

24^e Colloque International de l'ADMEE-Europe

L'évaluation des compétences en milieu
scolaire et en milieu professionnel

11-13 janvier 2012, Luxembourg

Résumés des ateliers (Ateliers 1-14)

	Titre de la communication	Auteur(s)
Atelier 1 : Autour de la notion de compétence	COMPETENCES ET / OU REPRESENTATIONS SOCIALES	Christian Béliston
	AU CARREFOUR DU CURRICULUM PRESCRIT ET DU CURRICULUM IMPLANTE : POLEMIQUE ET POLYSEMIE AUTOUR DU TERME DE COMPETENCE EN FEDERATION WALLONIE-BRUXELLES	Natacha Duroisin, Sabine Soetewey & Marc Demeuse
Atelier 2 : Outils d'évaluation	COMMENT EVALUER LE DEVELOPPEMENT DE COMPETENCES ETHIQUES EN FORMATION INITIALE D'ENSEIGNANTS ?	Antoine Sautelet
	EVALUATION DE LETTRES COMMERCIALES A VISEE PERSUASIVE (LE CAS DES ETUDIANTS DE L'ECOLE NATIONALE DE COMMERCE ET DE GESTION D'AGADIR / MAROC).	Amina Saoussany
Atelier 3 : Compétences conatives et sociales	L'EVALUATION DES COMPETENCES EN CONTEXTE COLLECTIF : POIDS DES RELATIONS SOCIO-AFFECTIVES SUR LES CONDUITES.	Lucie Mougenot
	DETERMINANTS PERSONNELS DES COMPETENCES DE L'ECOLIER : LE CAS DE SA PERFORMANCE SCOLAIRE ET DE SES REACTIONS ENVERS LA CLASSE	Halim Bennacer
Atelier 4 : Evaluation diagnostique	EVALUER DES COMPETENCES LATENTES PAR L'ANALYSE DES ERREURS : L'EXEMPLE D'EPREUVES DE RAISONNEMENT	Philippe Chartier
	LA FACE CACHÉE DES COMPÉTENCES EN LECTURE	Patricia Schillings, France Neuberg & Ariane Baye
	QUELLE ARTICULATION ENTRE LES COMPETENCES PAR DOMAINE ET PROFILS DE COMPETENCES : L'EXEMPLE DE PISA EN SUISSE	Jean Moreau & Christian Nidegger
Atelier 5 : Dispositifs de formation des adultes	DES OUTILS D'EVALUATION DES COMPETENCES EN LANGUE FRANCAISE DES ENSEIGNANTS DU PRIMAIRE	Wajiha Smaili
	L'ÉVALUATION D'UN DISPOSITIF D'INSERTION ; ÉTUDE DU DÉVELOPPEMENT PROFESSIONNEL PERÇU DES ANIMATEURS PAR L'ANALYSE DES COMPÉTENCES	Lucie Aussel
Atelier 6 : Référentiels et outils d'évaluation	LES EVALUATIONS A FAIBLES ENJEUX : QUEL ROLE JOUE LA MOTIVATION ? UNE EXPERIENCE A PARTIR DE PISA.	Saskia Keskpaik & Thierry Rocher
	UN DISPOSITIF INTÉGRÉ POUR L'ÉVALUATION DES ACQUIS DES ÉLÈVES DE 7E ANNÉE À MADAGASCAR	François-Marie Gerard
	DETERMINATION DU SEUIL DE MAITRISE POUR L'ATTRIBUTION DES COMPETENCES DU SOCLE COMMUN COMPARAISON DE DIFFERENTES METHODES	Pascal Bessonneau & Thierry Rocher

Atelier 7 : Outils et dispositifs d'évaluation	EVALUER DES COMPETENCES DANS UNE OPTIQUE SITUATIONNELLE : UN GAGE DE VALIDITE ? L'EXPERIENCE DE DEVELOPPEMENT D'OUTILS D'EVALUATION DES COMPETENCES EN HISTOIRE AU SECONDAIRE EN BELGIQUE FRANCOPHONE	Jean-Louis Jadoulle
	EVALUER L'HISTOIRE DES ARTS A L'ECOLE PRIMAIRE, DISCOURS ET DISPOSITIFS	Gaëlle Boubennec, Laurent Séjourné & Cécile Vendramini
	LES COMPETENCES PERSONNELLES ET SOCIALES EN QUESTION. ILLUSTRATION A TRAVERS DEUX OUTILS D'EVALUATION EN EDUCATION PHYSIQUE ET SPORTIVE	Lenzen Benoît, Dénervaud Hervé & Poussin Bernard
Atelier 8 : Compétences des enseignants	EVALUATION DE LA COMPETENCE A COMMUNIQUER A L'ECRIT ET A L'ORAL EN FORMATION INITIALE DES MAITRES : ETAT DE LA SITUATION AU QUEBEC	Christian Dumais & Nancy Goyette
Atelier 9 : Outils informatiques	L'ÉVALUATION DES COMPÉTENCES EN MILIEU SCOLAIRE ET EN MILIEU PROFESSIONNEL	Mohammed Talbi & Faouzia Messaoudi
	LA DEMARCHE EPORTFOLIO DANS L'ENSEIGNEMENT SUPERIEUR FRANÇAIS : PREMIERS RETOURS D'EXPERIENCES ET PROPOSITION DE CLARIFICATION	Jean Heutte & Annie Jézégou
Atelier 10 : Portfolio	ÉVALUER LE PORTFOLIO POUR ÉVALUER DES COMPÉTENCES. UNE GRILLE D'ÉVALUATION DOUBLEMENT CRITÉRIÉE	Evelyne Vanden Aynde
	LE PORTFOLIO DE DEVELOPPEMENT PROFESSIONNEL DU BACHELOR EN SCIENCES DE L'EDUCATION DU L'UNIVERSITE DU Luxembourg : UN MEDIA ET UN PROCESSUS EN CONSTANTE EVOLUTION	Débora Poncelet, Michelle Brendel, Viviane Bourg & Sylvie Kerger
	EVALUATION DU PROCESSUS DE MISE EN PLACE DE DEUX OUTILS PEDAGOGIQUES (PORTFOLIO ET GRILLE D'EVALUATION DES STAGES) EN LIEN AVEC LE REFERENTIEL DE COMPETENCES INTEGRE EN KINESITHERAPIE.	Helyett Wardavoir, Karin Van Loon, Catherine Romanus, Dominique Peeters, Patrick Parmentier & Florence Parent
Atelier 11 : Testing assisté par ordinateur	EVALUATION DES ENSEIGNEMENTS AU FORMAT PAPIER ET EN LIGNE : IMPACTS SUR LES PROCEDURES ET RESULTATS	Mallory Schaub & Marine Chittaro
	ANALYSE DE VALIDITE DES TESTS DE CLASSEMENT EN LIGNE POUR LES LANGUES A L'AIDE DU MODELE DE RASCH	Cornelia Gick, Hervé Platteaux, Sergio Hoein, Catherine Blons-Pierre & Patricia Kohler
	ANALYSE QUALITATIVE DE TESTS DE CLASSEMENT EN LIGNE POUR LES LANGUES : ACCEPTANCE ET ROLE DE L'AUTOEVALUATION	Catherine Blons- Pierre, Patricia Kohler, Cornelia Gick, Sergio

