté à l'utiliser. Une dizaine d'étudiants ont réellement alimenté les forum. Cela représente 25 % des étudiants de la préparation. Les difficultés sont liées à la capacité à communiquer. D'après certaines réponses collectées au niveau du questionnaire, ils ne se sentaient pas capables de débattre grâce à cet outil. Les entretiens réalisés avec les étudiants dans ce cas de figure confirment cette difficulté à communiquer. Malgré l'enjeu évaluatif, ils n'ont pas cherché à dépasser cet obstacle. Ils l'ont contourné en participant plus activement à l'oral lors des séances en « présentiel ».

Face à ces éléments, les quelques étudiants qui ont fait vivre ces forums ont une position différente. Ils ont tous été attirés par l'activité proposée (production de savoir et débat en forum) trouvant la situation de communication intéressante. Mais parmi cette population on retrouve aussi une certaine hétérogénéité. Tous n'ont pas eu le même rendement en forum ou en production de documents. Ils se sont astreints à ces travaux parce qu'ils en ressentaient les effets sur leur progression en formation (retour de notation en épreuves blanches) et pour se ménager une possibilité de redoublement en cas d'échec au concours.

Ces éléments nous permettent de dresser une typologie provisoire des étudiants dans cette préparation.

Près de 75 % d'entre eux sont arrivés avec des attentes précises quant à la fonction de l'évaluation de leur travail. Si nous reprenons les paradigmes évoqués précédemment, un premier groupe d'étudiants attend une évaluation sommative, sur les produits, formaliste, objective et structurelle. Ils tentent de reproduire « un candidat idéal » par syncrétisme. La collaboration même entre étudiants ne semble pas les attirer. Ils sont dans une logique cumulative.

Un second groupe d'étudiants, conscients des enjeux de l'évaluation que ce soit en termes d'étalonnage des performances ou en termes de perspectives d'accès au concours par un éventuel redoublement, va bâtir des stratégies d'apprentissage en fonction de ce que les étudiants sont en mesure d'accepter des pratiques d'évaluation des formateurs.

Schématiquement, ils vont par exemple contourner une faible activité sur les forum par une participation accrue en présentiel. Il s'agit d'arrangements individuels et locaux qui induisent des comportements visant à l'autonomisation dans les stratégies d'apprentissage. Les étudiants ont compris que le candidat idéal n'existe pas et ils utilisent l'évaluation des enseignants pour se bâtir un profil de lauréat sans pour autant saisir la totalité des outils mis à leur disposition. Les redoublants, de l'année précédent le démarrage de la FOAD, sont dans cas de figure.

Enfin, un dernier groupe d'une dizaine d'étudiants va se saisir des ressources du dispositif pour s'émanciper du modèle idéal du lauréat. Cette

émancipation se joue dans leur capacité à trouver de l'autonomie entre les couples paradigmatiques à l'œuvre dans les pratiques d'évaluation. Leur activité en forum va dans ce sens. Leurs stratégies d'apprentissage les conduisent à une autonomie telle qu'ils dépassent le simple travail de collaboration. Cette dynamique s'obtient aussi parce qu'ils se projettent en fonction de leur connaissance du concours dans une perspective d'un éventuel redoublement.

Toutefois, nous ne voudrions pas conclure hâtivement en pointant que pour une minorité d'étudiants une démarche d'enseignement mutuel émerge d'un tel dispositif.

Lors de cette démarche exploratoire, nous n'avons pas questionné les caractéristiques liées à l'histoire personnelle des étudiants, à leur cursus universitaire. Ces éléments peuvent agir aussi dans leur capacité à prendre en charge des activités mutualisantes dans un forum et ce, en dehors des pratiques d'évaluation des formateurs.

Nous avons présenté notre recherche sur un dispositif émergeant dans lequel aucun enseignant n'a l'expérience de la FOAD et qui fonctionnait très différemment les années précédentes. Le contexte reste à investiguer afin d'identifier ces influences sur les stratégies de quelques étudiants.

Il nous reste aussi à questionner la notion de forum dans ce type de préparation. Ne révèle-t-il pas simplement des échanges entre étudiants qui existaient déjà dans le précédent dispositif mais qui restaient inaccessibles aux enseignants.

On le voit cette recherche est dans sa phase exploratoire. Il reste à approfondir certaines directions de recherche avant de pouvoir se prononcer sur une relation entre évaluation et émergence d'une forme d'enseignement mutuel en FOAD.

Références

ARDOINO, J. (1976). *Au filigrane d'un discours : la question du contrôle et de l'évaluation*. Préface à Morin (M), *L'imaginaire dans l'éducation permanente* Paris : Gauthier-Villars.

AUBEGNY, J. (2004). *L'évaluation : un modèle de lecture et d'aide à la décision* in Aubégny J. &Clavier L., *L'Évaluation en IUFM* , 33-52, Paris : L'Harmattan

CLAVIER, L. (2003). *L'évaluation au service de l'émancipation*, 18ème Colloque ADMEE Liège 2003

MARTIN, J. P. (2004) *La méthode Lernen durch lehnen : quand les apprenants font la classe* paru dans les cahiers de l'APLIUT. Actes du

XXVème congrès *A la recherche de situations communicatives authentiques : l'apprentissage des langues par les tâches.*

EVALUATION DE L'INTÉRÊT POUR LES BRANCHES SCIENTIFIQUES PRÉSENTÉES DANS UN CONTEXTE MASCULIN ET DANS UN CONTEXTE FÉMININ : CONTENUS NON-PARALLÉLISÉES ET CONTENUS PARALLÉLISÉES

Kerger Sylvie, Université du Luxembourg

Mots clés : contenus parallélisés, intérêts scolaires, branches scientifiques

Résumé :

Cette étude essaye d'évaluer l'intérêt pour les branches scientifiques. Afin de rendre les branches plus intéressantes pour les filles, nous nous sommes référée à des recherches utilisant le contexte de présentation masculin et féminin afin d'adapter les branches scientifiques aux intérêts des filles (Todt, 2000). Dans notre recherche, nous avons fait deux enquêtes qui mesurent chacune les intérêts pour les branches scientifiques. La différence entre les deux enquêtes est celle que dans la première enquête les branches, présentées dans un contexte masculin et dans un contexte féminin, n'englobaient pas les mêmes contenus. Tandis que dans l'enquête 2, les mêmes contenus des branches ont été présentés dans un contexte masculin et dans un contexte féminin. En comparant les deux enquêtes, les résultats montrent que les contenus n'avaient pas une grande influence, mais que le contexte de présentation jouait un rôle très important dans l'évaluation des intérêts pour les branches scientifiques. Dans les deux enquêtes, les filles s'intéressent davantage au contexte féminin des branches et les garçons s'intéressent davantage au contexte masculin des branches. Dans notre évaluation, les deux enquêtes montrent des résultats similaires. Les contextes de présentation des branches scientifiques ont une influence marquée sur l'intérêt des sujets et cette influence est similaire dans l'enquête avec les contenus parallélisés et dans l'enquête avec les contenus non-parallélisés.

Problématique

Le but de cette recherche est de trouver des moyens de rapprocher les filles des sciences, afin d'inciter les jeunes filles à s'engager davantage dans les carrières scientifiques.

Au lieu d'essayer de changer les intérêts des sujets, nous avons employé une méthode pour adapter les branches scientifiques aux intérêts des sujets. Nous nous sommes appuyée sur des recherches qui ont essayé de remédier aux différences des intérêts en essayant de rapprocher les contenus des branches scientifiques davantage aux intérêts des filles, en présentant les branches scientifiques dans un contexte féminin (Häuβler & Hoffmann, 1995; Hoffmann & Häuβler, 1993; Todt, 2000; Todt & Händel, 1988).

Enquête 1 : les branches scientifiques avec contenus non-parallélisées

Les branches, notamment la biologie, la physique, l'informatique et les sciences statistiques ont été présentées dans un contexte masculin (10 items) et dans un contexte féminin (10 items).

Tableau 1 : Exemples d'items avec les contextes non-parallélisés :

Biologie contexte féminin : Examiner quelles parties de la langue humaine sont les plus sensibles à différents goûts (sucré, amer).
Biologie contexte masculin : Apprendre de quelles parties le système nerveux humain est constitué.

Résultats de l'étude avec les branches non-parallélisées

Les filles s'intéressent significativement plus à la biologie dans un contexte féminin et aux sciences statistiques dans un contexte féminin. L'intérêt des garçons est significativement plus élevé pour la physique dans un contexte masculin, l'informatique dans un contexte masculin et pour les sciences statistiques dans un contexte masculin. Pas de différences sont à noter pour la biologie dans un contexte masculin et pour la physique et l'informatique dans un contexte féminin.

Tableau 2 : Analyse de la variance selon le sexe pour l'intérêt les branches (enquête 1, n=295).

variables	**moyenne (écart-type) des garçons**	**moyenne (écart-type) des filles**	**significativité**	**eta²**
Biologie, f	2.46 (0.92)	**3.14** (0.92)	++	.119
Biologie, m	2.62 (1.02)	2.53 (1.04)	/	.002
Physique, f	2.51 (0.95)	2.60 (0.94)	/	.002
Physique, m	**3.28** (1.15)	2.09 (0.99)	++	.233
Informatique, f	3.20 (1.81)	3.03 (0.79)	/	.007
Informatique, m	**3.86** (1.06)	2.23 (1.04)	++	.374
Sciences statistiques, f	2.48 (0.94)	**3.13** (1.01)	++	.101
Sciences statistiques, m	**2.73** (1.03)	1.99 (0.77)	++	.138

m=contexte masculin, f=contexte féminin
échelle : 1 = pas du tout intéressant à 5 = très intéressant
significativité : ++ = p<.010 ; / = non significatif
Pillai's Trace = 10.277 ; p<.001

Dans cette enquête nous avons présenté les branches dans deux contextes différents, notamment le contexte masculin et le contexte féminin, mais nous n'avons

pas présenté les mêmes contenus dans les deux contextes. Les résultats obtenus pourraient donc être remis en question en raison du choix des items. A cet effet, nous avons élaboré dans l'enquête 2 des contenus parallélisés.

Enquête 2 : les branches scientifiques avec contenus parallélisées

Nous avons voulu vérifier, si les branches présentées lors de l'enquête 1, mesuraient bien le contexte féminin et le contexte masculin. Les items n'étaient pas parallélisés entre les contextes féminins et les contextes masculins, c'est-à-dire que les contenus d'une branche n'étaient pas tout à fait les mêmes pour les deux contextes de présentation.

Afin de mieux pouvoir comparer les deux contextes de présentation, nous avons élaboré des items parallélisés. Les branches, notamment la biologie, la physique, l'informatique et les sciences statistiques ont été présentées dans un contexte masculin (5 items) et dans un contexte féminin (5 items). Selon notre hypothèse, les items parallélisés montrent les mêmes résultats que les items non parallélisés.

Tableau 3 : Exemples d'items avec les contextes parallélisés :

Physique contexte féminin : Savoir plus sur le fonctionnement du cœur artificiel.
Physique contexte masculin : Savoir plus sur l'extraction du pétrole de grandes profondeurs (3000 m).

Comparaison entre les résultats de l'enquête 1 et de l'enquête 2

Noue avons élaboré l'hypothèse que les branches avec les items parallélisés montrent les mêmes effets que les branches de l'enquête 1, dans laquelle les items entre les branches n'étaient pas parallélisés.

Ces échelles comprennent des items pour lesquelles seulement le contexte de présentation change et non les contenus des branches proposés aux élèves. Nous avons donc parallélisé les items des branches entre le contexte de présentation masculin et le contexte de présentation féminin.

Dans le tableau suivant, les moyennes et les résultats de l'analyse de la variance sont présentés. Pour l'intérêt des branches scientifiques, les différences dues au sexe sont généralement plus élevées dans cette étude que dans l'enquête 1.

Tableau 4 : Analyse de la variance selon le sexe pour l'enquête 2 (enquête 2, n=190) et pour l'enquête 1 (enquête 1, n=295).

variables	moyenne (écart-type) des garçons	moyenne (écart-type) des filles	significativité enquête 2 (enquête1)	eta²
Biologie, f, p	2.62 (0.82)	**2.86** (0.81)	+ (++)	.021 (.119)
Biologie, m, p	2.71 (1.01)	2.85 (0.98)	/ (/)	.005 (.002)
Physique, f, p	2.69 (0.99)	**3.17** (1.07)	++ (/)	.052 (.002)
Physique, m, p	**3.33** (0.91)	2.17 (0.71)	++ (++)	.326 (.233)
Informatique, f, p	**3.41** (0.94)	3.03 (0.97)	++ (/)	.038 (.007)
Informatique, m, p	**3.89** (1.04)	2.23 (1.05)	++ (++)	.385 (.374)
Sciences statistiques, f, p	2.68 (0.98)	**3.75** (0.92)	++ (++)	.237 (.101)
Sciences statistiques, m, p	**2.98** (0.92)	2.19 (0.89)	++ (++)	.158 (.138)

m=contexte masculin, f=contexte féminin
p= contenus parallélisées
échelle : 1 = pas du tout intéressant à 5 = très intéressant
significativité : ++ = p<.010 ; + = <.050 ; / = non significatif
Pillai's Trace = 10.177 ; p<.001
Pour la significativité et l'eta², les **valeurs entre parenthèses** sont les résultats de l'enquête 1, dans laquelle les contenus des branches n'étaient pas parallélisés.

Dans l'analyse de la variance selon le sexe pour l'intérêt aux branches, nous avons trouvé les mêmes différences entre les sexes, sauf pour la physique et pour l'informatique dans un contexte féminin. Pour ces deux branches, nous avons noté aucune différence de sexe dans l'enquête 1, mais dans cette enquête, nous avons constaté que les filles sont davantage intéressées à la physique dans un contexte féminin que les garçons, tandis que les garçons sont plus intéressés à l'informatique dans un contexte féminin que les filles.

Ces résultats montrent donc, que la présentation des mêmes contenus dans un contexte féminin peut avoir une influence positive sur les intérêts des filles. Tandis que les contextes féminins des branches trouvent généralement plus d'intérêts chez les filles, les contextes masculins trouvent généralement plus d'intérêts chez les garçons. La présentation des branches dans les contextes masculins et féminins ont pratiquement les mêmes effets dans l'enquête 1 que dans l'enquête 2.

Conclusion

Ces résultats indiquent que les principales tendances sont similaires entre les échelles non-parallélisées et les échelles parallélisées. Donc le fait de présenter les branches scientifiques avec des contenus parallélisés ou non-parallélisés montrent généralement les mêmes intérêts des filles et des garçons aux branches scientifiques. Ceci nous mène à la conclusion que la mesure des intérêts pour les branches scientifiques est surtout dépendante du contexte de présentation masculin et féminin et non du fait que les contenus sont parallélisés ou non. L'intérêt des filles pour les branches scientifiques augmente en prenant en compte leurs intérêts et en présentant les branches scientifiques dans un contexte féminin.

Références

Häuβler, P., & Hoffmann, L. (1995). Physikunterricht- an den Interessen von Mädchen und Jungen orientiert. *Unterrichtswissenschaft, 23*, 107-127.

Hoffmann, L., & Häuβler, P. (1993). *Der blk-modellversuch: Chancengleichheit-Veränderung des Anfangsunterrichts Physik/Chemie unter besonderer Berücksichtigung der Kompetenzen und Interessen von Mädchen: Bericht und erste Ergebnisse.*: Tischvorlage für die Sitzung des Sachverständigenrats des IPN. Kiel: IPN.

Todt, E. (2000). Geschlechtsspezifische Interessenentwicklung und Möglichkeiten der Modifikation. *Zeitschrift zur Theorie und Praxis erziehungswissenschaftlicher Forschung, 14*, 215-254.

Todt, E., & Händel, B. (1988). Kontextabhängigkeit von Physikinteressen. *Der mathematische und naturwissenschaftliche Unterricht, 41*, 137-140.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

POSITIONNEMENT DES PARENTS LUXEMBOURGEOIS FACE A PLUSIEURS SCENARIOS D'EVOLUTION POSSIBLE DU SYSTEME EDUCATIF

Dierendonck, C. - Université du Luxembourg
Poncelet, D. - Université du Luxembourg
Martin, R. - Université du Luxembourg

Mots-clés : Ecole de demain, évolution du système éducatif, Luxembourg, parents, lien école-familles

Résumé :

A l'instar des enseignants et des élèves, les parents constituent des acteurs incontournables du processus éducatif. L'école ne peut en effet être considérée comme l'unique facteur influençant la réussite scolaire et professionnelle des individus. Le rôle de l'environnement familial joue, au contraire, un rôle crucial dans le développement et la socialisation de l'enfant. Il s'agit donc d'apporter une contribution à la compréhension du processus éducatif à l'intérieur des unités familiales et de comprendre la nature des relations que les parents entretiennent avec l'école.

La prise d'informations a été réalisée dans le cadre de la recherche « *La place de l'école dans la société luxembourgeoise de demain* » auprès d'un échantillon de 120 classes sélectionnées de manière aléatoire à trois niveaux de la scolarité (5e primaire, 1ère année du secondaire et 3e année du secondaire). Les enseignants, les élèves et les parents d'élèves de ces classes ont été interrogés par questionnaire.

L'étude auprès des familles poursuivait un triple objectif. Il s'agissait tout d'abord d'offrir aux parents l'occasion de s'exprimer sur le fonctionnement actuel et futur du système éducatif luxembourgeois (Partie 1). Il s'agissait ensuite pour l'équipe de recherche de mieux comprendre le fonctionnement des sphères familiales en s'intéressant particulièrement aux pratiques éducatives mises en place dans les foyers (Partie 2). Il s'agissait enfin de cerner la nature du lien qui existe entre les familles et le système scolaire (Partie 3).

Notre communication se focalise sur le positionnement des parents par rapport aux scénarios d'évolution possible du système luxembourgeois (partie 1).