		Hoein & Hervé Platteaux
Atelier 12 : Compétences, hétérogénéité et équité	LES COMPÉTENCES FONDAMENTALES ET LE DÉVELOPPEMENT : PEUT-ON ÉVALUER LES SYSTÈMES ÉDUCATIFS PAR LE NIVEAU D'ACQUISITION HOMOGENÈ D'UN BLOC DE COMPÉTENCES DE BASE ?	Nadir Altinok & Jean Bourdin
	LA CLASSE COMME « ZONE DE RESPONSABILITÉ » : VERS LA FLEXIBILITÉ SOCIOCOGNITIVE EN TANT QUE COMPÉTENCE TRANSVERSALE DE L'ENSEIGNANT(E)	Aneta Mechi
	LES COMPETENCES DES ETUDIANTS A L'ENTREE A L'UNIVERSITE : QUELLE INCIDENCE SUR LA REUSSITE?	Sophie Morlaix & Bruno Suchaut
Atelier 14 : Dispositifs d'évaluation	L'APPRENTISSAGE PAR LES COMPETENCES LE RÔLE DE L'ÉVALUATION AU COURS DE L'APPRENTISSAGE PAR COMPÉTENCE	Zoualfakar Jammoul
	L'APPROCHE PAR COMPETENCES ET L'ÉVALUATION CLINIQUE OBJECTIVE STRUCTUREE EN SOINS INFIRMIERS.	Corinne Bay
	UNE EVALUATION « PAR LA BANDE » DES COMPETENCES EN MILIEU PROFESSIONNEL. FORMATEURS D'ADULTES ET CHARGES D'INSERTION SOUS EVALUATION	Cédric Frétygné
Atelier 15 : Modèles de mesure et analyse statistique	HIERARCHISATION DE COMPETENCES PAR L'INTERMEDIAIRE DU LOGICIEL C.H.I.C. : EXEMPLES D'ANALYSES DE DONNEES DANS DEUX DOMAINES DIFFERENTS	Stéphanie Malaise & Dominique Casanova
	COMMENT ASSURER LA COMPARABILITE DES SCORES ISSUS D'ÉVALUATIONS NATIONALES ANNUELLES ET EXHAUSTIVES ? COMPARAISON DE DIFFERENTES METHODES D'AJUSTEMENT DES METRIQUES (« EQUATING »).	Thierry Rocher
	LE SYSTÈME D'ÉVALUATION NATIONALE ITALIEN (INVALSI) COMME OUTIL POUR METTRE EN ÉVIDENCE UN DÉFICIT DE COMPÉTENCES	Margherita Vitale
Atelier 16 : Outils d'évaluation et de formation	L'IMPACT DE L'ÉVALUATION QUALITATIVE SUR LE DEVELOPPEMENT ET L'IDENTIFICATION DES COMPETENCES PROFESSIONNELLES DES FUTURS ENSEIGNANTS DE LANGUES	Frédérique Longuet & Claude Springer
Atelier 17 : Référentiels et pratiques d'évaluation	EVALUATION PRAGMATIQUE DES COMPETENCES	Christèle Billing
	ARTICULATION ENTRE REFERENTIELS, PRATIQUES D'ENSEIGNEMENT, DISPOSITIFS DE FORMATION ET PRATIQUES D'ÉVALUATION	Evelyne d'Hoop, Dominique Lemenu, Christine Malhomme & Michel Coupremagne

Atelier 18 : Reconnaissance et validation des acquis de l'expérience	INTRODUCTION D'UNE PROCEDURE DE VAE POUR LA FORMATION A L'ENSEIGNEMENT DANS LE SECONDAIRE II	Alessandra Di Maria & Nicole Rege Colet
	DES EFFETS DE LA VALIDATION DES ACQUIS DE L'EXPERIENCE EN ETABLISSEMENTS DE SERVICES ET D'AIDE AU TRAVAIL	Olivier Grammelspacher
Atelier 19 : Portfolio	LE PORTFOLIO COMME OUTIL DE REFLEXION DANS UNE COMMUNAUTE D'APPRENT-ES	Fabienne Breedijk-Goedert
	COMPETENCES SITUEES ET COMMUNAUTE D'APPRENANT-E-S INCLUSIVE	Michelle Brendel
	LE RECOURS AUX MODES D'EXPRESSION NON ECRITE DANS L'APPROCHE PORTFOLIO DANS LE CADRE D'UN APPRENTISSAGE INCLUSIF DES LANGUES	Melanie Noesen
Atelier 20 : Dispositifs de formation	PRATIQUES INNOVANTES DANS L'ACCOMPAGNEMENT DES ÉTUDIANTS DE MASTER EN ENSEIGNEMENT SPÉCIALISÉ	Pierre Petignat & Marie-Paule Matthey
	L'ANALYSE DES SITUATIONS D'ENSEIGNEMENT/APPRENTISSAGE EN EDUCATION VISUELLE AU 1ER CYCLE DANS UNE APPROCHE PAR COMPETENCES	Ana-Vanêssa Lucena
Atelier 21 : Evaluations externes nationales	UNE EVALUATION CERTIFICATIVE EXTERNE DES COMPETENCES EN SCIENCES : MISSION (IM)POSSIBLE EN BELGIQUE FRANCOPHONE ?	Sabine Soetewey & Natacha Duroisin
	QUELS OBJECTIFS POUR LES EVALUATIONS NATIONALES EN FRANCE ?	Jean-Pierre Jeantheau & Sandra Johnson
Atelier 22 : Référentiels et dispositifs d'évaluation	EVALUATION DE COMPETENCES DANS LA REALISATION/EVALUATION D'UNE RECHRECHE EN SOINS ET SANTE (TFE DE BACHELIER EN SOINS INFIRMIERS) : OUTILS ET DISPOSITIF EVALUATION.	Dumont Valérie, Corman Cécile, Fastré Anne-Cécile & Faufra Laurence
	LE TRAVAIL DE FIN D'ETUDES COMME EPREUVE D'INTEGRATION PROFESSIONNALISANTE. POUR QUELLES COMPETENCES DU METIER D'ENSEIGNANT ? AVEC QUELS OUTILS D'EVALUATION ?	Isabelle Bastin, Colette Leunus & Marie-Hélène Straeten
	LES ATTITUDES COMME INDICATEURS DES COMPETENCES POUR L'EDUCATION MULTILINGUE	Anna Volodina & Pierre Valois
Atelier 23 : Evaluation en situation authentique	EDUCATION EN SANTE : MODELISATION DES PRATIQUES DES SAGES-FEMMES FRANCAISES	Marie-Reine Bernard & Chantal Eymard
Atelier 24 : Dispositifs d'évaluation	EN QUOI LES DONNÉES LONGITUDINALES PERMETTENT-ELLES DE MIEUX ÉVALUER LES RETOURS SUR LES INVESTISSEMENTS EN FORMATION AU SEIN DES ORGANISATIONS ?	Amélie Bernier

Atelier 25 : Portfolio	SAVOIR DECRIRE UNE SITUATION PROFESSIONNELLE ; UNE CONDITION CLE DANS L'EVALUATION DES COMPETENCES PROFESSIONNELLES A L'AIDE D'UN PORTFOLIO	Chantal Eymard, Caroline Ladage & Michel Vial
	ENJEUX AUTOUR DE LA DECLARATION DES COMPETENCES VIA L'E-PORTFOLIO (NOUVEL OUTIL NUMERIQUE)	Laurence Puissant & Grosjean
Atelier 26 : Plurilinguisme	TRANSITIONS ET PLURILINGUISME : EVALUATION DES JEUNES ENFANTS ALLOGLOTTES ISSUS DES MINORITES LINGUISTIQUES AUX PAYS-BAS ET EN SUEDE A L'ECOLE PUBLIQUE	Emmanuelle Le Pichon-Vorstman
	EVALUER LES PERFORMANCES SCOLAIRES D'ELEVES PRESENTANT UNE DYSLEXIE SCOLARISES EN CLASSE ORDINAIRE D'ECOLE PRIMAIRE	Catherine Martinet & Valérie Hauser
Atelier 27 : Référentiels et dispositifs de formation	CONCEPTION D'UN REFERENTIEL COMME COLONNE VERTEBRALE D'UN MASTER EN ALTERNANCE EN GENIE ANALYTIQUE	Chantal Dupont, Dominique François, Thérèse Walravens, Rose-May Delrue, Daniel Schiavon, Christelle Maes, Béatrice Pirson, François Georges & Marianne Poumay
	LA CONCEPTION D'UN DISPOSITIF D'EVALUATION DES COMPETENCES AU SERVICE DE L'OBJECTIVATION ET D'UNE APPROCHE QUALITE	Véronique Gérard
Atelier 28 : Analyse de pratique d'évaluation	LA COMPETENCE EVALUATIVE DES ENSEIGNANTS : DE SON EVALUATION A SON DEVELOPPEMENT	Yann Mercier-Brunel & Anne Jorro
	L'EVALUATION EXTERNE STANDARDISEE : UN LEVIER DIDACTIQUE VERS L'APPROCHE PAR COMPETENCES ?	Gonzague Yerly
	PRATIQUES EVALUATIVES ENSEIGNANTES ET SOCLE COMMUN DES CONNAISSANCES ET DES COMPETENCES : PRATIQUES EFFICACES ? PRATIQUES INNOVANTES ?	Sandra Safourcade
Atelier 29 : Référentiels et dispositifs de formation	ÉVALUATION DES COMPÉTENCES DANS UN PROGRAMME DE FORMATION À L'ENSEIGNEMENT : UNE LECTURE PSYCHOMÉTRIQUE	Léon Harvey & Sylvie Bonin
	UNE CO-CONSTRUCTION DE FICHES METHODOLOGIQUES, UNE FAÇON DE S'APPROPRIER ET DIFFUSER L'APPROCHE PAR COMPETENCES	Dominique Lemenu, Elfried Heinen & Vinciane De Keyser
Atelier 30 : Dispositifs d'évaluation	EVALUER DES COMPETENCES PROFESSIONNELLES DANS DES DISPOSITIFS DE FORMATION EN ALTERNANCE : LE CAS	Lucie Petit & Anne- Catherine Oudart