Présentation

Les données de cette communication s'inscrivent dans le cadre plus large de la consultation des intervenants du système éducatif luxembourgeois au centre de la

recherche « *La place de l'école dans la société luxembourgeoise de demain* ». Nous présentons ici le positionnement de l'échantillon de parents par rapport à une série de propositions de réforme de l'école luxembourgeoise en le mettant en perspective avec les prises de position de nos échantillons d'enseignants, de directions d'écoles et d'inspecteurs.

Les propositions de réforme peuvent être regroupées en deux catégories : les propositions de la première catégorie sont d'ordre structurel ou pédagogique tandis que les propositions de la deuxième catégorie[1] se focalisent sur la relation entre l'école et les familles.

1. Positionnement des acteurs par rapport aux propositions d'ordre structurel ou pédagogique

Les propositions d'ordre structurel ou pédagogique ont en partie été définies dans le cadre de la consultation par *focus groups*[2] à partir de l'analyse de l'état actuel du système éducatif luxembourgeois et de ce qui se fait dans d'autres systèmes éducatifs. Nous avons choisi de reprendre ces mêmes scénarios et de les soumettre à nos échantillons de parents, d'enseignants, de directions d'école[3] et de membres de l'Inspection[4.]

Il faut cependant souligner que seuls les enseignants, les directions et les membres de l'Inspection ont reçu le document complet qui accompagnait les différentes propositions de réforme. Les parents n'ont pas reçu ce document argumentatif et ont été invités à se positionner à partir des seuls énoncés de scénarios.

Lors de la rédaction des scénarios, l'équipe de recherche a distingué trois axes potentiels de changement :

- un axe visant la gestion de l'hétérogénéité des élèves ;
- un axe centré sur la redéfinition du métier d'enseignant et du fonctionnement des écoles ;
- un axe focalisé sur la mise en place d'un dispositif d'assurance-qualité de l'école.

1.1. Vers une gestion de l'hétérogénéité des élèves

Au Luxembourg, les taux de retard scolaire sont particulièrement élevés. A l'entrée de la sixième année primaire, les statistiques du Ministère (année 2002-2003) révèlent que près de 20 % des élèves ont déjà redoublé au minimum une année scolaire. Ces statistiques montrent également que la proportion d'élèves d'origine

---

1 Dans cette catégorie, on retrouve deux scénarios à propos desquels tous les échantillons (parents, enseignants, inspecteurs, directions) ont été interrogés. Les autres propositions n'ont été soumises qu'aux parents et aux enseignants.

2 Voir à ce sujet la communication de Meyers, C.

3 Treize directions sur 26 ont accepté de participer à cette étude.

4 Neuf inspecteurs/trices sur 18 ont accepté de participer à cette étude.

étrangère ayant redoublé (25 à 30 %) est nettement supérieure à celle des élèves d'origine luxembourgeoise (13 à 15 %). Ventilés selon l'arrière-fond d'immigration ou la nationalité des élèves, les chiffres débouchent sur un état des lieux alarmant puisqu'on constate qu'un Portugais sur trois, un ex-Yougoslave sur trois, un Français sur cinq, un Italien sur quatre, un Capverdien sur deux et un Luxembourgeois sur sept sont en situation de retard scolaire.

On peut évidemment faire l'hypothèse que cette hétérogénéité observée au niveau des résultats scolaires s'explique en partie par une autre hétérogénéité se marquant à plusieurs niveaux :

- aux niveaux socio-économique et socio-culturel où on relève l'existence de familles défavorisées ou favorisées avec des arrière-fonds socio-culturels différents ;
- au niveau linguistique où on constate chez les élèves l'existence d'autres langues dominantes à côté des 3 langues officielles du pays ;
- au niveau des besoins spécifiques de certains élèves pour qui l'enseignement ordinaire semble être inadéquat.

Pour combattre cette hétérogénéité, nous avons formulé plusieurs scénarios de réformes potentielles. Voici comment nos échantillons se sont positionnés par rapport à chacun d'entre eux.

1.1.1. Supprimer le redoublement

Faut-il ou non faire redoubler les élèves ? Telle est la question qui suscite depuis plusieurs décennies un débat animé entre chercheurs et enseignants. Dans les deux camps, on trouve à la fois des partisans du redoublement et des adeptes d'un système éducatif basé sur la promotion automatique des élèves mais on remarque surtout que le débat a évolué différemment selon les pays considérés et selon les cultures d'évaluation. Ainsi, en définissant le redoublement comme le fait pour un élève de devoir recommencer une année scolaire suite à un constat d'échec à une ou plusieurs évaluations sommatives, on peut définir trois types de systèmes éducatifs : ceux qui font redoubler les élèves (groupe A), ceux qui organisent un tronc commun à l'intérieur duquel le redoublement est aboli (groupe B) et ceux qui présentent une configuration mixte (groupe C).

La majorité des systèmes éducatifs peuvent être classés dans le groupe A. On y retrouve la plupart des pays en voie de développement ainsi que des pays comme les Etats-Unis, le Canada, la Corée du Sud, l'Allemagne, les Pays-Bas, l'Italie ou le Luxembourg. Le redoublement y est considéré comme une mesure efficace. Le groupe B rassemble essentiellement les pays scandinaves ainsi que le Royaume-Uni. La promotion automatique des élèves y est de rigueur. Les élèves ne connaissent pas le redoublement tout au long de leur scolarité de base (jusque 15 ans). Dans le groupe C, on trouve des pays comme la Suisse, la Belgique, la France ou l'Espagne dans lesquels le redoublement est possible mais uniquement au terme d'étapes ou de cycles bien définis.

Cette diversité dans le fonctionnement des systèmes éducatifs montre à quel point la question du redoublement est centrale dans la façon dont les uns et les autres luttent contre l'échec scolaire et les résultats des dernières enquêtes internationales (PISA 2000 et 2003) sur le niveau scolaire des élèves laissent apparaître que les structures adoptées dans les pays du groupe B sont relativement plus efficaces.

Enfin, *last but not least*, la quasi-totalité des études menées sur l'efficacité du redoublement aboutissent à la conclusion d'un impact nul de la mesure sur les performances académiques et sur un impact négatif au niveau des variables socio-affectives.

Dès lors, en instaurant une promotion automatique des élèves accompagnée d'un système efficace de gestion de l'hétérogénéité des élèves à l'intérieur d'un même groupe d'âge[5], on pourrait permettre à tous les élèves de progresser à leur rythme tout en restant intégrés dans un groupe social de référence qui correspond à leur classe d'âge.

Tableau 1 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Supprimer le redoublement à l'école primaire"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1140	46,6	30	8,9	9,6	4,9	-
Inspecteurs	9	11,1	22,2	22,2	33,3	11,1	-
Directions	13	-	15,4	7,7	61,5	15,4	-
Enseignants	108	23,1	23,1	22,2	19,4	8,3	3,7
Enseignants du primaire	45	20	20	22,2	**24,4**	13,3	-
Enseignants du secondaire	63	25,4	25,4	22,2	15,9	4,8	6,3

5 Par exemple, abolir le système de notation basé sur des niveaux de compétence "normés" à partir de programmes préétablis et le remplacer par une évaluation formative des élèves basée sur leurs progrès individuels. On peut également souhaiter une collaboration plus étroite entre les enseignants auxquels on adjoint du personnel psychopédagogique supplémentaire.

Tableau 2 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Supprimer le redoublement à l'école secondaire"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1136	48,2	32	9,6	8,3	1,9	-
Inspecteurs	9	11,1	33,3	55,6	-	-	-
Directions	13	7,7	7,7	30,8	38,5	15,4	-
Enseignants	108	35,2	28,7	17,6	10,2	6,5	1,9
Enseignants du primaire	45	28,9	20	**28,9**	8,9	8,9	4,4
Enseignants du secondaire	63	39,7	34,9	9,5	11,1	4,8	-

Au vu des tableaux 1 et 2 qui portent respectivement sur la scolarité primaire et la scolarité secondaire, parents et enseignants ne semblent pas tout à fait réjouis par une réforme qui abolirait le redoublement au Luxembourg. En effet, quel que soit le niveau d'enseignement envisagé, on relève que 75 à 80 % des parents et 46 à 64 % des enseignants s'y opposent. On signalera toutefois que 38 % des enseignants du primaire sont favorables à la proposition de supprimer le redoublement dans l'enseignement primaire (contre 40 % de désaccords).

Les membres de l'Inspection (n=9) sont partagés quant à la suppression du redoublement au primaire (3 désaccords, 2 positions mitigées, 4 accords) et plutôt défavorables à sa disparition au secondaire (4 désaccords, 5 positions mitigées).

Au niveau des directions d'école (n=13), il semble qu'une majorité se dégage en faveur de la proposition (10 accords pour la suppression du redoublement au primaire et 7 accords pour le secondaire).

1.1.2. Retarder le moment de l'orientation vers les filières d'enseignement (classique, technique et professionnel) en instaurant un tronc commun jusqu'à 15 ans (même formation de base pour tous) et des modules de formation permettant une individualisation ponctuelle des apprentissages

Le système éducatif luxembourgeois figure parmi les systèmes éducatifs qui orientent le plus précocement ses élèves lors de la procédure dite du « passage primaire-post-primaire » en fin de sixième année primaire. L'orientation des élèves s'opère vers différentes filières hiérarchisées, drainant un nombre considérable d'élèves en difficulté à la fin du primaire vers un enseignement allégé censé remédier ou simplement tenir compte des « lacunes » existantes. Malheureusement, cette

stratégie d'homogénéisation des publics scolaires ne semble pas porter ses fruits. Au contraire, les taux d'abandon et de retard scolaire ne font qu'augmenter au fil de l'enseignement secondaire. Les statistiques sur les notes insuffisantes montrent clairement que les difficultés des élèves persistent ou s'aggravent dès le début du secondaire.

En retardant le moment de la première orientation par la création d'un tronc commun jusque l'âge de 15 ans, il s'agit donc pour le système éducatif de viser un objectif ambitieux : permettre à TOUS les élèves à l'âge de 15 ans de maîtriser un socle commun de compétences tout en offrant à chaque enfant la possibilité de développer un profil hautement individualisé.

Selon le tableau 3, 52 % des parents et 39 % des enseignants consultés se montrent plutôt favorables à cette proposition de réforme. Ils sont respectivement 33 % chez les parents et 38 % chez les enseignants à s'opposer à ce changement potentiel. Visiblement, ce sont les enseignants qui sont le plus partagés sur la pertinence d'un tel scénario d'évolution. Et lorsqu'on ventile leurs réponses selon le niveau d'enseignement, on met en lumière un positionnement tout à fait différent entre, d'un côté, les enseignants du primaire (49 % d'accords, 24 % de désaccords) et, de l'autre côté, les enseignants du secondaire (33 % d'accords, 48 % de désaccords).

Si les membres de l'Inspection semblent assez partagés sur la question (3 désaccords, 2 positions mitigées et 4 accords), on constate par contre que les directions d'établissement ont marqué massivement leur accord avec la proposition (2 désaccords, 11 accords).

Tableau 3 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Retarder le moment de l'orientation vers les filières d'enseignement (classique, technique et professionnel) en instaurant un tronc commun jusqu'à 15 ans (même formation de base pour tous) et des modules de formation permettant une individualisation ponctuelle des apprentissages"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1132	14,5	18,6	14,8	33	19,2	-
Inspecteurs	9	22,2	11,1	22,2	11,1	33,3	-
Directions	13	15,4	-	-	38,5	46,2	-
Enseignants	108	21,3	16,7	17,6	24,1	15,7	4,6
Enseignants du primaire	45	11,1	13,3	20	26,7	22,2	6,7
Enseignants du secondaire	63	28,6	19	15,9	22,2	11,1	3,2

1.1.3. Laisser aux parents le choix de l'établissement scolaire de leur enfant

Depuis la loi scolaire du 10 août 1912, les enfants en âge de scolarisation doivent fréquenter l'école communale établie dans le ressort scolaire de la résidence du père, de la mère, du tuteur ou de toute autre personne ayant la garde de l'enfant. Il en va de même au début de l'enseignement secondaire où l'élève est inscrit en priorité dans un lycée situé dans la zone de proximité de sa commune de résidence[6]. Dans l'enseignement secondaire supérieur et aux cycles moyen et supérieur de l'enseignement secondaire technique, les élèves s'inscrivent en fonction des formations offertes par les lycées et de leurs capacités d'accueil.

La répartition des élèves dans les écoles luxembourgeoises s'opère par sectorisation jusqu'à l'âge d'environ 15 ans. Nous avons voulu remettre en question cette pratique en proposant aux parents et aux enseignants de se positionner par rapport à un libre choix de l'établissement fréquenté. Le tableau 4 synthétise leurs réponses.

Tableau 4 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Laisser aux parents le choix de l'établissement scolaire de leur enfant"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1136	3,9	13,3	8	39,9	34,9	-
Inspecteurs	9	11,1	11,1	33,3	-	44,4	-
Directions	13	-	7,7	23,1	30,8	38,5	-
Enseignants	108	12	13,9	16,7	27,8	25,9	3,7
Enseignants du primaire	45	22,2	26,7	17,8	17,8	8,9	6,7
Enseignants du secondaire	63	4,8	4,8	15,9	34,9	38,1	1,6

A la lecture de ce tableau, on peut dresser un constat fort : près de trois quarts des parents et plus de la moitié des enseignants (54 %) semblent favorables à l'abandon de la carte scolaire, ce qui pose question sur les raisons d'une telle prise de position. On notera encore une fois que le positionnement est différent selon que l'on se situe au primaire ou au secondaire puisque qu'ils sont 49 % à être en désaccord avec la proposition au primaire et seulement 10 % au secondaire.

[6] Les zones de proximité sont définies par règlement grand-ducal. A la demande de l'élève, l'inscription peut se faire dans un autre lycée si les capacités d'accueil de ce lycée le permettent.

Une majorité semble également se dégager en faveur de la proposition au niveau des membres de l'Inspection (2 désaccords, 3 positions mitigées et 4 accords) et des directions (1 seul désaccord).

1.1.4. Laisser aux élèves et aux parents le choix de la langue d'alphabétisation à l'école primaire (allemand ou français)

Avec 40 % d'élèves étrangers, dont une grande partie de romanophones, on pourrait également proposer que le système éducatif luxembourgeois laisse aux élèves et à leurs parents le choix de la langue d'alphabétisation (allemand ou français) des jeunes qui deviendrait langue dominante pour la suite de leur parcours scolaire, avec une deuxième langue (français ou allemand) qui prendrait le statut de langue étrangère et par rapport à laquelle les exigences de maîtrise formelle seraient moins élevées.

La réaction des parents et des enseignants à cette proposition est assez nette au vu du tableau 5. Ils ne sont en effet que 38 % chez les parents et 30 % chez les enseignants à trouver le scénario potentiellement positif.

Au niveau de l'Inspection, on relève un positionnement plutôt négatif (6 désaccords, 2 positions mitigées et 1 accord) tandis qu'au niveau des directions d'école la question est moins tranchée (8 désaccords, 5 accords).

Tableau 5 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Laisser aux élèves et aux parents le choix de la langue d'alphabétisation à l'école primaire (allemand ou français)"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1133	22,7	24,9	13,9	24,7	13,8	-
Inspecteurs	9	44,4	22,2	22,2	-	11,1	-
Directions	13	38,5	23,1	-	30,8	7,7	-
Enseignants	108	24,1	24,1	18,5	19,4	10,2	3,7
Enseignants du primaire	45	26,7	24,4	20	17,8	11,1	-
Enseignants du secondaire	63	22,2	23,8	17,5	20,6	9,5	6,3

1.1.5. Supprimer les devoirs à domicile

Pour les parents, le moment des devoirs à domicile constitue dans le meilleur des cas une occasion de suivre ce que l'enfant fait à l'école et de communiquer à ce sujet. Pour d'autres parents par contre, c'est un moment redouté (enfant fatigué, exercices trop difficiles, …), manqué (parents qui travaillent le soir, parents occupés) voire évité (parents dépassés).

Pour les enseignants, c'est trop souvent un moyen de rattraper ou d'avancer dans la matière, de finir les exercices qui n'ont pas pu être terminés en classe.

Ce qui est certain, c'est qu'à côté du facteur socio-économique bien connu, l'encadrement des devoirs à domicile par les parents constitue un facteur de discrimination supplémentaire, aggravant parfois les différences déjà importantes entre les élèves issus de milieu socio-économique défavorisé et les élèves plus favorisés. C'est pour cette raison que nous avons proposé aux parents et aux enseignants de se positionner par rapport à la suppression éventuelle des devoirs à domicile.

Tableau 6 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Supprimer les devoirs à domicile"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1140	43,5	38,9	4,6	8,9	4	-
Inspecteurs	9	22,2	44,4	11,1	11,1	11,1	-
Directions	13	53,8	23,1	23,1	-	-	-
Enseignants	107	63,6	25,2	7,5	-	3,7	-
Enseignants du primaire	44	54,5	29,5	11,4	-	4,5	-
Enseignants du secondaire	63	69,8	22,2	4,8	-	3,2	-

Les chiffres du tableau 6 sont des plus clairs : parents, inspecteurs, directions et enseignants ne sont pas prêts à abandonner les devoirs à domicile. Seuls 13 % des parents et 4 % des enseignants y sont favorables.

1.1.6. Organiser un soutien spécifique aux élèves d'origine étrangère en proposant un soutien systématique à la langue maternelle (ex : des cours de portugais pour les élèves d'origine portugaise)

La scolarisation des enfants étrangers constitue sans aucun doute un des problèmes majeurs auquel doivent faire face nombre d'enseignants, surtout au primaire. Or on sait que l'apprentissage d'une langue étrangère se déroule dans de meilleures conditions si la langue maternelle est elle-même bien maîtrisée. Il nous est dès lors apparu intéressant d'envisager, en complément des mesures classiques de soutien scolaire (des cours d'appui en luxembourgeois par exemple), l'organisation d'un soutien systématique à la langue maternelle en ayant recours à des enseignants issus du groupe d'immigrés en question par exemple.