	D'UN MASTER ET D'UNE LICENCE PROFESSIONNELS	
Atelier 31 : Développement professionnel	LES COMPETENCES PROFESSIONNELLES DE STAGIAIRES DE MASTER 2 PROFESSIONNEL : REGARDS EVALUATIFS	Ranim Ayoubé
	L'EVALUATION LONGITUDINALE PAR ETAPES DE LA COMPETENCE (ELEC) A MENER UNE RECHERCHE ET A EN PRODUIRE LE RAPPORT	Séverine Delcomminette, Bertrand Chapellier & Dieudonné Leclercq
	LES COMPETENCES PROFESSIONNELLES REQUISES POUR LA GESTION DES CONFLITS DANS LES EQUIPES DE TRAVAIL AU COLLEGIAT	Marie-Julie Babin & Lucie Héon
Atelier 32 : Attitude et compétences des enseignants	LES ATTITUDES COMME INDICATEURS COGNITIFS ET AFFECTIFS DES COMPETENCES RELATIVES AU DEVELOPPEMENT DURABLE	Astrid de Leeuw & Pierre Valois
	ÉVALUER LA FORMATION PRATIQUE DES ENSEIGNANTS : ENJEUX, DÉFIS ET PROPOSITIONS	Bernard André
Atelier 33 : Dispositifs d'évaluation des enseignants	QUEL SENS DE FUTURS ENSEIGNANTS ATTRIBUENT-ILS A L'EVALUATION PAR TACHE COMPLEXE ?	Marc Luisoni & Isabelle Monnard
	EVALUATION COLLEGIALE DE PRODUCTIONS COMPLEXES - A LA RECHERCHE D'INDICATEURS VALIDES DE NIVEAUX DE DEVELOPPEMENT DES COMPETENCES	François Georges & Marianne Poumay
	QUELLES CONNAISSANCES LES ENSEIGNANTS ONT-ILS DE LEURS ELEVES ET COMMENT S'EN SERVENT-ILS?	Elise Farinelle & Philippe Wanlin
Atelier 34 : Référentiels et pratique d'enseignement	POUR CONSTRUIRE ET UTILISER UN NOUVEAU REFERENTIEL METIER	Catherine Van Nieuwenhoven, Olivier Maes, Bertrand Monville & Jean-Marc Vifquin
	FORMER A LA PRATIQUE A L'AIDE D'UN REFERENTIEL DE COMPETENCES PROFESSIONNELLE ET D'UN OUTIL D'AUTOEVALUATION	Lucienne Kuffer & Chantal Tièche Christinat
	L'ÉVALUATION DES COMPÉTENCES PROFESSIONNELLES : DES APPROCHES THÉORIQUES À LEUR MISE EN OEUVRE DANS L'ÉVALUATION DES COMPÉTENCES À ENSEIGNER L'ÉCONOMIE GESTION EN LYCÉE	Solange Ramond
Atelier 35 : Outils d'évaluation des enseignants	ECHELLE D'AUTOEVALUATION DE COMPETENCES ET CARTOGRAPHIE DE TEXTES REFLEXIFS : PRESENTATION ET PARTAGE D'EXPERIENCE	Christophe Gremion

	EVALUER LA GESTION D'UNE SITUATION D'APPRENTISSAGE : D'UN MODELE THEORIQUE A SON OPERATIONNALISATION	Arnaud Dehon & Antoine Derobertmeasure
Atelier 36 : Dispositifs de formation et d'évaluation	EVALUATIONS DES COMPETENCES PROFESSIONNELLES DANS LE CADRE D UNE FORMATION A L INGENIERIE DE L INTERVENTION. DES PISTES POUR UNE RECHERCHE SUR L ETHIQUE DE L'INTERVENTION ?	Michel Crunenbergh
	L'ENTRETIEN D'EVALUATION DE FIN DE STAGE : UNE ACTIVITE DIALOGIQUE POUR INFERER LES COMPETENCES	Pasquale Chilotti
Atelier 37 : Dispositifs d'évaluation à l'université	UNE GRILLE CRITERIEE GENERIQUE POUR L'EVALUATION DES COMPETENCES EXPERIMENTALES DES ETUDIANTS A L'UNIVERSITE	Julien Douady, Christian Hoffmann, Stéphane Baup, Sébastien Marc & Marie-Françoise Soulage
	EXPOSE ORAL : EVALUER LA MAITRISE DE LA COMPETENCE COMMUNICATIONNELLE	Christian Michaud
	EVALUER LES COMPÉTENCES PROFESSIONNELLES DES CONSEILLERS PÉDAGOGIQUES À L'UNIVERSITÉ À L'AIDE D'UN RÉFÉRENTIEL	Sophie Lecloux & Dorothee Kozlowski
Atelier 38 : Référentiels et pratique d'enseignement/évaluation	DE LA MISE EN PLACE DE SITUATIONS DIDACTISEES EN FORMATION PROFESSIONNELLE AU DEVELOPPEMENT DE COMPETENCES EXTRASCOLAIRES EN MILIEU SCOLAIRE	Stéphane Vince
	LES OUTILS D'EVALUATION DES COMPETENCES SCRIPTURALES DANS LES MANUELS SCOLAIRES DE FIN D'ECOLE ELEMENTAIRE FRANÇAISE : QUELS FONDEMENTS THEORIQUES ET INSTITUTIONNELS ?	Anne Leclaire-Halté & Fabienne Rondelli
Atelier 39 : Référentiels et pratiques d'évaluation	EVALUATION DES COMPETENCES EN LANGUE ETRANGERE : UN REGARD SUR LES PRATIQUES DES ENSEIGNANTS DANS UNE ECOLE PORTUGAISE	Graciète Gaspar & Rosa Bizarro
	COMPETENCES : PLAN D'ETUDES, EVALUATIONS ET PRATIQUES DECLAREES A GENEVE	Anne Soussi & Christian Nidegger
	L'EVALUATION DE LA COMPREHENSION DE L'ECRIT AU PRIMAIRE : QUELLE (S) COMPETENCE(S) POUR QUEL(S) MODELE (S) DE COMPREHENSION ?	Zineb Haroun
Atelier 40 : Modèles de mesure	MESURE DE L'EVOLUTION DES COMPETENCES DES ELEVES DANS LE CADRE D'UNE EXPERIMENTATION PEDAGOGIQUE. UNE ILLUSTRATION PAR L'EXEMPLE AVEC	Marion Le Cam, Pascal Bessonneau & Thierry Rocher