Voici comment parents, inspecteurs, directions et enseignants de nos échantillons se sont positionnés par rapport à cette proposition (tableau 7).

Tableau 7 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Organiser un soutien spécifique aux élèves d'origine étrangère en proposant un soutien systématique à la langue maternelle (ex. : des cours de portugais pour les élèves d'origine portugaise)"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1132	17,6	24,2	16,5	**28,9**	12,8	-
Inspecteurs	9	11,1	-	44,4	22,2	22,2	-
Directions	13	7,7	46,2	23,1	23,1	-	-
Enseignants	107	5,6	7,5	29	31,8	22,4	3,7
Enseignants du primaire	45	8,9	2,2	40	28,9	17,8	2,2
Enseignants du secondaire	62	3,2	11,3	21	33,9	25,8	4,8

Les chiffres du tableau 7 laissent apparaître un positionnement différencié entre parents et enseignants. Alors que les parents semblent partagés sur la question (42 % de désaccords, 41 % d'accords), les enseignants sont 54 % à marquer leur accord avec le scénario proposé (29 % d'avis mitigés et 13 % de désaccords). La position des membres de l'Inspection est mitigée voire favorable à la proposition (1 désaccord, 4 positions mitigées, 4 accords). Les directions par contre sont majoritairement opposées à ce scénario (7 désaccords, 3 positions mitigées et 3 accords).

1.2. Vers une redéfinition du métier d'enseignant et du fonctionnement des écoles

L'art d'enseigner est une tâche à la fois noble et difficile à réaliser. Elle nécessite des professionnels de l'apprentissage capables de travailler en équipe et d'analyser leur action pour l'améliorer continuellement. Pourtant, les enseignants ont souvent le sentiment d'être laissés seuls face à la tâche. Les travaux réalisés sur la satisfaction professionnelle des enseignants montrent d'ailleurs que le phénomène de *burn out* gagne en ampleur : l'enseignant adopte une attitude de résignation par rapport aux exigences non maîtrisées et trop lourdes de sa vie professionnelle. D'autres enseignants déclarent ouvertement qu'ils se sentent mal préparés par leur formation à la complexité de leur tâche professionnelle. Bref, le malaise enseignant est réel et l'école luxembourgeoise se doit de le prendre en compte. C'est dans cette optique que nous avons formulé plusieurs scénarios d'évolution possible du métier tant en termes de fonctionnement que de pratiques.

1.2.1. Permettre aux écoles d'afficher, à côté de la formation de base garantie, des profils particuliers (langues, sciences, arts, travail manuel, sports, …) afin de recruter des élèves qui se sentent attirés par ces profils.

La proposition d'instaurer un tronc commun pour tous les élèves jusqu'à 15 ans (cf. 1.1.2.) a été faite en vue de combattre les phénomènes de relégation des élèves les plus faibles vers les filières techniques suite à la procédure d'orientation conduite en fin d'école primaire. Ce scénario propose non seulement de retarder le moment de l'orientation mais il se caractérise avant tout par la volonté de faire acquérir à tous les élèves un socle minimal de compétences essentielles à la vie en société. Cela ne veut pour autant pas dire que l'uniformisation des apprentissages est recherchée, loin de là.

C'est pourquoi, dans la foulée de la précédente proposition, nous avons suggéré qu'à côté de la formation de base dispensée par les écoles à tous les élèves, soit envisagée une structure en modules ou en options qui permettrait une individualisation ponctuelle des apprentissages tout en gardant les élèves dans une structure unique. Les écoles pourraient également afficher des profils particuliers, par exemple en mettant l'accent alternativement sur l'enseignement des langues ou sur l'enseignement des sciences, afin de recruter des élèves qui se sentent attirés par ce profil spécifique.

Comme le montre le tableau 8, il apparaît que près de 87 % des parents et 69 % des enseignants consultés sont en accord avec la proposition formulée. Avec plus de 70 % d'accords, il en va de même au niveau des membres de l'Inspection et des directions d'école.

Tableau 8 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Permettre aux écoles d'afficher, à côté de la formation de base garantie, des profils particuliers (langues, sciences, arts, travail manuel, sports, …) afin de recruter des élèves qui se sentent attirés par ces profils"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1141	1,8	3,7	7,8	52,1	34,7	-
Inspecteurs	9	-	-	22,2	44,4	33,3	-
Directions	13	-	7,7	7,7	46,2	38,5	-
Enseignants	107	3,7	6,5	17,8	45,8	23,4	2,8
Enseignants du primaire	45	6,7	4,4	15,6	44,4	24,4	4,4
Enseignants du secondaire	62	1,6	8,1	19,4	46,8	22,6	1,6

1.2.2. Réorganiser la gestion des ressources consacrées aux enfants à besoins spécifiques en créant une équipe pluridisciplinaire au sein de chaque école

La question est ici de savoir s'il est préférable d'intégrer les enfants à besoins spécifiques dans l'enseignement ordinaire ou s'il est nécessaire de créer des classes d'enseignement spécialisé qui permettraient un encadrement plus adéquat et plus efficace pouvant, le cas échéant, déboucher sur une réinsertion dans l'enseignement ordinaire. Cette question touche aux finalités visées par l'école. Actuellement, le Luxembourg adopte une stratégie mixte, en gardant d'un côté, avec l'enseignement différencié, une structure d'enseignement spécialisé pour les enfants qui sont jugés incapables de suivre un régime d'enseignement ordinaire et en intégrant, de l'autre côté, des enfants à besoins spécifiques dans les classes qui peuvent alors bénéficier du soutien ponctuel de services externes aux écoles (SREA, …). On constate cependant assez souvent que ces interventions ponctuelles externes sont difficiles à intégrer dans les activités que l'enseignant réalise par ailleurs dans sa classe, ce qui peut amener à un manque de cohérence entre les activités d'apprentissage à l'intérieur de la classe et les activités d'apprentissage réalisées au cours des interventions ponctuelles externes.

Dans les pays où l'intégration semble bien fonctionner, on observe un co-enseignement (enseignant et éducateur spécialisé, Grundschul/SekundarstufenlehrerIn et SonderpädagogIn) avec des classes de +/- 20 élèves dont 5 à 6 enfants à besoins spécifiques afin d'assurer une hétérogénéité dans les classes qui permet d'éviter la stigmatisation de « l'enfant intégré ». Dès lors, dans l'optique d'une intégration des

enfants à besoins spécifiques dans la structure ordinaire, ne serait-il pas plus indiqué d'affecter les ressources humaines destinées au soutien des enfants à besoins spécifiques directement aux écoles en y créant des équipes pluridisciplinaires (y inclus les enseignants de l'école) chargées de gérer au mieux et dans une action concertée la diversité des élèves ? Il s'agirait donc de remplacer les services externes aux écoles par des équipes internes à l'école composées à la fois d'enseignants et de spécialistes ayant la co-responsabilité de gérer les différences au sein de leur école.

Tableau 9 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Réorganiser la gestion des ressources consacrées aux enfants à besoins spécifiques en créant une équipe pluridisciplinaire au sein de chaque école"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1121	2,3	6	27,9	41,4	22,4	-
Inspecteurs	9	-	-	22,2	33,3	44,4	-
Directions	13	-	-	**38,5**	23,1	38,5	-
Enseignants	106	2,8	3,8	24,5	34	31,1	3,8
Enseignants du primaire	45	2,2	2,2	20	37,8	37,8	-
Enseignants du secondaire	61	3,3	4,9	27,9	31,1	26,2	6,6

Au vu du tableau 9, il semble que face à cette proposition de réorganiser la gestion des ressources consacrées aux enfants à besoins spécifiques, les différents intervenants soient en majorité d'accord. On relève en effet que près de 65 % des parents et des enseignants consultés y sont favorables. On soulignera aussi l'absence de désaccords chez les inspecteurs et les directions.

1.2.3. Adjoindre à l'enseignant du personnel psychopédagogique supplémentaire

La réflexion menée au sujet des enfants à besoins spécifiques peut être élargie à la problématique de la gestion des différents rythmes d'apprentissage des élèves. Il est en effet illusoire de croire encore que l'ensemble des élèves de la classe va apprendre la même chose en même temps. Inévitablement, des différences apparaissent et l'enseignant se retrouve souvent seul face à la gestion de celles-ci. Nous avons dès lors proposé d'adjoindre du personnel psychopédagogique supplémentaire dans les écoles pour, le cas échéant, soutenir l'enseignant dans cette gestion si difficile de l'hétérogénéité des apprentissages des élèves.

Tableau 10 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Adjoindre à l'enseignant du personnel psychopédagogique supplémentaire"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1128	4,3	8,4	16	43,6	27,7	-
Inspecteurs	9	-	-	**44,4**	33,3	22,2	-
Directions	13	-	7,7	7,7	15,4	69,2	-
Enseignants	107	4,7	8,4	21,5	31,8	30,8	2,8
Enseignants du primaire	45	8,9	6,7	13,3	33,3	35,6	2,2
Enseignants du secondaire	62	1,6	9,7	27,4	30,6	27,4	3,2

A la lumière du tableau 10, on s'aperçoit que plus de 70 % des parents et plus de 60 % des enseignants seraient favorables à cette aide complémentaire au travail de l'enseignant. Au niveau de l'Inspection, on relève 4 positions mitigées et 5 accords. La prise de position des directions est quant à elle plus nette puisqu'on comptabilise 1 seul désaccord, 1 position mitigée et 11 accords avec la proposition.

1.2.4. Supprimer les notes scolaires et les remplacer par une évaluation formative des élèves (càd repérer les forces et les faiblesses des élèves et leur dire comment ils peuvent s'améliorer).

Les travaux en docimologie ont suffisamment mis en évidence les biais liés à l'évaluation, notamment celle menée par les enseignants, pour se poser la question de la pertinence des notes scolaires. Fondamentalement, on peut effectivement se demander ce que représentent réellement ces notes chiffrées du point de vue des enseignants, des parents et des élèves. Que signifient-ils vraiment pour chaque intervenant ? Qu'apportent-ils en termes de régulation de l'apprentissage ? De là à parler de la suppression des notes scolaires …

Tableau 11 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Supprimer les notes scolaires et les remplacer par une évaluation formative des élèves (càd repérer les forces et les faiblesses des élèves et leur dire comment ils peuvent s'améliorer"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1137	17,2	22,7	11,3	29	19,9	-
Inspecteurs	9	11,1	22,2	22,2	22,2	22,2	-
Directions	13	15,4	15,4	23,1	15,4	30,8	-
Enseignants	107	26,2	29	19,6	12,1	13,1	-
Enseignants du primaire	44	22,7	22,7	20,5	13,6	20,5	-
Enseignants du secondaire	63	28,6	33,3	19	11,1	7,9	-

On s'aperçoit à la lecture du tableau 11 que la question de la suppression des notes scolaires divise les parents et les enseignants. Du côté des parents semble émerger une majorité d'accords avec la proposition (50 % d'accords et 40 % de désaccords) tandis que du côté des enseignants la tendance s'inverse (55 % de désaccords, 25 % d'accords). Il y a environ 10 % d'indécis chez les parents et le double chez les enseignants. Chez les professionnels de l'éducation, le désaccord se marque surtout au niveau secondaire puisqu'ils sont plus de 60 % à s'opposer à la proposition contre 45 % au primaire. Au niveau de l'Inspection, le positionnement est des plus mitigés (3 désaccords, 2 positions mitigées et 4 accords). Les directions semblent quant à elles un peu plus favorables à la proposition (4 désaccords, 3 positions mitigées et 6 accords).

1.3. Vers un dispositif d'assurance-qualité de l'école

Depuis une vingtaine d'années, les pays démocratiques connaissent un mouvement de plus en plus marqué d'évaluation de leurs politiques publiques. Il s'agit non seulement de vérifier si les résultats obtenus correspondent effectivement aux objectifs fixés – on parlera dans ce cas d'efficacité – mais également de voir si ces résultats ont été obtenus au prix d'une mobilisation rationnelle et optimale des moyens mis à disposition – on parlera dès lors d'efficience. La demande d'évaluation de l'Ecole n'a jamais été aussi forte au vu des problèmes budgétaires auxquels de nombreux pays européens sont confrontés. Ceux-ci ne peuvent que vouloir l'efficacité de l'Ecole ou en tout cas une rentabilité des fonds investis dans celle-ci. S'y ajoutent d'une part la demande des familles, « consommateurs » ou usagers directs de l'Ecole et, d'autre part, les employeurs, tant publics que privés, qui ne peuvent plus se contenter d'une main d'œuvre non formée ou non qualifiée. Il s'ensuit que dans la

quasi-totalité des pays industrialisés, l'institution scolaire est sommée de rendre des comptes et d'être transparente tant sur le plan de l'utilisation des fonds publics que sur la qualité et l'efficacité de l'enseignement dispensé dans les écoles et les classes. Les études internationales telles que PISA sont le signe le plus tangible de cette orientation sur les résultats. Le pilotage se fait donc sur la base des résultats qu'un système scolaire arrive à engendrer auprès de sa population d'élèves. Les résultats sont opérationnalisés en termes de niveaux de compétence dans des domaines jugés comme fondamentaux pour la participation à la vie sociétale.

Cette orientation sur l'« output », sur les résultats, constitue pour un pays comme le Luxembourg un changement de paradigme par rapport à une logique de pilotage des politiques éducatives beaucoup plus orientée sur l'« input », c'est-à-dire sur la qualité «matérielle» de l'enseignement et de son cadre infrastructurel (par la fixation d'un nombre d'heures d'enseignement par branche, par l'élaboration de contenus et de programmes très précis, de critères de promotion très fins et constamment réadaptés, de directives d'examen et de notation, de manuels scolaires et d'autres matériaux pédagogiques obligatoires, etc.). Or, les résultats de ce pilotage par l'« input » tel qu'il était pratiqué au Luxembourg montrent qu'il ne s'avère pas être garant d'une formation de base indispensable pour la réussite dans la future société de la connaissance.

C'est pourquoi nous avons demandé aux parents et aux enseignants de nos échantillons de se positionner par rapport à la proposition suivante : « Créer un organisme national indépendant chargé d'évaluer de manière régulière l'efficacité du système scolaire ».

A quelques exceptions près (dont le Luxembourg), tous les pays démocratiques se sont dotés d'un dispositif de pilotage de leur système éducatif qui, dans les faits, se traduit par la création d'une structure officielle et permanente (pérenne aux différentes législatures) chargée d'évaluer continuellement si le système atteint bien les finalités qu'il s'est fixées (Opetushalitus en Finlande, OFSTED et QCA en Angleterre, CITO aux Pays-Bas, DEP en France, DVO en Communauté flamande de Belgique, Commission de pilotage en Communauté française de Belgique, …). Ces différentes instances peuvent conduire ou coordonner trois types de pilotage qui se différencient par leur fonction mais qui sont tous nécessaires : le pilotage administratif, le pilotage formatif (ou pilotage diagnostique) et le pilotage du rendement scolaire (ou pilotage par les résultats).

Le pilotage d'un système éducatif suppose que des décisions soient prises et que des interventions soient menées si l'on constate que les objectifs du système ne sont pas rencontrés. Cependant, pour que les décisions d'intervention soient fondées, il est vital que les acteurs de la régulation puissent baser leurs décisions sur des informations fiables. Ce besoin d'informations fiables sur le fonctionnement et l'efficacité du système doit conduire les acteurs du pilotage à recueillir, de manière récurrente afin qu'un suivi dans le temps soit possible, un ensemble d'indicateurs permettant de cerner avec précision les domaines ou les lieux dans lesquels une intervention est souhaitable. On pourrait ainsi imaginer la construction d'un tableau

d'indicateurs comme les « Indicateurs pancanadiens de l'Education » (Canada) ou la publication française « L'Etat de l'école » qui, en donnant une série d'indicateurs de fonctionnement et de résultat, permettent un pilotage du système.

Tableau 12 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Créer un organisme national indépendant chargé d'évaluer de manière régulière l'efficacité du système scolaire"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1124	3,8	4	19	40	33,2	-
Inspecteurs	9	11,1	11,1	11,1	44,4	22,2	-
Directions	13	-	7,7	7,7	38,5	46,2	-
Enseignants	107	3,7	11,2	27,1	33,6	15	9,3
Enseignants du primaire	45	4,4	11,1	22,2	33,3	20	8,9
Enseignants du secondaire	62	3,2	11,3	30,6	33,9	11,3	9,7

Comme l'indique le tableau 12, 73 % des parents et 49 % des enseignants se sont montrés favorables à la création d'une structure officielle permanente chargée du pilotage du système éducatif luxembourgeois. Ils sont moins de 8 % chez les parents et moins de 15 % chez les enseignants à avoir exprimé un désaccord. L'accord est assez net également chez les inspecteurs et les directions d'établissement.

2. Positionnement des acteurs par rapport aux propositions visant à créer ou renforcer le lien entre les écoles et les familles

Depuis la fin de la seconde guerre mondiale, l'ensemble du système éducatif a été confronté à une massification de l'enseignement. Les écoles qui hier formaient une minorité doivent aujourd'hui nécessairement brasser un public plus large. Les enseignants sont dès lors amenés à faire face à une diversité de jeunes et par conséquent, à élargir le répertoire de leurs « pratiques » pour gérer cette hétérogénéité. La révision de leurs « pratiques » concerne aussi bien les méthodes d'enseignement et d'apprentissage auxquelles ces professionnels ont recours dans leur classe que le panel d'attitudes à adopter pour gérer les relations avec les élèves et leurs parents.

L'institution scolaire n'est pas la seule à avoir fortement évolué ces soixante dernières années. Les familles d'aujourd'hui diffèrent dans une large mesure des familles de l'après-guerre. La baisse de la natalité a entraîné la réduction de la taille

des familles. Le statut des couples transforme lui aussi la vie familiale : l'union libre remplace dans de nombreux cas le mariage. Les cas de divorce et de séparation se sont multipliés, entraînant de ce fait une augmentation du nombre de familles monoparentales. Outre les modifications relatives à la composition familiale, on relève aussi des évolutions touchant au fonctionnement familial : les femmes, traditionnellement moins présentes sur le marché du travail, occupent désormais de plus en plus des emplois rémunérés. Toutes ces mouvances modifient les rapports entre membres de la famille ainsi que leur statut mais aussi les pratiques éducatives des parents (Conseil Supérieur de l'Education du Québec, 1994, pp. 5-18).

Pour Terrail (1997, pp. 78-79), l'essor pris par l'école dans les années 60 expliquerait en partie les modifications observées au sein des familles. Selon l'auteur, des bénéfices pourraient être retirés de la scolarisation de masse tels que le développement des capacités individuelles, l'élargissement des trajectoires sociales destinées jusqu'alors à une minorité privilégiée, et partant, la mobilisation des familles autour du projet scolaire destiné à faire émerger l'individu par sa réussite.

L'évolution des modes d'échanges école-famille trahit une partie des bouleversements relevés au sein des deux environnements éducatifs. Jusque dans les années 30 à 50, la sphère privée et l'institution scolaire sont cloisonnées ; chacune d'elle œuvre à l'éducation de l'enfant de façon indépendante : à la première revient l'éducation et la transmission des valeurs ; à la seconde, l'instruction et la transmission des savoirs. A partir des années 50, les missions de l'une comme de l'autre s'élargissent mais les deux pôles éducatifs demeurent relativement éloignés l'un de l'autre. L'école n'assigne qu'aux parents des tâches de soutien et de contrôle (encouragement, sanction, signature du bulletin). Dans les années 60, une étape est franchie : les programmes scolaires sont rendus accessibles aux parents. Mais c'est réellement dans les années 70 et 80 que les parents franchissent les portes de l'école, que la possibilité de vrais échanges apparaît et que les premières théories sur les responsabilités de chacun des acteurs voient le jour. Toutefois, malgré des rapprochements certains entre l'école et la famille, la communication et les échanges école-famille ne vont pas nécessairement de soi.

S'il est difficile de ne pas être en accord avec le principe selon lequel tout type d'éducation cherche à rendre les jeunes autonomes, dans les faits, les choses sont bien différentes. Lorsque l'enfant devient élève, il relève de deux mondes éducatifs distincts : la famille et l'école. « *Aussi n'est-il pas rare de voir les uns et les autres se disputer le bien des enfants et s'opposer sur les limites de leurs territoires respectifs. Si la guerre n'est pas ouverte entre l'école et les parents, règne parfois la paix armée à la frontière des compétences attendues des uns et des autres* » (Dubet, 1997, p. 13). Or, dans ce jeu d'influences, il semble que l'institution scolaire bénéficie des meilleurs atouts parce qu'elle « *exerce un pouvoir institutionnel garanti par l'Etat* » (Gayet, 1999, p. 11). C'est un peu comme si les enseignants oeuvraient logiquement pour l'intérêt général alors que l'on pourrait craindre des parents, même au niveau de leurs représentants légaux, une défense de l'intérêt particulier au détriment du collectif. Sur le plan participatif, force est de constater que les parents n'ont en effet pas souvent voix au chapitre.

Il faut dès lors s'interroger sur la place des parents au sein de l'institution scolaire luxembourgeoise. A l'instar de nombreux pays européens, il semble que le monde scolaire luxembourgeois se montre frileux à l'ouverture de ses portes aux parents d'élèves. Les textes officiels, il est vrai, ne facilitent pas la mise en place d'une communication fluide et efficace entre les deux environnements éducatifs. A ce propos, Pourtois, Desmet et Nimal (1999) remarquent que « *la politique de tous les Etats européens ne permet pas aux parents de participer à l'acte d'enseignement qui relève des seuls enseignants. Au niveau des politiques locales, par contre, les parents peuvent être invités à participer, en tant que partenaires, au processus de production des actes d'enseigner* » (p. 6).

Au Luxembourg, une revue des textes officiels montre que le seul règlement grand-ducal du 3 mai 1989 fixant la tâche des enseignants de l'éducation préscolaire et de l'enseignement primaire envisage brièvement les échanges école-famille. A l'article 2, point c de ce règlement, relatif à la définition de la tâche de l'enseignant, on peut lire que la tâche d'orientation et de concertation comprend, en autres choses, la consultation obligatoire[7] des parents à raison d'une heure par quinzaine ainsi que l'information des parents au début de l'année scolaire. Par ailleurs, les parents sont présents dans les différents organes de participation au niveau national, communal et de l'établissement (Bogdanowicz, 1997). Au niveau national, quatre représentants de parents, nommés par le ministre pour une durée de 2 ans, siègent au conseil supérieur de l'éducation (42 membres au total) ; leurs compétences sont consultatives. Pour le primaire, il existe une commission d'instruction, au sein de laquelle, le ministre nomme, sur une base volontaire, un représentant des parents (10 membres au total) pour une durée de 4 ans alors que la loi n'en prévoit pas. Cette commission est chargée de donner un avis sur toute question concernant l'enseignement primaire mais aussi d'approuver le choix des manuels scolaires. Au niveau des communes, il existe une commission scolaire pour le préscolaire et pour le primaire au sein de laquelle on trouve au moins un père et une mère sur les 5 à 7 membres (selon le nombre d'habitants) nommés pour 6 années par le conseil communal. Ces parents ont un rôle consultatif en ce qui concerne les problèmes en rapport avec l'organisation scolaire et un rôle décisionnel en ce qu'il s'agit de veiller à l'exécution de la loi sur l'instruction obligatoire (bonne fréquentation de l'école, répression des absences) et à l'entretien des constructions scolaires. Au niveau des établissements secondaires, on trouve le conseil de l'éducation, composé de 9 membres dont 2 parents, élus pour une durée de 2 ans. Ces parents s'impliquent directement dans la gestion et le fonctionnement quotidien de l'établissement et dans l'élaboration des projets d'établissement. Notons encore l'existence d'associations locales de parents d'élèves de l'enseignement primaire et de l'enseignement post-primaire ainsi que deux fédérations regroupant, l'une, des associations locales du primaire et, l'autre, des associations locales du post-primaire. Enfin, il faut savoir que les enseignants ne reçoivent aucune formation en ce qui concerne les relations parents-enseignants.

[7] Que recouvre cette notion de consultation obligatoire ?

Les familles sont pourtant parmi les premiers usagers du système. Ce sont leurs enfants que l'école forme. Et il est aujourd'hui démontré que la réussite scolaire est intimement liée à l'intérêt que les parents manifestent pour les activités scolaires de leur enfant que ce soit par l'investissement cognitif de ces derniers dans le suivi scolaire de leur enfant, leur participation à la vie scolaire (présence aux réunions de parents, aux fêtes scolaires, …) ou encore les aspirations qu'ils développent à l'égard de l'avenir scolaire ou professionnel de leur progéniture. Il serait donc tout à fait logique, comme dans toute société démocratique, qu'ils puissent intervenir et faire entendre leur position dans les processus de décision en matière d'éducation. Même si, par leur diversité, les familles ne sont représentatives que d'elles-mêmes alors que l'on peut penser qu'il existe une relative unité au sein du corps enseignant, il est essentiel, pour une prise en compte globale de l'enfant, qu'école et famille prennent part à l'éducation scolaire du jeune.

C'est dans cet esprit de renforcer le lien école-famille que nous avons formulé dix propositions à propos desquelles parents et enseignants ont été invités à prendre position. Ces dix propositions sont classées dans quatre catégories distinctes :

- La part de responsabilité des parents à l'école
- La présence des parents aux activités de l'école
- La communication et les échanges école-famille
- Le degré d'information des parents

2.1. La part de responsabilité des parents à l'école

Un item entre dans cette dimension. Il traite de la part de responsabilité et de décision prise par les parents au sein des établissements : « *Donner aux parents une responsabilité et donc une part de décision dans la définition du projet pédagogique et des objectifs de l'établissement fréquenté par leur enfant* ».

Les chiffres du tableau 13 laissent entrevoir des positionnements assez contrastés par rapport au scénario proposé.

Tableau 13 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Donner aux parents une responsabilité et donc une part de décision dans la définition du projet pédagogique et des objectifs de l'établissement fréquenté par leur enfant"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1125	6	16,7	24	38,8	14,4	-
Inspecteurs	9	22,2	11,1	22,2	33,3	11,1	-
Directions	13	15,4	7,7	46,2	15,4	15,4	-
Enseignants	106	18,9	31,1	25,5	18,9	3,8	1,9
Enseignants du primaire	44	20,5	31,8	31,8	9,1	2,3	4,5
Enseignants du secondaire	62	17,7	30,6	21	25,8	4,8	-

Ainsi, on constate tout d'abord que les parents d'élèves sont plutôt favorables à la proposition (23 % de désaccords, 25 % de positions mitigées et plus de 50 % d'accords). Les membres de l'Inspection sont davantage partagés (3 désaccords, 2 positions mitigées, 4 accords) tout comme les directions d'école (3 désaccords, 6 positions mitigées, 4 accords). Enfin, avec 50 % de désaccords, les enseignants semblent majoritairement hostiles à la proposition.

2.2. La présence des parents aux activités de l'école

Quatre items sont réunis dans cette section. Ils ont trait à la mise en place d'activités visant l'augmentation de la présence des parents sur la scène scolaire et la sensibilisation des parents au rôle qu'ils tiennent dans la réussite scolaire de leur enfant.

2.2.1. Renforcer le lien entre l'école et les familles en organisant des rencontres et des échanges au sein même des écoles

En ouvrant l'école aux parents par l'organisation de rencontres et d'échanges au sein même des écoles, nous faisons ainsi l'hypothèse qu'il est possible de réduire l'écart entre le milieu familial et le milieu scolaire. L'enfant ne doit plus être le seul à tenter de combler ce fossé, il gagne à être encadré par un dispositif participatif qui fait le lien entre son milieu familial et l'école.

Face à cette proposition de renforcer le lien entre l'école et les familles, parents et enseignants semblent massivement favorables puisque près de 75 % des parents et 73 % des enseignants interrogés marquent leur accord (tableau 14). Il en est

de même pour les inspecteurs et les directions dont l'accord avec la proposition est presque total.

Tableau 14 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Renforcer le lien entre l'école et les familles en organisant des rencontres et des échanges au sein même des écoles"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord	Ne souhaite pas se prononcer
Parents	1130	1,6	8,1	16	53,8	20,4	-
Inspecteurs	9	-	11,1	-	55,6	33,3	-
Directions	13	-	7,7	-	53,8	38,5	-
Enseignants	107	3,7	5,6	16,8	49,5	23,4	0,9
Enseignants du primaire	44	4,5	9,1	20,5	45,5	18,2	2,3
Enseignants du secondaire	63	3,2	3,2	14,3	52,4	27	-

Si les intervenants scolaires semblent tous d'accord pour renforcer le lien entre l'école et les familles, reste à cerner les démarches à entreprendre et les efforts à consentir par les différents partenaires. C'est l'objet des propositions qui suivent.

2.2.2. Faciliter la mobilisation et la participation des parents aux activités de l'école

La recherche en sciences de l'éducation a mis en évidence le lien qui existait entre la présence des parents sur la scène scolaire et la réussite scolaire des jeunes. Par ailleurs, il apparaît que le rapprochement entre l'école et les familles, outre l'influence constatée sur la scolarité de l'élève, peut avoir des retombées positives pour l'équipe éducative. Notons par exemple, *« une meilleure connaissance réciproque, la découverte de goûts communs et le plaisir de connaître l'autre sous un jour nouveau. »* (CRIRES, 1993). Ces échanges favorisent aussi l'élargissement de perspective et l'amélioration de la relation éducative. Quel est l'avis de différents partenaires sur le sujet ?

Tableau 15 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Faciliter la mobilisation et la participation des parents aux activités de l'école"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1123	1,6	10,7	20	54,6	13,1
Enseignants	107	4,7	24,3	-	58,9	12,1
Enseignants du primaire	45	2,2	31,1	-	57,8	8,9
Enseignants du secondaire	62	6,5	19,4	-	59,7	14,5

Ici encore, la proposition recueille près de 70 % d'accords chez les parents et les enseignants, ce qui témoigne peut être d'une volonté de renforcer la participation des parents à la vie de l'école. On soulignera cependant que 30 % des enseignants s'y opposent.

2.2.3. Faire l'inventaire en début d'année de la disponibilité, des talents et des centres d'intérêt des parents et organiser des activités à l'école ou en dehors impliquant les parents

Permettre aux parents d'être présents sur la scène scolaire offre encore l'avantage de constituer de nouvelles ressources pour l'école. Il s'agit d'une contribution éducative intéressante qui donne à l'école de nouvelles possibilités de contacts avec l'extérieur. Par conséquent, pourquoi ne pas faire le bilan avec les parents en début d'année de leur disponibilité, de leurs talents et de leurs centres d'intérêt en vue d'organiser des activités en dehors ou dans l'école en collaboration avec les parents ?

Tableau 16 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Faire l'inventaire en début d'année de la disponibilité, des talents et des centres d'intérêt des parents et organiser des activités à l'école ou en dehors impliquant les parents"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1121	4,2	16,1	**25,9**	41	12,8
Enseignants	107	11,2	**31,8**	-	46,7	10,3
Enseignants du primaire	44	9,1	27,3	-	56,8	6,8
Enseignants du secondaire	63	12,7	34,9	-	39,7	12,7

L'idée d'établir en commun en début d'année la disponibilité, les talents et les centres d'intérêt des parents en vue d'organiser des activités à l'école impliquant

les parents reçoit l'aval de la majorité des parents et des enseignants. Toutefois, notons quand même la position mitigée d'un quart des parents et le fait que 3 à 4 enseignants sur 10 marquent leur désaccord à l'égard de cette démarche.

2.2.4. Organiser des activités à caractère social qui rassemblent les élèves, le personnel scolaire, les familles, les partenaires sociaux et les citoyens du quartier

Pourquoi ne pas encore élargir les possibilités d'échanges et de rencontres en mettant en place des activités à caractère social qui rassemblent l'ensemble des acteurs de l'école et qui gravitent directement autour de l'école : les élèves, le personnel scolaire, les partenaires sociaux et les citoyens du quartier ?

Tableau 17 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Organiser des activités à caractère social qui rassemblent les élèves, le personnel scolaire, les familles, les partenaires et les citoyens de du quartier"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1123	2,5	12	19,5	50	16
Enseignants	106	12,3	33	-	40,6	14,2
Enseignants du primaire	45	6,7	24,4	-	48,9	20
Enseignants du secondaire	61	16,4	39,3	-	34,4	9,8

Un fossé clair oppose les réponses des parents, des enseignants du primaire et les enseignants du secondaire. Ces derniers se disent le plus souvent peu prêts à mettre en place des activités à caractère social réunissant les élèves, le personnel scolaire, les familles, les partenaires et les citoyens du quartier.

2.2.5. Organiser des activités en parallèle pour les élèves et les parents afin de montrer aux élèves l'importance que leurs parents accordent à l'école

Tableau 18 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Organiser des activités en parallèle pour les élèves et les parents afin de montrer aux élèves l'importance que leurs parents accordent à l'école"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1124	2,8	11,2	17,5	50,7	17,7
Enseignants	107	6,5	28	-	53,3	12,1
Enseignants du primaire	45	8,9	24,4	-	53,3	13,3
Enseignants du secondaire	62	4,8	30,6	-	53,2	11,3

Une position clairement positive réunit parents et enseignants autour de l'organisation d'activités en parallèle pour les élèves et les parents en vue de montrer aux éduqués toute l'importance que leurs parents accordent à l'école.

2.3. La communication et les échanges école-famille

2.3.1. Mettre en place des moyens de communication plus efficaces et plus variés entre l'école et les familles

Pour renforcer les échanges entre l'école et la famille, des leviers à la communication doivent être mis en place et les obstacles doivent être levés. C'est pour cette raison qu'il importait de vérifier le degré d'accord des familles et des enseignants à l'égard de la proposition relative à la mise en place de moyens de communication plus efficaces et plus variés entre les deux partenaires éducatifs.

Tableau 19 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Mettre en place des moyens de communication plus efficaces et plus variés entre l'école et les familles"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1129	1	5,3	13,5	58,5	21,8
Enseignants	106	2,8	11,3	-	67,9	17,9
Enseignants du primaire	43	2,3	11,6	-	74,4	11,6
Enseignants du secondaire	63	3,2	11,1	-	63,5	22,2

En répondant massivement qu'ils sont d'accord avec le scénario formulé (plus de 80 % d'accords), parents et enseignants semblent ne pas être satisfaits des moyens de communication actuellement mis en place entre l'école et les familles. Au vu de leur prise de position, ils semblent en tous cas prêts à communiquer davantage.

2.3.2. Nommer dans chaque classe un délégué des parents qui servirait de lien entre les parents et l'enseignant.

Une solution possible, susceptible de renforcer la liaison école-famille, serait de nommer au sein de chaque classe un délégué de parents qui deviendrait un médiateur utile dans la communication école-famille.

Tableau 20 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Nommer dans chaque classe un délégué des parents qui servirait de lien entre les parents et l'enseignant"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1126	8,4	18,7	19,6	36,3	17
Enseignants	106	20,8	37,7	-	30,2	11,3
Enseignants du primaire	43	30,2	48,8	-	14	7
Enseignants du secondaire	63	14,3	30,2	-	41,3	14,3

Les enseignants du primaire, contrairement à leurs collègues du secondaire et aux parents de leurs élèves, ne se disent pas prêts à désigner au sein de chaque classe un délégué de parents qui pourrait être utilisé comme lien entre l'école et la famille.