	LE DISPOSITIF L'ENSEIGNEMENT INTEGRE DE SCIENCES ET TECHNOLOGIE (EIST).	
	MODÉLISATION PRÉDICTIVE DU RENDEMENT DES ÉLÈVES DE L'ONTARIO : APPROCHE PRATIQUE ET INDIVIDUELLE PERMETTANT DE CIBLER LES INTERVENTIONS PÉDAGOGIQUES	Serge Boulé & Mario Gagnon
Atelier 41 : Reconnaissance et validation des acquis de l'expérience	L'ÉVALUATION DES COMPÉTENCES DANS LA FORMATION PROFESSIONNELLE EN BELGIQUE FRANCOPHONE	Mathieu Gaillard, Bérenger Dufromont, Catherine Deruisseau, Isabelle Verbruggen & Christophe Mattart
	BILAN DE COMPETENCES DANS LE CADRE DE LA VALORISATION DES ACQUIS DE L'EXPERIENCE : AUTO-EVALUATION ET ROLE DE L'ACCOMPAGNANT	France Dantinne & Evelyne Charlier
Atelier 42 : Analyse de pratiques d'évaluation	L'ÉVALUATION DES COMPETENCES EN LANGUES ANCIENNES D'UNE NOTE « SUBJECTIVE » A L'APPLICATION DE GRILLES CRITERIEES	Thomas Debrux
	ETUDE DES ANNOTATIONS D'UN ENSEIGNANT A LA SUITE DE L'ENSEIGNEMENT EXPLICITE DES STRATEGIES DE RESOLUTION DE PROBLEMES MATHEMATIQUES	Joanne Lapointe
Atelier 43 : Testing assisté par ordinateur	L'ÉVALUATION DES COMPETENCES SCIENTIFIQUES AU 21IEME SIECLE : CADRE CONCEPTUEL ET PROTOTYPES POUR UNE EVALUATION TRANSFORMEE-PAR-ORDINATEUR	Katja Weinerth, Robert Reuter, Vincent Porro, Thibaud Latour & Romain Martin
	EVALUATION DE LA COMPREHENSION DE L'ECRIT SUR SUPPORT INFORMATIQUE ET COMPARAISON AVEC DES EPREUVES DE TYPE PAPIERCRAYON	Pascal Bessonneau
	LES JEUNES FACE A L'ECRIT – UN NOUVEAU DISPOSITIF D'EVALUATION MASSIVE DE LEURS PERFORMANCES EN LECTURE	Thierry Rocher
Atelier 44 : Outils d'évaluation	ADAPTABILITE INDIVIDUELLE : SA STRUCTURE ET SES RELATIONS AVEC D'AUTRES ECHELLES D'ADAPTABILITE	Armanda Hamtiaux & Claude Houssemand
Atelier45 : Autour de la notion de compétence	PROPOSITION D'ANALYSE DE LA COMPETENCE A PARTIR DE SA MOBILISATION PAR DES PROFESSIONNELS DE L'ENSEIGNEMENT	Starck Sylvain
	FORMER DES ENSEIGNANTS A DES PRATIQUES PERTINENTES A L'APPROCHE PAR COMPETENCES : RECIT D'UNE EXPERIENCE	Philippe Péaud
Atelier 46 : Dispositifs d'évaluation	L'ÉVALUATION DES COMPETENCES PROFESSIONNELLES DES ETUDIANTS INSPECTEURS DE L'ENSEIGNEMENT EN FIN	Mohamed Essaoudi, Raja Lotfi, Said Lotfi,

	DE FORMATION AU CENTRE NATIONAL DE FORMATION DES INSPECTEURS DE L'ENSEIGNEMENT (CNFIE) DE RABAT (MAROC).	Mohamed Radid & Mohammed Talbi
	LES COMPETENCES PROFESSIONNELLES INFIRMIERES : UNE INTELLIGENCE PRATIQUE DES SITUATIONS	Bahilila Daineche
	L'EVALUATION DES COMPETENCES EN MATIERE DE CONDUITE D'ENTRETIEN DE RELATION D'AIDE EN TRAVAIL SOCIAL	Evelyne Charrière & Martine Bürger-Heim
Atelier 47 : Analyses de pratiques	LA SUBJECTIVITE : UNE COMPETENCE A METTRE AU TRAVAIL	Claudine Manier

COMPETENCES ET / OU REPRESENTATIONS SOCIALES
Christian Béliston*

* Université de Rouen
Laboratoire CIVIIC –
christian.beliston@val-de-marne.gouv.fr

Mots-clés : compétence, représentations sociales, évaluation,

Résumé.

La compétence est un concept qui englobe de nombreux phénomènes. Est-ce un concept d'objet, dans ce cas observable, ou une catégorie, qui aurait alors plus une fonction générique et heuristique ? Dans ce dernier cas, est-il adapté aux situations d'évaluation ? Après une brève approche étymologique et une rapide revue de la littérature, nous aborderons les différentes problématiques que pose ce concept, à partir d'un rapport de l'Inspection Générale de l'Education Nationale. L'analyse des discours nous offrira les outils pour approfondir l'étude, soit de documents officiels (fiches du RNCP), soit d'entretiens réalisés auprès d'examineurs d'un diplôme professionnel d'animateur (le BEATEP).