2.4. Le degré d'information des parents

2.4.1. Informer les parents en organisant des réunions ou des conférences sur leur rôle et leur responsabilité par rapport à la scolarité de leur enfant (ex : conseils en matière de communication parents-enfant, conseils d'aide aux devoirs, …)

Les parents sont-ils informés de leur rôle et de leur responsabilité par rapport à la scolarité de leur enfant ? Cette information est, nous semble-t-il, un préalable indispensable à la mise en place d'une communication fluide et efficace entre l'école et la famille mais aussi à toute démarche collaborative.

Tableau 21 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Informer les parents en organisant des réunions ou des conférences sur leur rôle et leur responsabilité par rapport à la scolarité de leur enfant (ex. conseils en matière de communication parents-enfant, conseils d'aide aux devoirs, …)"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1139	1,2	6,9	11,6	59,1	21,2
Enseignants	107	0,9	12,1	-	65,4	21,5
Enseignants du primaire	45	-	15,6	-	62,2	22,2
Enseignants du secondaire	62	1,6	9,7	-	67,7	21

Les réponses données par parents et enseignants sont massivement positives même si on observe des positions mitigées chez presque 12% des parents et des positions plus en désaccord chez 13% des enseignants. Les partenaires se montrent plutôt demandeurs d'une plus grande transparence et d'une clarification des rôles mais aussi des responsabilités que le monde scolaire attend chez les parents en ce qui concerne le suivi scolaire des jeunes.

2.4.2. Informer les parents de l'existence d'autres partenaires ou organismes de leur quartier susceptibles de les aider (école de devoirs, centre de ressources, …)

Certains parents se trouvent parfois démunis devant le suivi scolaire de leur enfant. Leur indiquer la démarche à suivre constitue une manière indirecte d'aider le jeune dans sa scolarité.

Tableau 22 – Degré d'accord des différents intervenants scolaires avec l'affirmation "Informer les parents de l'existence d'autres partenaires ou organismes de leur quartier susceptibles de les aider (école de devoirs, centre de ressources, …)"

	N	Pas du tout d'accord	Plutôt pas d'accord	Position mitigée	Plutôt d'accord	Tout à fait d'accord
Parents	1128	0,9	2,8	9,9	58,7	27,7
Enseignants	107	0,9	4,7	-	63,6	30,8
Enseignants du primaire	45	-	2,2	-	66,7	31,1
Enseignants du secondaire	62	1,6	6,5	-	61,3	30,6

Parents et enseignants se montrent en très grande majorité d'accord en ce qui concerne l'identification de partenaires susceptibles d'aider les parents dans le suivi scolaire de leur enfant.

3. Eléments conclusifs

Le présent article porte sur une partie des informations récoltées dans le cadre de recherche « *La place de l'école dans la société luxembourgeoise de demain* ». Plus précisément, il s'agit de rendre compte du positionnement de quatre échantillons d'acteurs du système éducatif luxembourgeois (parents, enseignants, directeurs et inspecteurs) par rapport à plusieurs scénarios d'évolution possibles. Ces scénarios sont répartis en deux catégories : les premiers relèvent du structurel et du pédagogique tandis que les seconds ont trait aux relations école-familles. Les conclusions qui suivent s'articulent selon cette distinction.

3.1. Positionnement des acteurs par rapport aux propositions d'ordre structurel ou pédagogique

Les différents scénarios soumis à l'évaluation des parents, des enseignants, des directions et des membres de l'Inspection ont été rédigés dans la cadre de la consultation par *focus group* à partir de l'analyse de l'état actuel du système éducatif luxembourgeois et de ce qui se fait dans d'autres systèmes éducatifs. Ces propositions recouvrent trois axes distincts qui visent la gestion de l'hétérogénéité des élèves, la redéfinition du métier d'enseignant et du fonctionnement des écoles ainsi que la mise en place d'un dispositif d'assurance-qualité de l'école.

3.1.1. Vers une gestion de l'hétérogénéité des élèves

Six aspects sont envisagés dans cette section. Le tableau qui suit rappelle les thèmes abordés et les tendances principales observées.

Tableau 23 - Thèmes abordés dans les scénarios relatifs à la gestion de l'hétérogénéité et les tendances principales observées

Thèmes	Pas d'accord	Position mitigée	D'accord
Supprimer le redoublement au primaire	Parents et Enseignants	------	Direction et Inspection
Supprimer le redoublement au secondaire	Parents et Enseignants	Inspection	Direction
Retarder le moment de l'orientation vers les filières d'enseignement	Enseignants du secondaire	------	Parents Enseignants du primaire Inspection Direction
Libre choix de l'établissement	Enseignants du primaire	------	Parents Enseignants du secondaire Inspection Direction
Le choix laissé aux familles de la langue d'alphabétisation à l'école primaire	Tous les acteurs	------	------
Supprimer les devoirs à domicile	Tous les acteurs	------	------
Organiser un soutien spécifique en langue maternelle aux élèves d'origine étrangère	Parents et Direction	Inspection	Enseignants

Si le choix de la langue d'alphabétisation laissé aux familles ainsi que la suppression des devoirs à domicile se voit rejeter de manière unanime, les quatre autres propositions divisent davantage.

A l'inverse de tous les autres acteurs interrogés, les enseignants du secondaire sont clairement opposés au retardement du moment de l'orientation vers les filières d'enseignement. En ce qui concerne le choix laissé aux parents de l'établissement fréquenté par leur enfant, ce sont les enseignants du primaire qui affichent un avis défavorable alors que les autres acteurs se positionnent plutôt positivement.

Le scénario relatif à la suppression du redoublement oppose clairement deux blocs d'acteurs : les enseignants et les parents qui sont en désaccord avec la proposition *vs* les directions d'établissement et les membres de l'Inspection qui

semblent plus favorables à l'abolition du redoublement. Ces derniers adoptent toutefois une position intermédiaire lorsque la problématique se situe au secondaire.

Alors que parents et directions d'établissement sont visiblement contre l'organisation d'un soutien spécifique en langue maternelle aux élèves d'origine étrangère, les membres de l'Inspection occupent une position intermédiaire et les enseignants se montrent plutôt en accord avec le scénario.

3.1.2. Vers une redéfinition du métier d'enseignant et du fonctionnement des écoles

Quatre propositions entrent dans la composition de cet axe de changement. Le tableau qui suit reprend les thèmes abordés et les tendances principales observées.

Tableau 24 : Thèmes abordés dans les scénarios relatifs à la redéfinition du métier d'enseignant et du fonctionnement des écoles et les tendances principales observées

Thèmes	Pas d'accord	Position mitigée	D'accord
Laisser la possibilité aux écoles d'afficher, à côté de la formation de base garantie, des profils particuliers (langues, sciences, arts, travail manuel, sports, …) afin de recruter des élèves qui se sentent attirés par ces profils	------	------	0Tous les acteurs
Réorganiser la gestion des ressources consacrées aux enfants à besoins spécifiques en créant une équipe pluridisciplinaire au sein de chaque école	------	------	Tous les acteurs
Adjoindre à l'enseignant du personnel psychopédagogique supplémentaire	------	------	Tous les acteurs
Supprimer les notes scolaires et leur remplacement par une évaluation formative	***0Enseignants***		Parents Inspection Direction

A l'exception de la proposition relative à la suppression des notes scolaires et leur remplacement par une évaluation formative où se distinguent une majorité d'enseignants défavorables, les trois autres propositions recueillent l'assentiment de l'ensemble des intervenants scolaires. Ainsi, la spécialisation des offres d'enseignement tout en garantissant le maintien d'une formation de base aux compétences essentielles, la réorganisation de la gestion des ressources des enfants à besoins spécifiques et l'adjonction de personnel psychopédagogique supplémentaire à

l'enseignant obtiennent l'approbation générale de nos quatre échantillons. Notons toutefois que 5 directions sur 13 choisissent la tendance centrale par rapport à la deuxième proposition et que 4 inspecteurs interrogés sur 9 semblent indécis en ce qui concerne la troisième proposition.

13.1.3. Vers un dispositif d'assurance-qualité de l'école

Un seul aspect est soumis à l'évaluation des acteurs, les tendances enregistrées pour ce point sont reprises dans le tableau ci-dessous.

Table 25 - Thème abordé dans le scénario relatif au dispositif d'assurance-qualité de l'école et les tendances principales observées

Thèmes	Pas d'accord	Position mitigée	D'accord
Créer un organisme national indépendant chargé d'évaluer de manière régulière l'efficacité du système scolaire	------	------	2Tous les acteurs

La majorité des acteurs interviewés affichent une attitude favorable à la mise en place d'une structure officielle permanente qui aurait en charge le pilotage du système éducatif luxembourgeois.

En définitive, sur les onze propositions soumises à évaluation se dégage un accord majoritaire, favorable ou défavorable, dans nos quatre échantillons pour six scénarios. Le tableau ci-dessous reprend ces six propositions.

Table 26 : Liste des propositions pour lesquelles il y a un accord ou un désaccord majoritaire uniforme

Approbation majoritaire et uniforme pour :

- Créer un organisme national indépendant chargé d'évaluer de manière régulière l'efficacité du système scolaire
- Permettre aux écoles d'afficher, à côté de la formation de base garantie, des profils particuliers (langues, sciences, arts, travail manuel, sports, …) afin de recruter des élèves qui se sentent attirés par ces profils.
- Réorganiser la gestion des ressources consacrées aux enfants à besoins spécifiques en créant une équipe pluridisciplinaire au sein de chaque école
- Adjoindre à l'enseignant du personnel psychopédagogique supplémentaire

Désapprobation majoritaire et uniforme pour :

- Laisser aux élèves et aux parents le choix de la langue d'alphabétisation à l'école primaire (allemand ou français)
- Supprimer les devoirs à domicile

Il ressort des résultats obtenus que l'ensemble des acteurs de la scène scolaire souhaiterait voir la mise en place d'un organe permanent d'évaluation du système scolaire luxembourgeois chargé d'évaluer l'atteinte des finalités éducatives que le système s'est fixées.

Il apparaît également que parents, enseignants, direction et membres de l'Inspection soient particulièrement favorables à l'idée que toutes les écoles garantissent la même formation de base à tous les élèves mais qu'elles puissent se distinguer en offrant des formations complémentaires spécialisées.

En ce qui concerne l'accueil des enfants à besoins spécifiques, il apparaît que les divers intervenants manifestent un souci certain d'intégration de ce type de population scolaire au sein des classes de l'enseignement ordinaire. Ce type de constat joue en faveur de l'affectation des ressources humaines dévolues aux enfants à besoins spécifiques directement aux écoles et en instaurant des équipes pluridisciplinaires chargées de gérer, de façon concertée, l'hétérogénéité du public dont ils ont la charge.

Nos quatre échantillons s'accordent encore à considérer que sur l'adjonction de personnel psychopédagogique serait nécessaire pour aider l'enseignant principal dans sa tâche d'enseignement.

Pour ce qui est de la proposition de laisser aux familles le choix de la langue d'alphabétisation, nous observons un désaccord net de l'ensemble des répondants, y compris les parents. Personne ne semble souhaiter que les parents optent pour la langue principale d'enseignement de leur enfant, langue qui deviendrait dominante pour la suite de leur parcours scolaire avec une deuxième langue qui prendrait le statut de langue étrangère et par rapport à laquelle les exigences de l'enseignant seraient moindre.

Enfin, quelle que soit l'origine des prises de position, un désaccord unanime est relevé en ce qui concerne la suppression éventuelle des devoirs à domicile.

Six propositions laissent la possibilité d'un débat entre les différents acteurs. Le tableau ci-dessous les présente.

Table 27 : Liste des propositions pour lesquelles les répondants sont indécis ou partagés

- Supprimer le redoublement à l'école primaire
- Supprimer le redoublement à l'école secondaire
- Retarder le moment de l'orientation vers les filières d'enseignement (classique, technique ou professionnel) en instaurant un tronc commun jusqu'à 15 ans (même formation de base pour tous) et des modules de formation permettant une individualisation ponctuelle des apprentissages
- Laisser aux parents le choix de l'établissement scolaire de leur enfant
- Organiser un soutien spécifique aux élèves d'origine étrangère en proposant un soutien systématique à la langue maternelle (ex. : cours de portugais pour les élèves d'origine portugaise)
- Supprimer les notes scolaires et les remplacer par une évaluation formative des élèves (càd repérer les forces et les faiblesses des élèves et leur dire comment ils peuvent s'améliorer)

Parents et enseignants ne sont pas systématiquement opposés dans les divergences observées autour de ces six scénarios. Ce constat est de bon augure pour la mise en place d'une communication fluide et efficace entre parents et enseignants (voir point 2). S'entendre sur certains aspects fondamentaux du système scolaire rend sans doute plus aisée l'entame des discussions. Voyons à présent comment se distribuent les positions des différents acteurs selon les six propositions ciblées.

Lorsqu'il s'agit de la suppression du redoublement, au primaire ou au secondaire, parents et enseignants adoptent une attitude concordante : ils sont plutôt contre la suppression du redoublement. A l'inverse, les directions sont pour l'abolition du redoublement quel que soit le niveau tandis que les inspecteurs sont soit « pour » lorsque le débat se situe au niveau du primaire, soit « contre » lorsqu'il s'agit du secondaire.

L'idée de retarder le moment de l'orientation scolaire voit s'opposer les enseignants du secondaire à tous les autres types d'intervenants. Au niveau du scénario « *Laisser aux parents le choix de l'école secondaire* », ce sont les enseignants du primaire qui, en marquant leur désaccord, se distinguent de tous les autres acteurs.

Pour ce qui est de l'organisation d'un soutien systématique à la langue maternelle aux élèves d'origine étrangère, les enseignants du primaire et du secondaire qui se montrent favorables à une telle proposition font front aux directions et aux parents qui sont plutôt en désaccord. L'Inspection adopte quant à elle une attitude intermédiaire.

Enfin, on observe une scission en deux blocs distincts au niveau de la proposition relative à la suppression des notes scolaires et à leur remplacement par une évaluation formative : d'un côté les enseignants qui y sont majoritairement favorables et d'un autre côté, tous les autres acteurs qui y sont défavorables.

3.2. Positionnement des acteurs par rapport aux propositions visant à créer ou à renforcer le lien entre les écoles et les familles

Quel est l'avis des principaux acteurs de la scène scolaire en ce qui concerne les relations entre l'école et la famille ?

Quatre aspects distincts de la relation école-famille chapeautent les dix scénarios proposés aux quatre échantillons : 1) La part de responsabilité des parents sur la scène scolaire ; 2) La présence des parents aux activités de l'école ; 3) La communication et les échanges école-famille et 4) Le degré d'information des parents. Le tableau ci-dessous reprend ces quatre aspects, les propositions qui s'y rapportent ainsi que les tendances centrales observées.

Les tableaux ci-dessous reprennent les quatre aspects.

Tableau 28 - Scénario relatif à la part de responsabilité des parents sur la scène scolaire

Thèmes	Pas d'accord	Position mitigée	D'accord
Donner aux parents une part de responsabilité et donc une part de décision dans la définition du projet pédagogique et des objectifs de l'établissement fréquenté par leur enfant	Enseignants	Directions	3Parents et Inspection

Tableau 29 - Scénarios relatifs à la présence des parents aux activités de l'école

Renforcer le lien entre l'école et les familles en organisant des rencontres et des échanges au sein même de l'école	------	------	Tous les acteurs
Faciliter la mobilisation et la participation des parents aux activités de l'école	------	------	Tous les acteurs
Faire l'inventaire en début d'année de la disponibilité, des talents et des centres d'intérêt des parents et organiser des activités à l'école ou en dehors impliquant les parents	------	------	0Tous les acteurs
Organiser des activités à caractère social qui rassemblent les élèves, le personnel scolaire, les familles, les partenaires sociaux et les citoyens du quartier	Enseignants du secondaire	------	Parents Enseignants du primaire
Organiser des activités en parallèle pour les élèves et les parents afin de montrer aux élèves l'importance que leurs parents accordent à l'école	------	------	1Tous les acteurs

Tableau 30 - Scénarios relatifs à la communication et aux échanges école-famille

Mettre en place des moyens de communication plus efficaces et plus variés entre l'école et les familles	------	------	2Tous les acteurs
Nommer dans chaque classe un délégué des parents qui servirait de lien entre les parents et l'enseignant	Enseignants du primaire		Parents Enseignants du secondaire

Tableau 31 - Scénario relatif au degré d'information des parents

Informer les parents en organisant des réunions ou de conférences sur leur rôle et leur responsabilité par rapport à la scolarité de leur enfant (ex. : conseils en matière de communication parents-enfant, conseils aux devoirs, …)	------	------	3Tous les acteurs
Informer les parents de l'existence d'autres partenaires ou organismes de leur quartier susceptibles de les aider (école de devoirs, centre de ressources, …)	------	------	4Tous les acteurs

On constate à la lecture des différents tableaux que sur les dix propositions, sept reçoivent l'assentiment de la majorité des acteurs interviewés. Il semble que l'idée de renforcer les échanges entre l'école et la famille et de mettre en place des moyens de communication plus efficaces et plus variés soit acceptée favorablement. Cette intensification passerait par l'organisation de rencontres parents-enseignants, la participation des parents à des activités organisées par l'école voire l'implication de ces derniers aux activités scolaires en fonction de leur disponibilité, talents ou centres d'intérêt mais aussi via l'enrichissement des informations transmises (conseils en matière de communication parents-enfant, conseils aux devoirs, …) et la

multiplication des voies utilisées pour véhiculer ces données (écoles de devoirs, centre de ressources, …)

Certaines propositions opposent cependant les acteurs en présence. Ainsi, accroître la part de responsabilité des parents et donc, la part de décision de ces derniers dans la définition du projet pédagogique et des objectifs de l'établissement fréquenté par l'enfant divise les enseignants (plutôt contre l'idée), les directions (position mitigée) ainsi que les parents et les membres de l'Inspection (pour un tel changement). Les parents, jusqu'ici tenus à distance de ce genre d'investissement, se disent prêts à réfléchir à certains sujets pédagogiques avec les enseignants. Or ces derniers se montrent massivement frileux à l'égard de cette proposition.

L'organisation d'activités à caractère social qui rassembleraient les élèves, leurs parents, l'équipe éducative mais également des partenaires sociaux reçoit un accueil favorable auprès des parents et des enseignants du primaire. Les enseignants du secondaire se disent réticents à la mise sur pied de telles animations. Peut-être est-ce le manque d'habitude de ces démarches qui poussent les enseignants du secondaire à réagir ainsi ? Les enseignants du primaire, plus coutumiers à ce genre d'événements, sont par contre tout à fait preneurs de ce type d'initiatives.

La proposition de nommer au sein de chaque classe un délégué de parents susceptible de faire le lien entre les parents et les enseignants est appréciée par les parents et les enseignants du primaire mais ne l'est pas par les enseignants du secondaire. Peut-être les enseignants du primaire, responsables, dans la plupart des cas, d'une même classe du matin au soir craignent-ils d'être « envahis » par les parents ? Au niveau du secondaire, on peut parler d'équipe éducative – une même classe est gérée par un ensemble d'enseignants – et donc sont peut-être moins effrayés par ce genre d'initiative puisqu'ils peuvent répondre de façon consensuelle et groupée aux demandes parentales.

Pour conclure, nous rappellerons que la majorité des propositions ont été évaluées de façon positive par la majorité des répondants. Dans tous les cas, elles ont reçu l'accord massif des parents. Il semble donc que la mise en place de projets susceptibles de rapprocher l'école et la famille et partant, de renforcer la communication et les rapports entre les deux principaux éducateurs de l'enfant, trouverait un terrain fertile dans les classes luxembourgeoises. Contrairement aux *a priori* couramment répandus, nous ne pouvons pas parler de désinvestissement des parents à l'égard de l'éducation scolaire de leur enfant. Au contraire, les résultats obtenus attestent d'une volonté certaine de répondre aux appels de l'institution scolaire et des enseignants et de se montrer actifs sur la scène scolaire.

Références bibliographique

Bogdanowicz, M. (1997). *La place des parents dans les systèmes éducatifs de l'Union européenne*. Site de Eurydice – Le réseau d'information sur l'éducation en Europe, [En ligne]. http://www.eurydice.org/Documents/Parents/fr/FrameSet.htm (Page consultée le 19 septembre 2001).

Conseil Supérieur de l'Éducation (1994). *Être parent du primaire : une tâche éducative irremplaçable*. Québec : Ministère de l'Éducation.

Desmet, H., Nimal, P. et Pourtois, J.-P. (1999), Les partenaires de l'Éducation. Famille-Ecole-Société : un contrat social et pédagogique qui se bâtit, *La Nouvelle Revue de l'AIS*, 7, 49-60.

Règlement grand-ducal du 3 mai 1989 fixant la tâche des enseignants de l'éducation préscolaire et de l'enseignement primaire. Site de Technolink - Centre de coordination du Service de l'Enseignement pour l'introduction progressive des TIC dans les classes préscolaires et primaires de la Ville de Luxembourg [En ligne] http://www.technolink.lu/teachers/comite/reglement/cogestion/reglement_tache.pdf (Page consultée le 30 août 2006).

MENFPS, SCRIPT (2003), Les chiffres clés de l'Education nationale : statistiques et indicateurs : année scolaire 2002-2003. Site du SCRIPT [En ligne] http://www.script.lu/documentation/publication_indicateurs_statistiques_chiffres_cles.phtml (page consultée le 7 septembre 2006).

Saint-Laurent, L., Royer, E., Hebert, M. et Tardif, L. (1993). La réussite scolaire et la collaboration entre l'école et la famille, *Bulletin du CRIRES*, 1, 8 pages. Site de l'Université de Laval [En ligne] http://www.ulaval.ca/crires/pdf/CRIRES1.pdf (Page consultée le 30 août 2006).

Terrail, J.-P. (1997). La sociologie des interactions famille/école, Sociétés Contemporaines, 25, 37-83.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

METHODOLOGIE D'EVALUATION DES COMPETENCES DE CONDUITE D'UN GROUPE D'ANALYSE DE PRATIQUES

Richard ÉTIENNE, Équipe CERFEE, Master *Conseil et Formation en Èducation*, directeur du département de sciences de l'éducation, université Montpellier 3

Mots-clés : pratiques, analyse, groupe.

Résumé:
Le recours de plus en plus fréquent à l'analyse de pratiques pour la formation s'explique par la relative faillite de l'ingénierie de formation dont les analyses de besoins ne prennent pas en compte les personnes. Elle pose un problème d'évaluation, ne serait-ce que pour en assurer le déploiement. Pour le traiter, nous nous appuyons sur l'expérience de deux unités d'enseignement d'un Master de formation de formateurs visant à développer la « *pratique réflexive dans le métier d'enseignant* » (Perrenoud, 2001)… et de formateur ! Nous faisons le point sur les principaux obstacles rencontrés et nous analysons les nouvelles pratiques d'évaluation co-construites dans ce Master.

La règle de confidentialité s'oppose à la communication de ce qui se passe dans le groupe ; elle garantit le droit à l'erreur de chacun et crée un espace propice au développement réflexif. Cette contrainte impose un secret professionnel sur le contenu des échanges mais aussi sur l'atmosphère des séances. L'absence de jugement est une autre constante qui vise l'autonomisation des participants et rend délicate la certification des animateurs. Or, c'est à travers eux que les dispositifs sont jugés ; leur défaillance les discrédite. Nous avons dû réfléchir à l'évaluation formative de ces compétences et décider de leur validation. Une tension forte existe entre leur présentation et leur développement progressif et le contrôle de leur présence que traduit une note.

La discussion porte sur les effets de normes universitaires dans un cadre d'éducation des adultes : réservée à des initiés qui prenaient le temps de se former, l'analyse des pratiques se tenait loin de la frénésie certificative. Aujourd'hui objet d'injonctions institutionnelles, elle tarde à affronter les difficultés induites par les compromis d'une évaluation « *professionnalisée* » dans le contexte d'une présence de plus en plus marquée de l'université sur le champ de la formation continue.

Introduction

Le recours de plus en plus fréquent à l'analyse de pratiques pour assurer ou structurer la formation initiale ou continue des enseignants mais aussi des travailleurs des « *métiers de l'humain* » (Blanchard-Laville & Fabre, 1996 ; Fumat, Vincens & Étienne, 2003 ; Giust-Desprairies, 2005) s'explique par la relative faillite de l'ingénierie de formation : les analyses de besoin ne tiennent généralement pas compte des différences interpersonnelles ; elles décrivent un sujet à former sans tenir compte de son histoire ou de son expérience. L'analyse de pratiques, malgré ses

formes très variées (Étienne, 2004), prend en compte l'histoire et le cheminement de la personne dans la mesure où elle fait l'hypothèse d'interactions entre l'expérience de chacun et son activité professionnelle. Ce type de dispositif pose donc inéluctablement le problème de son évaluation (avec un double objet : l'efficacité ou l'efficience de cette modalité et les compétences individuelles ou collectives développées), ne serait-ce que pour en assurer le déploiement. Pour le traiter en cherchant à ne pas réduire sa complexité, nous nous appuierons sur la mise en place de deux unités d'enseignement d'un Master professionnel de formation de formateurs visant à développer la « *pratique réflexive dans le métier d'enseignant* » (Perrenoud, 2001)… et de formateur ! Intitulé *Conseil et formation en éducation*, il existe depuis septembre 2001 et a formé cent cinquante professionnels de la formation d'enseignants, de personnels de santé, du social et d'autres personnels comme les officiers d'infanterie. En utilisant la première personne du pluriel, nous sommes porteur de l'expérience de ce groupe stable de formateurs et d'enseignants puisque seul Marc Durand nous a quittés pour rejoindre l'université de Genève et il a été remplacé par Dominique Bucheton, responsable du LIRDEF et co-responsable du Master.

Nous commencerons par faire le point sur les principaux obstacles inhérents au dispositif lui-même puis nous analyserons l'aspect problématique et provisoire de nouvelles pratiques d'évaluation co-construites par le groupe de pilotage de la formation et les étudiants concernés. Enfin, nous mettrons en discussion les points les plus délicats liés à la prise en charge par l'université d'une formation qui échappait à la certification.

Les obstacles inhérents aux valeurs et références théoriques de l'analyse de pratiques

L'analyse de pratiques peut être considérée comme une galaxie dans laquelle plusieurs typologies sont possibles, qu'elles soient fondées sur le support (écrit, vidéo, audio, oral, voire un mixte), sur le nombre de participants (un seul, deux, un groupe), sur les finalités (recherche, formation, échange) ou sur les références (psychologiques, pédagogiques, sociologiques, multiréférentielles). Dans une communication aussi courte, nous insisterons surtout sur les noyaux communs et nous viserons la mise en évidence d'obstacles inhérents à ce genre de dispositif, marqué par son histoire mais aussi lié à l'hypothèse fondatrice de son efficacité. Nous en avons repéré deux : la confidentialité et l'absence de jugement ; ils rendent difficile, voire impossible, la communication de l'évaluation et son élaboration sous forme d'un jugement de valeur si l'on se réfère à Barbier (1984, 64) : « *(…) l'évaluation se manifeste donc, socialement et historiquement, comme un ensemble d'actes, de pratiques, aux formes certes variées mais aboutissant tous à la production de jugements de valeurs (…)* ». Nous ne traiterons pas ici du troisième obstacle qui fait du volontariat la clé de la participation ; il est mis à mal dans les formations professionnelles obligatoires mais nous l'aborderons au moment où nous établirons les fragiles équilibres liés à l'attribution de notes pour cet enseignement dans le cadre universitaire d'un Master.

La confidentialité ou la mise en sécurité des personnes et du groupe

La règle première et la plus fréquente de l'analyse de pratiques est la confidentialité. Difficile à définir, elle se distingue du *secret* dans la mesure où ce dernier aboutit à l'enfermement définitif d'une partie retranchée d'un tout : *secretum* vient de *secerno* qui veut dire couper et mettre à part ; sécession est de la même famille. La confidentialité repose sur la notion de confiance entre les participants à l'analyse (et ce quel que soit leur statut par ailleurs) et cette règle s'oppose à la communication extérieure de ce qui se passe dans le groupe ; elle garantit le droit à l'erreur de chaque participant, source de son développement dans les théories socio-constructivistes, et crée un espace de sécurité propice à un épanouissement de sa réflexivité. Héritée des groupes Balint (1957) et de la non directivité (Rogers & Pagès, 1968), cette contrainte impose une sorte de secret professionnel sur le contenu des échanges mais aussi sur l'atmosphère des séances. Les étudiants sont très rapidement informés de ce qu'ils considèrent comme une contrainte avec laquelle il faudra composer. Dans le groupe, la confidentialité a pour effet secondaire d'établir un savoir vivre ensemble qui n'est pas la moindre des qualités recherchées dans une promotion qui élabore, grâce à cette norme, un sentiment de sécurité et qui contrebalance bien des porosités de notre univers actuel. Toutefois, au moment de l'évaluation, à laquelle on ne peut pas ne pas se soumettre, ainsi qu'à l'occasion de l'écriture de l'ouvrage (Fumat, Vincens & Étienne, 2003), nous avons rencontré la contrepartie de cet enfermement dans le groupe d'entraînement à l'analyse de situations éducatives (GEASE) qui est notre dispositif de référence, même s'il n'est ni unique ni monopolistique.

Les récits, les sentiments et les difficultés évoquées par les narrateurs, même si nous leur demandons d'évoquer une situation qui n'est pas forcément difficile, n'ont pas à être connus de tout le monde ; a fortiori, les employeurs, même s'ils sont commanditaires de la formation à et par l'analyse de pratiques, restent dans l'ignorance de ce qui s'échange dans l'analyse de pratiques et nous sommes vigilants pour éviter tout rapport hiérarchique au sein du groupe, quand bien même il n'est réuni qu'une fois, pour une simple initiation. La confiance ne se décrète pas et cette obligation de confidentialité est bien un obstacle pour la validation des compétences acquises par la participation ou l'animation d'un groupe d'analyse de pratiques. Nous avons utilisé les ressources de la recherche en sciences humaines pour préserver l'anonymat des étudiantes et étudiants, de formation initiale ou continue. Les narratrices de la discrimination et de l'exclusion (*op. cit.*, 66-94) ont donné leur accord pour parution mais elles ont revu l'ensemble du document, ont choisi leur nouveau nom et ont négocié des modifications permettant de ne pas les identifier.

Le développement exponentiel de ces dispositifs ne permet sans doute pas à tous les formateurs-animateurs des groupes de marquer une attention rigoureuse à cette protection de la confidentialité qu'il est nécessaire de voir comme un élément essentiel au fonctionnement du dispositif quand bien même les responsables institutionnels font pression sur eux pour obtenir des éléments susceptibles d'éclairer les instances décisionnelles. Imaginons (cas vécu) un groupe dit d'analyse de pratiques dans lequel l'inspectrice de la circonscription somme une enseignante d'analyser les raisons de son échec en enseignement de la lecture au cours préparatoire. En revanche, il faut avouer que le souci d'en savoir le plus possible

avant de titulariser un fonctionnaire pour quarante ans est légitime mais tout autant légitime est la fin de non recevoir réservée à cette demande par tout spécialiste de la question.

L'absence de jugement ou le renoncement à l'évaluation immédiate

Dans un groupe d'analyse de pratiques, la première pente naturelle que va essayer de faire remonter le formateur-animateur est la propension au jugement et au conseil qui caractérise le monde de l'enseignement. À l'intérieur du groupe, la confidentialité fait l'objet d'un consensus tel qu'elle est oubliée s'il n'y a pas une ritualisation qui permet de séparer les phases de travail en signalant qui a droit à la parole ou pas, en sachant qu'aucune interprétation n'est plus vraie qu'une autre. En revanche, sitôt exposée une situation, parfois même avant la fin du récit, fusent des expressions du type : « *moi, à ta place, j'aurais…* ». C'est ce que nous essayons d'éviter car, pour reprendre Lacan, si nous nous mettons à la place de l'autre, quelle place lui laissons-nous ? Aucune ! Nous repartons dans notre prurit de jugement à l'emporte-pièce. Et nous ne constituons pas les bases de toute évaluation qui se manifeste toujours comme un jugement de valeur, mais la valeur en moins si on nous accorde ce jeu sur les mots.

Depuis les histoires de vie jusqu'à la visite de classe, toute analyse de pratiques suppose de réunir un *référé*, donc de se mettre en position de recueillir par l'observation ou l'entretien des faits sur lesquels portera le travail d'analyse. C'est là, en partie que le bât blesse : l'ambiguïté du langage mise en évidence par les linguistes interdit de prendre les mots pour la réalité (Jacques, 1985) et les points de vue différents font que jamais un objet unique ne remplacera les visions que nous en avons (Husserl, 1950). Il n'y a donc pas de vérité de la situation mais des situations qui sont *situées* par celui qui les narre et en fait des objets dont la décomposition (l'analyse) devient l'objectif de travail du groupe. S'interdire le jugement pendant ce temps n'est pas renoncer à tout jamais au jugement ; bien au contraire, c'est établir une priorité, une antériorité à l'écoute et au recueil d'information avant de passer à une seconde phase de travail toujours préalable au jugement : la construction du référentiel.

Nous pouvons nous inspirer du plan de Figari (1994, quatrième de couverture) et faire de l'analyse de pratiques, quand elle se soumet à l'interdiction de jugement, un prolégomène à tout procès en éducation « *ce qui implique de comprendre la notion de "référent" (1ère partie) ; d'élucider le référentiel d'une situation donnée (2ème partie) ; d'interroger les approches du diagnostic et de l'audit (3ème partie) ; de disposer d'une méthode de construction d'un référentiel d'évaluation (4ème partie)* ». Comme nous nous reportons expressément à Ardoino (1992), nous pouvons avancer que les efforts du groupe et du narrateur ainsi que du formateur-animateur convergent vers la construction d'un système complexe, *hic et nunc*, de références qui permettront de commencer à comprendre et à interpréter la situation. Si la référentialisation est nécessaire pour l'évaluation, elle se révèle indispensable pour l'action et nous insistons sur le fait que nous nous livrons à un entraînement, ce qui est radicalement différent de la performance mais qui la prépare efficacement (*op. cit.*, p. 16-17 et 113-119).

Obstacle pour les personnes pressées d'émettre des jugements, l'interdiction de jugement réalise le tour de force de les faciliter en les prohibant aussi bien du côté du référé qui est patiemment construit dans les premiers temps de travail (très souvent par le récit et par le questionnement du narrateur) que du référent qui se vit sur un mode dynamique (la référentialisation) et complexe (la multiréférentialité). Il faut reconnaître qu'en l'état actuel l'absence de jugement est une constante de cette forme de travail ; bien qu'elle vise l'autonomisation des participants au groupe, elle rend délicate la certification des animateurs de ce type de dispositif que pour notre part nous considérons comme des formateurs, plus par leurs résultats que par leur mode d'action. Or, c'est à travers ces animateurs que les dispositifs sont évalués par les participants ou les institutions de formation. Leur défaillance ou simplement l'image qu'ils transmettent de leur intervention peut entraîner la mise à l'index d'une méthode longuement mise au point par ses concepteurs : « *c'est trop psy.* », « *on doit nous évaluer ; on ne voit pas pourquoi on donnerait de nous-mêmes des informations sur nos difficultés* ».