1. Un petit historique sur la notion de compétence :

Le concept de compétence s'est affirmé, un premier temps, dans son acception juridique, comme reconnaissance à accomplir un acte. De cette première acception est dérivée, à la fin du 17^{ème} siècle la reconnaissance des connaissances qui confèrent le droit de juger ou de prendre des décisions. J.P. Bronckart et J. Dolz nous brossent un rapide historique sur la genèse de cette notion :

« Le terme de *compétence* est attesté dans la langue française depuis la fin du XV^{ème} siècle ; il désignait alors la légitimité et l'autorité conférée aux institutions pour traiter de problèmes déterminés (un tribunal est *compétent* en matière de ...) ; depuis la fin du XVIII^{ème} siècle, sa signification s'est étendue au niveau individuel et il désigne depuis lors toute capacité due au savoir et à l'expérience (...) L'expression de compétence linguistique a été introduite par Chomsky (1955) dans le cadre d'un article qui constitue l'un des textes fondateurs de la « révolution cognitive » en sciences humaines. L'objectif de l'auteur était alors de combattre le béhaviorisme linguistique et plus spécifiquement la thèse selon laquelle le langage s'apprend par essais/erreurs, conditionnements, renforcements » (dans J. Dolz & E. Ollagnier ; 2002 : 31, 32).

Le concept de compétence est ensuite repris par les ergonomes dans les années 1980 (M. de Montmollin, J. Leplat). Il s'agit alors d'interroger les processus cognitifs mis en œuvre par les acteurs, pour accomplir certaines tâches. Les ergonomes introduisent ainsi la différence entre le travail prescrit et le travail réel. Mais ce concept est alors « de l'ordre de la description et non de l'évaluation » (M. de Montmollin ; 2001 : 13). Ils observent les compétences à partir d'un ensemble d'activités des protagonistes : schèmes et routines intériorisées en fonction des situations ; ressources mobilisées, matérielles, organisationnelles, cognitives (connaissances déclaratives et procédurales) et humaines ; processus cognitifs pour analyser la situation, coordonner les schèmes et hiérarchiser les actions ; enfin, modes d'interactions et de communication organisés au sein d'une équipe. Le concept de compétence se généralise aussi au sein des milieux de la formation professionnelle, avec les travaux de G. Malglaive et de B. Rey. Il acquiert alors une autre signification, les performances apparaissant comme une expression des compétences. Ces auteurs s'émancipent ainsi du schéma réducteur de la pédagogie par objectifs pour une approche plus

globale de la formation et des comp  tences    acqu  rir. Mais c'est surtout    partir des assises du MEDEF en 1998    Deauville que le concept se popularise (A. Dupray ; C. Guitton ; S. Monchatre ; 2003). Au d  but du XXI  me si  cle, avec l'  laboration des fiches du RNCP (r  pertoire national de la certification professionnelle), le concept de comp  tence est entr   dans les m  urs. Mais d'un point de vue scientifique, l'  valuation des comp  tences n'est pas sans poser maints probl  mes. Un rapide parall  le   tymologique avec le concept de « fonction » a l'int  r  t de mettre en valeur comment les significations actuelles de la comp  tence se sont construites au fil des si  cles. Un concept n'est jamais isol  , il s'inscrit dans des ensembles conceptuels et ceux-ci refl  tent l'  volution des m  urs et de l'organisation sociale. Au XVI  me si  cle, la fonction est un « faire »    accomplir pour remplir un « r  le » (ensemble de missions copi  es dans un rotulus, parchemin). La fonction   volue    la fin de celui-ci, en particulier dans les essais de Montaigne (1580), elle devient une des propri  t  s des organes qui contribuent    un m  me but. C'est au XVII et XVIII  me si  cles, en particulier avec les travaux de Leibniz, qu'elle acquiert sa signification math  matique, celle de rapports fonctionnels entre des   l  ments. Au XIX  me si  cle, la fonction devient la profession contribuant    la vie sociale, de laquelle a   t   d  clin   le terme de « fonctionnaire ». Au XX  me si  cle, elle acquiert de nouvelles acceptions scientifiques, en particulier en linguistique o   elle signifie la propri  t   logico-grammaticale des unit  s langagi  res, ou en sciences de l'  ducation o   elle qualifie les principaux « faire » professionnels. Mais dans ce cas, une question fondamentale se pose : ces concepts, comp  tence et fonction, font-ils r  f  rence    des objets ou s'agit-il de cat  gories ? Cette question n'est pas neutre, car si, par essence, les objets poss  dent des attributs observables, ce n'est pas le cas d'une cat  gorie qui fait r  f  rence aux processus cognitifs mis en jeu pour conduire la m  thodologie. Ainsi, si la comp  tence est une cat  gorie, elle n'est pas observable et, de ce fait, elle n'est pas   valuable non plus. Si en revanche, il s'agit d'un concept d'objet, il est possible d'en pr  ciser les attributs qui lui donnent du sens dans nos r  alit  s quotidiennes. La conception de la comp  tence sera alors logiquement d  termin  e (au sens linguistique) par des concepts attributs, dont les r  f  rents sont identifiables    travers nos pratiques ou dans notre univers sensible.

2. Les limites de la comp  tence :

Lorsque N. Chomsky d  finit son concept de comp  tence, il r  interpr  te la dichotomie de F. de Saussure entre la parole et la langue : il oppose la performance, qui est de l'ordre de la production hic et nunc,    la comp  tence qui exprime l'  d  e d'une structure profonde. La performance langagi  re est l'actualisation de la comp  tence dans une situation de communication donn  e. N. Chomsky inf  re donc la comp  tence grammaticale    partir de la performance discursive. Mais ainsi con  u, le mod  le de la comp  tence a plus une fonction heuristique qu'une fonction objective et pr  dictive. Par ailleurs, D.H. Hymes (1973/1991), dont les travaux sont fond  s sur une approche ethnologique, a critiqu   l'insuffisance des conceptions de N. Chomsky, qui a r  duit le mod  le de la comp  tence    des acquisitions langagi  res coup  es de l'apprentissage de la vie en soci  t  . Or il existe aussi des comp  tences de communication sociale, non langagi  res, apprises par tous les membres d'une m  me communaut   culturelle. La comp  tence linguistique/ grammaticale, existerait-elle sans les comp  tences de communication, qui offrent du sens    tous les   v  nements qui surviennent dans notre environnement :

« Un enfant normal acquiert une connaissance des phrases, non seulement grammaticale, mais aussi comme   tant ou non appropri  es. Il acquiert une comp  tence qui lui indique quand parler, quand ne pas parler, et aussi de quoi parler, avec qui,    quel moment, o  , de quelle mani  re. Bref, un enfant devient    m  me de r  aliser un r  pertoire d'actes de parole, de prendre part    des   v  nements de parole et d'  valuer la fa  on dont d'autres accomplissent ces actions. Cette comp  tence, de plus, est indissociable de certaines attitudes, valeurs et motivations touchant    la langue,    ses traits et    ses usages et est tout aussi indissociable de la comp  tence et des attitudes relatives    l'interrelation entre la langue et les autres codes de conduite en communication (cf. Goffman) » (D.H. Hymes ; 1973/1991 : 74).

La comp  tence linguistique n'est donc qu'un aspect de la communication, autrement dit une des diverses comp  tences de communication qui s'expriment sous la forme de conduites normalis  es d'une communaut   culturelle. Ces modes de communication sont int  rioris  s tr  s t  t et ils acqui  rent du sens    travers les relations de r  f  rence avec des modes implicites : les syst  mes de r  les (routines), les fa  ons de faire, les attentes r  ciproques, etc. La comp  tence appara  t ainsi comme un concept g  n  rique pour signifier un ensemble de ph  nom  nes. Elle exprime l'organisation des divers modes de communication du milieu professionnel : proc  dures, processus cognitifs, syst  mes de r  f  rence, etc. Si on se r  f  re    la d  finition de G. Le Boterf (1994), la « comp  tence » est une « mobilisation de ressources » pour agir (p. 17), un « acte d'  nonciation » (p. 19), un « syst  me, une organisation structur  e qui associe de fa  on combinatoire divers   l  ments » (p. 22), un « savoir int  grer » (P. 23), « c'est *avoir autorit   pour...* et les moyens de l'exercer » (p. 28), c'est un « savoir agir » (p. 32) « reconnu » (p. 35), mais c'est aussi une « unit   bi-polaire » qui allie « situation professionnelle » et « sujet acteur » (p. 50), un ensemble construit de « repr  sentations op  ratoires » (p.53), une « intention » qui communique un sens, une « signification » aux actes (p. 66), des « savoirs mobilisables » (p.73) : « savoirs th  oriques » (p.73), « savoirs proc  duraux » (p.84), « savoirs faire proc  duraux » (p. 87), « savoirs exp  rientiels » (p. 89), « savoirs faire sociaux » (p. 102) ; des « savoirs faire cognitifs » (p. 108) ; mais aussi,    la marge du syst  me, des « pratiques professionnelles et performances » (p. 117) et des « boucles d'apprentissage » (p. 122). De telles d  finitions attirent notre attention sur la complexit   des ph  nom  nes et leur enchev  trement : elles sont certainement adapt  es    des interventions sur le terrain (formation, conseil, etc.), dans la mesure o   elles fournissent un mod  le syst  mique (sorte de check-list pour les praticiens) de tous les domaines    envisager, en fonction des probl  matiques pos  es par les partenaires. Mais une d  finition si large n'est gu  re satisfaisante pour la recherche, et elle est difficile    exploiter dans le domaine de l'  valuation (voire de la formation au niveau op  rationnel), dans la mesure o   il y a une trop grande diversit   de ph  nom  nes    observer. Quelles dimensions le concept de comp  tence apporte-t-il en sus des autres concepts qui d  finissaient les capacit  s    acqu  rir en cours de formation ? Cette probl  matique a   t   aussi formul  e, en d'autres termes, par J.P. Bronckart et J. Dolz :

« Sans verser dans le purisme conceptuel, il nous para  t   vident qu'on ne peut raisonnablement penser la probl  matique de la formation en usant d'un terme qui finit par d  signer tous les aspects de ce que l'on appelait autrefois les « fonctions psychologiques sup  rieures » » (*dans J. Dolz & E. Ollagnier ; 2002 : 35*).

La conception des ergonomes est, nous l'avons vu plus haut, d  j   beaucoup plus pr  cise. Mais l   encore, le concept appara  t beaucoup plus comme une entit   g  n  rique qui permet de regrouper certains ph  nom  nes pour r  pondre    certaines probl  matiques : adaptation des ressources mobilis  es aux routines, difficult  s de communication au sein des   quipes, d  calage entre sch  mas cognitifs et r  alit  s de la situation, etc. Il para  t bien difficile de qualifier tous les aspects d'une comp  tence,    plus forte raison de les   valuer, l'objectif des ergonomes est bien plus modeste. La comp  tence est certainement un concept g  n  rique tout    fait pertinent pour faire progresser la recherche sur les processus cognitifs mis en jeu au cours des situations professionnelles, mais tel qu'il est d  fini    l'heure actuelle, il est bien loin de signifier un objet identifiable... et   valuable.

3. L'  ducation nationale fran  aise et la notion de comp  tence :

Ce serait une erreur de penser qu'il ne s'agit l   que d'une critique des milieux scientifiques. Les milieux professionnels ont   t   aussi sensibles    ces contradictions. Ainsi un rapport de l'IGEN (Inspection G  n  rale de l'  ducation Nationale) ne manque pas de faire remarquer que la comp  tence est « une notion aux contours flous » (2007 : 10). Bien que les auteurs parviennent    cerner un noyau commun aux diverses d  finitions, qui s'articule autour de la mobilisation de diverses ressources et de l'adaptation de celles-ci    une situation donn  e (*idem* : 11), ils soulignent la complexit   de « d  finir une   chelle de niveau des comp  tences » :

« A partir de quelle quantit   de ressources mobilis  es pourra-t-on parler d'  laboration de comp  tences et non plus de restitution de processus automatis  s ? De quelle vari  t   de situation parle-t-on ? » (*Idem* : 12).