Le cadre de la recherche sur la création d'un dispositif d'évaluation respectant de fortes contraintes

Nous avons retrouvé ces difficultés dans notre pratique de formation de formateurs. Elles étaient fortement atténuées du fait du volontariat des personnes inscrites dans notre Master *Conseil et Formation en Èducation*. Nous nous attendions à une interrogation sur les épreuves et sur les modalités de notation mais en même temps nous avons voulu nous emparer de ce chantier pour en faire une recherche sur deux points que nous aborderons successivement ; en premier lieu, l'innovation dans le domaine des épreuves universitaires : comment imaginer et mettre en place des épreuves certificatives qui attestent de compétences à décrire puis analyser une situation éducative ? Comment certifier la présence de compétences permettant au sein d'un travail collectif de piloter et animer une analyse de situation de formation ? En second lieu, nous tenons à faire émerger des dispositifs de recherche utilisant l'ensemble des dispositifs d'analyse de pratiques selon un va-et-vient que l'on peut observer chez Vermersch (1994 & 1997). Depuis plusieurs années, nos étudiantes et étudiants s'inscrivent dans un ensemble de travaux qui visent surtout le travail enseignant, particulièrement en éducation prioritaire. Ils sont donc associés aux travaux de nos deux équipes de recherche (Laboratoire Interdisiplinaire de Recherche en Didactique, Èducation et Formation – LIRDEF – de l'IUFM de Montpellier et de Montpellier 2 et Centre de Recherches sur la Formation, l'Éducation et l'Enseignement – CERFEE – de l'université Montpellier 3). Plusieurs sont en thèse et participent à une Équipe de Recherche Technologique de Montpellier 3 sur le début de cours.

Des épreuves certificatives qui s'éloignent de l'isomorphisme pour s'approcher de l'activité

Au premier semestre, l'objectif est de développer les compétences de participation à un groupe d'analyse de situations éducatives. La didactisation du dispositif GEASE

nous a conduits à abandonner cette représentation selon laquelle pour apprendre à participer à un GEASE, il suffit d'y participer ! Nous ne traiterons ici que le volant concernant l'épreuve sommative puisque nous avons été confrontés à une situation relativement inédite : attester des compétences en place alors que le dispositif est dynamique et peut se poursuivre *sine die* dans une perspective de formation continue : un GEASE de psychologues scolaires animé par Guy Vincent se réunit régulièrement depuis dix-neuf ans. Nous étions en présence d'un des impératifs de la semestrialisation : noter au bout de treize séances de trois heures ce premier ensemble de compétences. Nous avons choisi une rupture avec les modalités de déroulement habituel de l'enseignement. La métaphore de l'entraînement ne nous semblait pas autoriser une approche isomorphique qui aurait consisté à noter une ou plusieurs séances de GEASE. Nous avons décidé de procéder à une épreuve au cours de laquelle l'étudiante ou l'étudiant ferait une performance : présenter une situation d'éducation ou de formation de son choix et l'analyser en présentant la ou les théories de références mises à contribution. Les premières épreuves ont établi qu'une liste de critères apparaissait indispensable aux intéressés pour éviter de tomber dans un sentiment d'arbitraire de la note attribuée. Depuis deux ans, nous avons mis au point un fiche-navette présentée en figure 1. En cas d'incompréhension ou de demande d'explications complémentaires, les membres du jury discutent avec l'étudiante ou l'étudiant à l'aide de cette fiche et du document remis au début de la prestation.

Master CFE
Épreuve d'analyse de situation
Recommandations et critères

Deux membres du jury – Salle **** – L'oral est public mais il de tradition de demander l'accord du candidat. Les candidats et le public quittent la salle lors de la délibération. Les notes sont communiquées après une nouvelle délibération faite à l'issue des deux journées.

Présentation : durée 15 minutes
Vous écrirez la situation choisie pour l'épreuve sur une page au maximum en lui donnant un titre. Ce récit sera remis aux examinateurs avant le début de l'épreuve. La présentation sera faite sous forme d'un récit oral et vous ne disposerez d'aucun document pendant l'exposition. Pour l'interprétation et pour elle seule, vous pourrez vous aider de vos notes écrites.

Discussion : durée 10 minutes
Vous répondrez brièvement aux questions posées par les deux membres du jury

Délibération : durée 5 minutes
Les deux membres du jury s'accordent sur une note provisoire qui sera révisée à l'issue des deux journées.

CRITÈRES : NOM Prénom :

Formalisation des critères	Commentaires
1. Qualité de la présentation orale (utiliser les critères de la communication) 2. Singularité et description de la situation (éviter la présentation d'un cas ou le récit autobiographique) 3. Lien entre la situation exposée et les hypothèses avancées dans un cadre théorique de votre choix 4. Lien avec la théorie assuré à l'aide de deux ou trois références bibliographiques précises (pages ou chapitre ou article) pour les principaux concepts utilisés lors de l'analyse 5. Reprise de l'analyse dans une conclusion synthétique dégageant quelques axes de travail de formation (de deux à quatre)	
NOTE :	

L'évaluation est réalisée par les deux membres du jury sous la forme d'une note qui sera communiquée par écrit avec un rapide commentaire global (cf. ci-dessus). L'utilisation des critères se fait sous forme synthétique et ne recourt pas à un barème pré-déterminé.
Figure 1 – Fiche-navette pour l'évaluation du premier semestre

Au second semestre, nous avons à répondre d'une certification plus globale ; si l'étudiante ou l'étudiant doit savoir animer un groupe d'analyse de pratiques et problématiser ou faire problématiser des évolutions, des améliorations du dispositif mais aussi de sa conduite, cette compétence se manifeste au sein d'une activité qui combine deux autres éléments que nous qualifierons provisoirement de repérage politique et institutionnel de l'action de formation dans son contexte d'une part et d'élaboration méthodologique se conformant aux théories et modèles de formation actuellement en œuvre d'une autre part. C'est la raison pour laquelle nous nous sommes efforcés d'imaginer et de faire vivre une situation professionnelle la moins artificielle possible et la plus proche de données de recherche. Schématiquement, nous donnons environ six semaines à nos étudiantes et étudiants groupés par affinités par quatre ou cinq pour analyser une situation de formation enregistrée sur un support de notre choix et proposer une problématisation de préférence à des recommandations proches d'un audit (figure 2).

Évaluation des trois UE 7, 8 et 9 du Master *Conseil et formation en éducation*

But de l'évaluation : attester les compétences individuelles et collectives d'étudiants du Master à présenter une analyse problématisante d'une situation de préparation, de réalisation, de régulation ou d'évaluation d'une action de formation et à tenir une discussion professionnelle sur elle en dégageant les questions qu'elle soulève.

Principe : un corpus est remis à cinq étudiants (l'harmonisation peut imposer la création de groupes de trois ou de quatre) au moins deux semaines à l'avance et les étudiants en présentent **une analyse individuelle** qui est suivie d'une **mise en discussion collective** et close par une conclusion conduite par chaque étudiant. La séance est enregistrée et donne lieu à une note attribuée par les trois responsables des unités d'enseignement concernées. Chaque étudiant obtient une note qui est répliquée pour chaque unité d'enseignement. **Si un étudiant n'assiste qu'à une partie des enseignements, il n'est noté que sur elle.**

Déroulement précis :

J-45 : distribution par tirage au sort des corpus (verbatims mais pas exclusivement, on peut aussi remettre la vidéo ou l'enregistrement audio) aux cinq étudiants qui se sont regroupés. Les enseignants peuvent modifier la composition des groupes.

J+H : les étudiants présentent successivement leur analyse de la situation en 8 mn avec Powerpoint ou un autre logiciel de Présentation Assistée par Ordinateur. À l'issue de ces présentations, les membres du jury lancent la discussion professionnelle qui dure trente minutes. À la fin de cette discussion, chaque étudiant dispose d'une minute pour présenter ce qu'il retire de l'analyse collective, notamment en termes de déplacement de son point de vue. Les étudiants se retirent et les membres du jury délibèrent sur l'attribution d'une note provisoire attribuée à chaque étudiant.

Conduite du dispositif :

Les trois membres du jury déterminent et présentent le rôle qu'ils vont assumer : **l'animateur** distribue la parole et effectue les relances, **le preneur de notes** met en mémoire écrite ce qui se dit et **l'observateur** s'intéresse à la dynamique verbale et non verbale du déroulement de l'épreuve. Les rôles tournent à chaque nouvelle discussion professionnelle. Une harmonisation des notes est pratiquée après l'épreuve.

Critères d'évaluation :

1. Capacité à conceptualiser et verbaliser les jeux et les enjeux de la situation de formation

2. Référence à des savoirs théoriques issus de l'enseignement et/ou des lectures personnelles
3. Aptitude à mettre en synergie les multiples approches conceptuelles de la formation
4. Capacité à conduire individuellement et collectivement une analyse d'une situation de formation
5. Le savoir questionner et le savoir problématiser ne seront pas évalués du point de vue de la résolution mais de la mise en discussion
6. Les capacités de communication sont évaluées (parler sans lire un papier, regarder son jury).

Figure 2 – Protocole d'évaluation du second semestre

Stabilisées depuis deux ans, ces modalités d'évaluation nous posent encore des questions liées à leur trop grande parenté avec les pratiques universitaires traditionnelles. Nous avons décidé d'en faire des objets de recherche et d'enregistrer les épreuves pour donner lieu à des autoconfrontations simples et croisées (Clot *et al.*, 2001) dont l'intérêt premier est de construire le métier, ici d'évaluateur mais aussi d'étudiant, sur la base d'une analyse de l'action, donc *a posteriori*, dans les écarts, là où se situent les routines que l'on voudrait modifier pour améliorer ses performances.

Une recherche sur la formation qui vise une double interaction

En France, le bon professionnel est censé devenir un bon formateur et seul le Conservatoire National des Arts et Métiers (CNAM) peut faire état d'une chaire (Guy Jobert) et d'un Centre de Recherches sur la Formation (Jean-Marie Barbier). Or, plus la formation devient un enjeu et une source de dépenses minutieusement contrôlées, moins cette *tabula rasa* se justifie. Nous essayons donc de faire progresser l'idée qu'il existe des formateurs-chercheurs tout comme nos universités sont peuplées d'enseignants-chercheurs et nous n'essayons pas d'opposer les deux. C'est à l'occasion du mémoire que les étudiantes et les étudiants font le choix d'une recherche ou d'une étude. Les normes et critères diffèrent alors. Nous ne parlerons ici que du premier choix qui est en rapport direct avec notre sujet. Cette question a été étudiée avec Lorraine Savoie-Zajc, professeure de l'Université du Québec en Outaouais (UQO) invitée en mai 2005 par notre département de sciences de l'éducation. Les étudiants lui ont d'abord présenté les compétences nécessaires à l'analyse de situations (figure 3), recensées par eux et lui ont demandé d'animer une séance de travail au cours de laquelle ces compétences se retrouveraient dans le cadre d'une action (de formation)-recherche que nous préférons appeler dans le cadre français une recherche sur l'action (de formation). Étant entendu que dans tous les cas, cette recherche se retrouverait dans la littérature scientifique (production et communication de savoirs) et professionnelle (partage de savoirs d'action) et qu'elle se soumettrait à des boucles de rétro-action, voire à une récursivité liée à l'inventivité professionnelle des praticiens (gens de métier, formateurs et chercheurs).

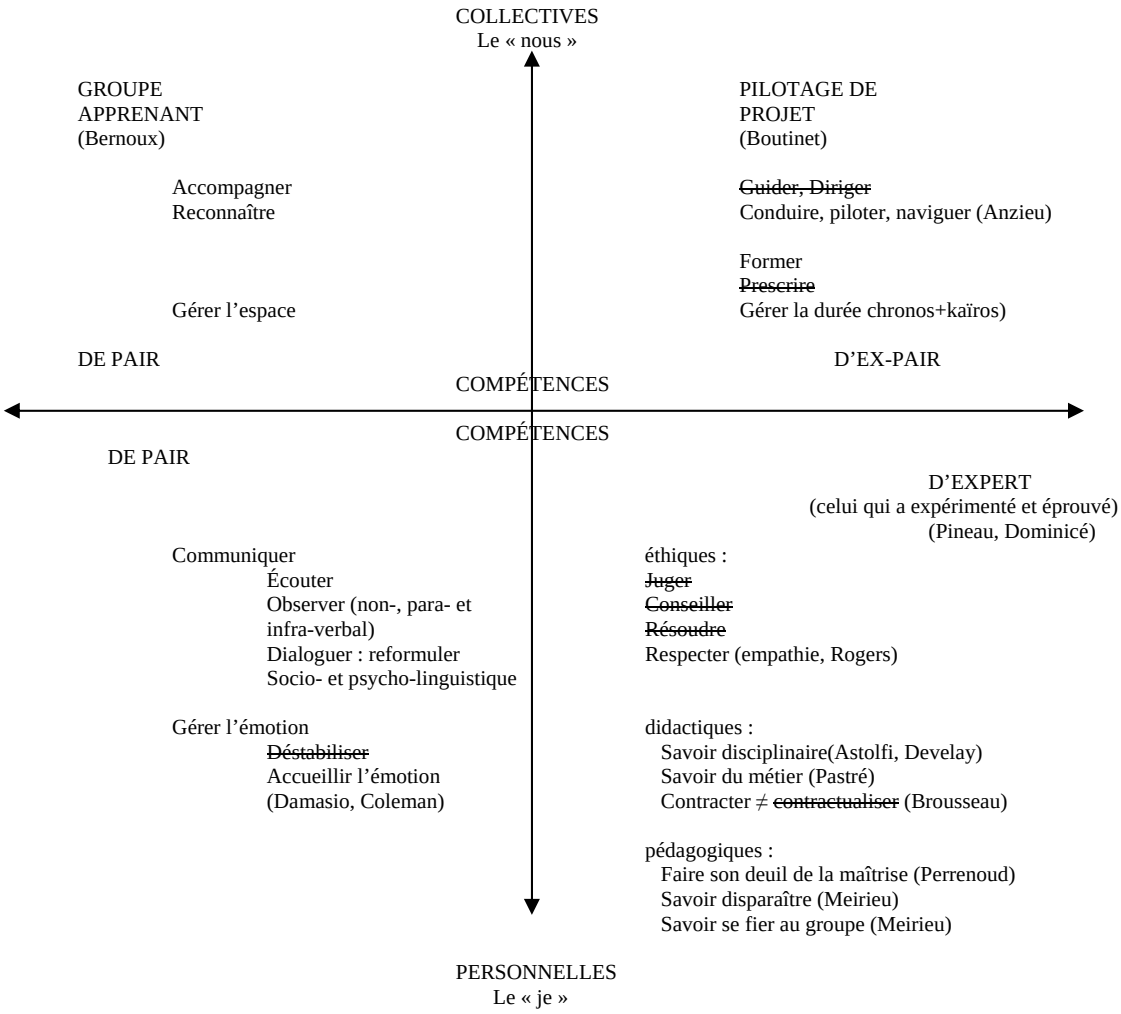


Figure 3 – Le réseau de compétences personnelles et collectives (Le Boterf, 1994) à l'œuvre dans l'analyse de situations

En réponse à cet exposé et à la nécessité de développer des recherches sur l'action, Savoie-Zjac (à paraître) a développé ce qu'elle expose dans un ouvrage issu des dernières rencontres du Réseau Éducation et Formation (REF) à Montpellier ; « *Il apparaît donc essentiel, dans une démarche visant à accompagner des ajustements de pratiques, de fournir aux participants l'espace nécessaire de réflexion, de discussion dans un climat ouvert et sécurisant pour que chacun puisse communiquer et partager avec les autres les résultats de ses expérimentations et les réflexions suscitées. C'est grâce à la réflexion et au partage avec les autres membres, que chacune des personnes sera en mesure de donner du sens aux changements attendus. Celui-ci sera communiqué, discuté, confronté, qualité essentielle pour un développement*

professionnel de qualité (Van den Berg, 2002), selon une visée socio-constructiviste (Windschitl, 2002), visant la croissance de la personne (Huberman & Guskey, 1995), perspective tout à fait cohérente avec la notion de communauté d'apprentissage, ajouterions-nous ». De fait, nous avons la confirmation qu'un vaste chantier s'ouvre pour l'évaluation : chantier heuristique d'abord puisqu'il va être question d'évaluer l'accompagnement, les ajustements de pratiques, la réflexion, le sens et le développement. Vaste programme ! Cette communication a donc pour but, dans sa dernière partie, de mettre en discussion les modalités de l'amélioration des pratiques liées à leur analyse par les acteurs, les chercheurs et les formateurs et son indispensable évaluation pour laquelle tout, ou presque, est à reprendre au début.

Un équilibre précaire et provisoire qui nécessite d'anticiper sur des développements ultérieurs

Les questions qui sont actuellement les nôtres portent en priorité sur les effets induits par l'introduction de normes universitaires dans un cadre initialement marqué par les valeurs égalitaires de l'éducation des adultes : tant que l'analyse des pratiques n'était pratiquée que par quelques groupes d'initiés qui prenaient le temps de se former progressivement et selon une méthode d'imprégnation douce, personne, et surtout pas ses thuriféraires, ne la mettait en cause au regard de ses résultats. Aujourd'hui, la formation devient le défi du vingt-et-unième siècle et l'analyse des pratiques est l'objet d'injonctions institutionnelles ou de demandes de professionnels désireux d'échanger sur leur métier mais aussi de déployer une analyse qu'ils espèrent réinvestir dans l'action. Dès lors, il importe de débattre dans la communauté scientifique et, plus précisément, dans le cadre de l'ADMEE, des moyens de redéployer l'évaluation pour qu'elle ne s'adapte plus seulement aux résultats, aux produits, mais s'intéresse à la dynamique des situations et du langage, qui est à l'origine de ces résultats. Pour nous, cette dynamique nécessite d'élaborer une réflexion épistémologique et méthodologique à deux niveaux, le premier, et le plus trivial, étant la formation des formateurs dont l'effet démultiplicateur mériterait une attention toute particulière, le second, et le plus prometteur, étant la reconfiguration des systèmes et des modèles de formation pour aller vers une plus grande intégration action-recherche-formation.

Comment évaluer une formation de formateurs à l'analyse de pratiques

Nous sommes encore en recherche d'une structuration de notre formation qui respecte les normes universitaires et développe une approche de la professionnalité du formateur qui se consacre essentiellement à la gestion des *imprévus* et des *événements* qui mobilisent 85%, voire plus, de son temps (Étienne & Jean, 2006). Ce constat nécessite de partir du réel, c'est-à-dire de ce qui est reçu et perçu comme tel par les acteurs pour aller vers une déconstruction-reconstruction de ce réel. Nous pensons avoir fourni quelques indicateurs pour cela, notamment en nous attachant à la notion de *situation* qui est plus identifiable que celle de *pratiques*, dans la mesure où ces dernières sont engluées dans la situation et où le travail d'analyse consiste

précisément à détisser ce que le réel présente comme un enchevêtrement. Toutefois, nous éprouvons encore des difficultés et tâtonnons pour bâtir une certification qui ne soit pas trop éloignée de notre conception radicalement recentrée sur l'expérience, la personne et le groupe que nous ajoutons et combinons à la *tâche prescrite* qui demeure mais fonctionne comme un miroir aux alouettes. En effet, le résultat (le produit ou le service) ne peut, par son seul démontage, fournir le secret de fabrication (le *savoir faire* ou l'*habitus* professionnel ou encore le *tour de main*). Nous imagions aisément que la première tâche d'une formation professionnelle universitaire est de se donner et de donner à son public les clés de l'évaluation. Or, malgré les progrès de la *didactique professionnelle* (Pastré, 1997), les raisons qui expliquent les courbes d'efficacité favorables aux vétérans restent mystérieuses, elles ne sont pas traitées en formation et peu en recherche, sauf en psychologie du travail. Comme universitaires, nous nous fixons comme horizon de recherche l'amélioration des procédures que nous avons mises en place, étant entendu que nous instituons par là-même notre travail d'enseignant et de formateur comme sujet de recherche. Si ce n'est pas original, nous pouvons affirmer que ce n'et pas si courant dans un cadre autre que celui de la thèse.

Comment intégrer l'évaluation dans une alternance intégratrice

La crise des modèles de formation a entraîné une sorte de retour à la case départ par réintégration du réel dans la formation. Il est vrai qu'il peut être convoqué sous deux formes : celle de la juxtaposition, ou pire de la simple application, dans laquelle des maîtres de la même appellation font faire par les élèves ce que leurs professeurs leur ont dit de faire. Nous nous intéressons plus particulièrement au second modèle plus pertinent parce qu'il intègre dans un même temps, celui de la réflexion, le souvenir de l'action menée individuellement ou collectivement, et les autoévaluations auxquelles se livrent les acteurs. La présence de chercheurs et de formateurs permet de ne pas enfermer les acteurs dans le solipsisme de leur analyse, dans un premier temps convenue et pauvre mais extrêmement riche et diversifiée quand elle bénéficie de l'effet-miroir que renvoient le chercheur ou le formateur ou, encore plus sûrement, le groupe de ceux qui pratiquent la même activité mais avec des nuances infinies d'exécution.

L'avantage de ces modalités d'évaluation par et avec les pairs en interaction avec des chercheurs et des formateurs mérite un temps d'arrêt : la parité (figure 3) permet l'examen de ce qui fait sens dans la réflexion et qui n'apparaissait pas dans l'action alors que l'expertise incite à trouver dans la théorie ce qui va éclairer le geste professionnel en montrant qu'il n'est pas mécanique mais issu du *bricolage* et du *braconnage* (Jorro, 2002) ou de *l'ajustement* (Savoie-Zajc, 2006). Si nous tenons depuis 1994 (Figari) une piste précieuse qui consiste dans la préférence marquée pour la *référentialisation* au détriment du référentiel, le chantier n'en reste pas moins impressionnant tant il va être indispensable de déterminer de nouvelles formes de formation et de nouvelles formes d'évaluation qui répondent à la convergence constatée entre l'approche nord-américaine du *praticien réflexif* et la vision européenne de *l'analyse de pratiques*.

En conclusion, il est tentant de se débarrasser du problème de l'évaluation des analyses de pratiques. Les premiers obstacles rencontrés incitent à baisser la pavillon évaluatif et à laisser l'évaluation *spontanée* ou *implicite* (Barbier, 1984) faire son œuvre. Ce ne sera pas au profit des praticiens de l'analyse de pratiques. Nous avons proposé une autre voie, celle qui conduit à *instituer* (*ibid.*) une évaluation spécifique aux analyses de pratiques en formation. Même si nous avons avancé quelques éléments d'action validés à l'heure actuelle, nous avons surtout essayé de poser la question de cette épistémologie et de cette méthodologie à la communauté scientifique, notamment à celle des chercheurs en évaluation dans la circonstance présente. C'est dire que cette démarche de recherche est possible. En revanche, l'on peut se demander aussi si elle est « *raisonnablement utile* ». Nous avons distingué deux niveaux d'utilité, même si, dans le fond, nous sommes convaincu de leur origine commune, la faillite des normes et des référentiels : le premier est celui de la certification des formateurs-animateurs de ces dispositifs. Vu les dégâts occasionnés à l'heure actuelle par les formateurs qui ne s'y sont pas formés et qui ne sont pas certifiés pour cette mission, nous pouvons avancer qu'elle est nécessaire et indispensable, faute de quoi l'analyse de pratiques rejoindra le cimetière des éléphants où gisent les bonnes réformes mal appliquées. Le second est beaucoup plus ambitieux et prospectif dans la mesure où il induit de véritables « *perspectives de changement* » (Étienne, 2005) qui font de l'analyse de pratiques, au sens large, le prétexte et le texte à venir de la formation, à commencer par celle des formateurs.

Références

ARDOINO, J. (1992). L'approche multiréférentielle (plurielle) des situations éducatives et formatives. *Pratiques de formation*, 15-34. Paris : Université Paris VIII.

BALINT, M. (Londres, 1957, Paris, 1960). *Le médecin, son malade et sa maladie*. Paris : PBP.

BARBIER, J.-M. (1984). *L'évaluation en formation*. Paris : PUF.

BLANCHARD-LAVILLE, Cl., FABLET, D. (1996). *L'analyse des pratiques professionnelles*. Paris : L'Harmattan.

BOTERF, G. le (1994). *De la compétence : essai sur un attracteur étrange*. Paris : Les éditions d'organisation.

ÉTIENNE, R. (2004). Dossier des *Cahiers du CERFEE*, n° 20 paru sous le titre « L'analyse des situations, des pratiques, de l'action et du travail en éducation et en formation dans la perspective d'une socialisation démocratique ». Montpellier : publication de l'université Paul Valéry.

ÉTIENNE, R. (2005). Perspectives de changement. *Cahiers pédagogiques*, n° 435, 31-33.

ÉTIENNE, R., JEAN, A. (2006). L'analyse de situations d'éducation et de formation entre la clinique et l'ingénierie. Questions actuelles sur l'invention d'une formation professionnelle reposant sur la singularité des situations étudiée en approche qualitative. Université Montpellier 3 : 1er colloque international sur la méthodologie qualitative pour la recherche en sciences humaines. Béziers, avec le concours de l'Association pour la Recherche Qualitative, les 27, 28 et 29 juin 2006.

FIGARI, G. (1994). *Évaluer : quel référentiel ?* Bruxelles : de Boeck.

FUMAT, Y., VINCENS, Cl., ÉTIENNE, R. (2003, rééd. en 2006). *Analyser les situations éducatives*. Paris : ESF.

GIUST-DESPRAIRIES, F. (2005, dir.). *Analyser ses pratiques professionnelles en formation*. Créteil : scéren-CRDP de l'Académie de Créteil.

HUSSERL, T. (1950). *Idées directrices pour une phénoménologie*. Paris : Gallimard.

JACQUES, F. (1985). *L'espace logique de l'interlocution*. Paris : PUF.

JORRO, A. (2002). *Professionnaliser le métier d'enseignant.* Paris : ESF.

PERRENOUD, Ph. (2001). *Développer la pratique réflexive dans le métier d'enseignant. Professionnalisation et raison pédagogique.* Paris : ESF.

ROGERS, C., PAGÈS, M. (1968). *Le développement de la personne*. Paris : Dunod.

SAVOIE-ZAJC, L. (à paraître). *L'accompagnement aux ajustements de pratiques professionnelles de personnels scolaires par la recherche-action. In* CROS, F. (dir.). *L'agir innovationnel: aux sources de la professionnalisation*. Bruxelles : de Boeck, perspectives en éducation et en formation.

VERMERSCH, P. (1994). *L'entretien d'explicitation en formation initiale et en formation continue*. Paris : ESF.

VERMERSCH, P., MAUREL, M. (1997). *Pratiques de l'entretien d'explicitation*. Paris : ESF.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

DRESSER LE PORTRAIT DES APPRENANTS POUR AMELIORER LEURS APPRENTISSAGES: UNE INTRODUCTION AUX MODELES PSYCHOMETRIQUES VISANT LE DIAGNOSTIC COGNITIF EN EDUCATION

Professeur André RUPP, Institut zur Qualitätsentwicklung im Bildungswesen

Nathalie LOYE, Université de Montréal

Depuis les quinze dernières années, le domaine des modèles à classes latentes restreintes (e.g., Nichols, Chipman, & Brennan, 1995) a connu un essor considérable aussi bien dans le domaine de la recherche théorique qu'appliquée. Ces modèles permettent de classer les apprenants relativement à un certain nombre de dimensions latentes. Ces dimensions sont habituellement appelées attributs cognitifs et sont supposées représenter les processus de réponse liés aux questions d'un test. Le portrait diagnostique établi pour chacun des sujets montre alors quels attributs il a déjà maîtrisés et quels attributs restent à maîtriser. Ces informations offrent donc la possibilité de déterminer la meilleure démarche que chaque sujet devrait adopter pour espérer maîtriser l'ensemble des attributs. Ainsi à partir de l'établissement de la structure des attributs d'un test, ces modèles donnent la possibilité de tester de nombreuses hypothèses à partir de données empiriques. Tous ces modèles cognitifs imposent d'établir au préalable la structure des attributs reliés au test. Cette structure prend la forme d'une matrice, notée Q (Tatsuoka, 1983), qui relie chaque attribut individuellement à chaque item. Les modèles permettent ensuite de spécifier si les attributs identifiés pour un item sont requis de façon conjonctive (i.e., s'il sont tous nécessaires pour obtenir une bonne réponse) ou de façon disjonctive (i.e., si au moins un de ces attributs est suffisant pour obtenir une bonne réponse). De plus, ces modèles estiment le degré d'incertitude grâce à plusieurs ensembles de paramètres. Bien que tous les modèles prennent en compte la possibilité de faire des erreurs d'étourderie (i.e., l'apprenant ne répond pas correctement malgré sa maîtrise des attributs relatifs à la question) et la possibilité de répondre par chance (i.e., l'apprenant répond correctement malgré un profil qui devrait donner une mauvaise réponse), chacun d'eux tient compte de ces comportements inattendus de façon différente. Par exemple, le modèle DINA (e.g., Junker & Sijtsma, 2001) exige l'estimation de deux paramètres concernant l'étourderie et de deux paramètres concernant la chance pour chaque item. Le modèle NIDA (e.g., Junker & Sijtsma, 2001) requiert leur estimation pour chaque attribut plutôt que pour chaque item. Et le modèle de Fusion (e.g., Hartz, 2002) impose deux paramètres pour chaque item associés à un paramètre pour chaque combinaison item-attribut, ce qui le rend beaucoup plus complexe que les deux premiers. En conséquence, ces modèles nécessitent des quantités de données variables afin de pouvoir estimer les paramètres. De plus le cadre probabiliste qui soutend ces modèles peut en général prendre en compte entre 4 et 8 attributs latents.

La vaste étendue des problèmes, aussi bien théoriques que pratiques, en lien avec ces modèles a conduit à une explosion du nombre de publications dans les revues scientifiques internationales et de conférences dans le cadre de congrès comme le NCME (National Council on Measurement in Education). En particulier, de nombreuses recherches visent à explorer les propriétés psychométriques liées à la robustesse des spécifications de la matrice Q (e.g., Jang & Roussos, 2006; Templin & Henson, 2006) et l'application de ces modèles à des données issues des évaluations à grande échelle (e.g., Ackerman, Zhang, Henson, & Templin, 2006; Rupp, Leucht, & Hartung, 2006). Pourtant, jusqu'à cette année, l'accès aux logiciels pour estimer les paramètres de ces modèles lors de leur application à des données empiriques était limité à un petit nombre d'experts de la mesure qui devaient soit écrire leur programmes eux-mêmes soit travailler pour l'Educational Testing Service qui possède un logiciel lié au modèle de Fusion. Ce panorama a changé avec la création d'une nouvelle macro pour estimer ces modèles, qui peut être utilisée avec le logiciel Mplus (Muthén & Muthén, 2006). Cette macro met enfin à la portée de nombreux chercheurs la possibilité d'appliquer ces modèles à des données (Douglas, de la Torre, Chang, Henson, & Templin, 2006).

Dans cette présentation, des considérations d'ordre aussi bien théoriques que pratiques seront abordées. Tout d'abord, les fondements conceptuels de ces modèles et les postulats de base liés à leur application seront passés en revue puis synthétisés. En particulier, l'importance des fondements théoriques liés à la cognition dans le cadre des processus de réponse et les limites pratiques imposées par la complexité seront contrebalancées. Dans un deuxième temps, les modèles DINA, NIDA et de Fusion seront présentés et appliqués à différentes données pour lesquelles une analyse dans une perspective diagnostique présente un intérêt. Une attention particulière sera portée à l'interprétation des paramètres, aux choix d'un modèle plutôt qu'un autre dans diverses situations et au format des portraits à fournir aux apprenants. Enfin, les avantages de ces modèles sur ceux qui sont basés sur le score total et les calibrations unidimensionnelles issus de la théorie de réponse à l'item (e.g., van der Linden & Hambleton, 1997) seront discutés, plus particulièrement dans le cadre des évaluations à grande échelle dans un contexte de réforme. Des périodes de discussions concernant de futures directions de recherche permettront aux participants intéressés de s'exprimer.

LE TESTING ADAPTATIF PAR ORDINATEUR: OUTIL D'EVALUATION FORMATIVE

Patrick PLICHART, Centre de Recherche Public Henri Tudor

Gilbert BUSANA, Université du Luxembourg

Les ateliers pratiques 2 et 3 proposent aux participants du colloque de s'initier au testing assisté par ordinateur au travers d'une plateforme informatique élaborée conjointement par l'unité de recherche EMACS de l'Université du Luxembourg et le Centre de Recherche Public Henri Tudor. Le dispositif permet des évaluations davantage centrées sur les processus cognitifs (et autres processus psychologiques) vus sous leur aspect dynamique et émergent.

La plateforme qui a été mise au point est entièrement basée sur Internet, ce qui permet une utilisation facile et sans connaissances informatiques particulières à partir d'un ordinateur disposant d'un accès Internet. L'originalité de la plateforme consiste dans le fait que les différents « métiers » du processus d'évaluation ont été séparés et implémentés dans des modules spécialisés: (1) la gestion et la création des items et de leurs contenus (2) la gestion et la création des tests (3) la gestion et la description des sujets (4) la gestion et la description des groupes de sujets (5) la mise à disposition et la passation des tests et (6) la gestion et la mise en forme des résultats. Les différents modules peuvent être installés de façon multiple et d'une manière délocalisée (sur des serveurs différents), ce qui permet une structure de communication et de collaboration ouverte qui est administrée par des abonnements mutuels et par l'attribution de droits d'accès. Cette architecture assure ainsi un maximum de flexibilité qui permet d'un côté la collaboration entre un grand nombre d'individus situés à des niveaux institutionnels variés, et de l'autre côté la possibilité d'adapter la plateforme à l'ensemble des besoins en évaluation.

D'un point de vue des fonctionnalités, la plateforme va donner accès à l'ensemble des fonctionnalités informatiques qui sont exploitables dans une perspective d'évaluation (création de nouveaux formats d'items utilisant les fonctionnalités multimédia de l'ordinateur, enregistrement des temps de réaction et de données comportementales, testing adaptatif, scoring automatique, etc.). L'adoption du paradigme « open-source » pour la programmation de la plateforme assure en plus la possibilité de répondre à des besoins très spécifiques par le développement collaboratif et délocalisé de nouvelles fonctionnalités sur le plan informatique.

http://www.jemacs.uni.lu/ 2007 volume 1, No 1

APPLICATION WEB POUR LA CREATION ET L'ADMINISTRATION EN LIGNE DE TESTS ADAPTATIFS

Nancy BOITEAU, Université du Québec à Rimouski

Dans le cadre d'une étude menée en 2004 et visant à comparer l'efficacité de plusieurs modalités de testing adaptatif, nous avons procédé, en collaboration avec une équipe d'experts en informatique, à l'élaboration d'une application Web permettant la création et l'administration en ligne de tests adaptatifs. Cette étude s'intéressait particulièrement à une modalité de testing adaptatif connue sous le nom de modalité à deux niveaux ; celle-ci est utilisée dans les enquêtes du Programme d'indicateurs de rendement scolaire (PIRS) au Canada. En combinant les modèles de la TRI à l'usage des TIC, cette étude visait à créer une modalité « optimale » à deux niveaux qui soit d'une efficacité comparable à celle offerte par la modalité de testing adaptatif par ordinateur (TAO), mieux connue sous le nom de Computerized Adaptive Testing (CAT). C'est pour répondre à l'objectif de cette étude qu'une plate-forme de création et d'administration de tests adaptatifs supportée par l'infrastructure réseau d'Internet a été développée en collaboration avec des spécialistes dans le développement d'applications Web.

L'atelier pratique propose d'initier les participants aux trois volets fonctionnels offerts par cette application Web : le volet Administration, le volet Gestion et le volet Administration de tests. Il vise ainsi à initier les participants aux fonctionnalités de l'application Web sous trois profils d'utilisateurs: administrateur du système, intervenant et répondant.