La comp  tence signifie un ensemble de processus complexes, qui varie en fonction des situations. En tant que concept g  n  rique, il a certainement un grand int  r  t pour r  fl  chir aux situations d'apprentissage, pour favoriser les acquisitions des   l  ves et leur adaptation aux diff  rentes situations, mais c'est un concept difficile    transformer en objet d'  valuation. D'ailleurs, les   volutions qui se sont op  r  es, en France, avec le socle commun de comp  tences (idem : 15) ou, en Angleterre, avec le National Curriculum (idem : 18), illustrent bien la tendance de ce concept    acqu  rir un caract  re g  n  rique :

« chaque grande comp  tence du socle commun est con  ue comme une combinaison de connaissances fondamentales pour notre temps, de capacit  s    les mettre en   uvre dans des situations vari  es, mais aussi d'attitudes indispensables tout au long de la vie comme l'ouverture aux autres, le go  t pour la recherche de la v  rit  , le respect de soi et d'autrui, la curiosit   et la cr  ativit   » (annexe du d  cret du 11 juillet 2006).

Le socle commun fran  ais est construit autour de sept comp  tences : « la ma  trise de la langue fran  aise, la pratique d'une langue vivante   trang  re, les comp  tences de base en math  matiques et la culture scientifique et technologique, la ma  trise des techniques usuelles de l'information et de la communication, la culture humaniste (...), les comp  tences sociales et civiques et, d'autre part, l'autonomie et l'initiative des   l  ves » (idem). De m  me le National Curriculum britannique s'articule autour des comp  tences cl  s suivantes : la communication, les applications num  riques, les technologies de l'information, le travail avec les autres, l'am  lioration de son propre apprentissage et de ses performances et la r  solution de probl  mes. On voit ainsi se dessiner, derri  re ce terme, un concept g  n  rique qui regroupe des connaissances, des capacit  s de r  alisation et des attitudes adapt  es    certaines situations. On retrouve ici, comme signification de ce concept, le triptyque savoir, savoir-faire et savoir   tre et les d  finitions propos  es d  s 1992 par les inspecteurs de l'  ducation nationale dans l'enseignement technique (B. Porcher, C. Letemplier, I. Rak et ass. ; 1992). Mais dans ce cas, ce sont des connaissances, des proc  dures et des attitudes (certainement serait-il plus exact de parler d'*habitus*) qui sont observ  es et   valu  es, le concept de comp  tence n'apportant que la dimension d'ajustement / d'adaptation de celles-ci au contexte propos   / impos      l'apprenant. G. Le Boterf a   mis des r  serves sur cette conception de la comp  tence, somme de savoirs, savoir-faire et savoir-  tre (2000/2009). P. Perrenoud, au contraire, l'a reprise    son compte (1999 : 17). Mais derri  re ces d  bats, ne faut-il pas identifier la complexe n  gociation, souvent implicite entre, d'une part, la dynamique de d  veloppement de cette notion au niveau des entreprises et, d'autre part, les modes d'  valuation traditionnels appliqu  s au sein du syst  me scolaire ?

Apr  s avoir pass   en revue la fa  on dont les autres pays francophones ont int  gr   ce concept dans leurs curricula scolaires (Qu  bec, Belgique, Suisse), puis avoir   tudi   son appropriation par les diverses disciplines du secondaire fran  ais, le rapport de l'IGEN souligne diff  rentes difficult  s. En analysant l'exemple Qu  b  cois, trois critiques sont relev  es : une incompr  hension de la part des familles, une difficult   technique pour les enseignants et un rejet g  n  ralis   de l'  valuation des attitudes (idem : 24). Des probl  mes similaires sont relev  s en Suisse et en Belgique, en particulier la r  sistance des familles. Par ailleurs, le rapport souligne « la d  stabilisation des enseignants, leur donnant le sentiment de ne plus comprendre les objectifs de l'  cole, de ne plus savoir ce que l'institution attendait d'eux » (idem : 45). « En France, la situation est encore plus complexe » puisque l'appropriation de ce concept est diff  rente, « non seulement entre le premier et le second degr   mais aussi d'une discipline    l'autre » (idem : 45). Les inspecteurs g  n  raux font ensuite diff  rents constats pour expliquer ces r  sistances : en particulier les « injonctions institutionnelles qui n'h  sitent pas    multiplier les formules et les prescriptions », mais aussi la confusion entre les concepts (comp  tence, savoir-faire, capacit  s, connaissances...), l'incoh  rence des programmes surtout en termes d'interdisciplinarit  , le d  faut de rep  res et d'outils de formation pour les enseignants, et le manque de communication avec les familles. Au del   des questions de confusion conceptuelle et du manque de rep  res, qui interpellent les sciences de l'  ducation, ces obstacles soul  vent un autre probl  me de fond : l'  valuation n'est-elle pas, avant tout, un syst  me de communication qui a une fonction sociale ? Cette dimension appara  t compl  tement occult  e de toutes les r  formes.    vouloir mesurer les performances des   l  ves, l'essentiel est ainsi oubli  . Ces

questions nous renvoient donc directement aux interrogations   pist  mologiques qui ont   t   soulev  es dans les deux premiers paragraphes : 1   La comp  tence n'est pas un objet, observable et identifiable, mais une cat  gorie : concept    la fois g  n  rique, qui regroupe de nombreux ph  nom  nes, et    la fois fonctionnel, qui interroge les pratiques professionnelles. 2   L'  valuation n'a pas pour seule fonction de mesurer des acquisitions, c'est avant tout un syst  me de communication qui a pour vocation de transmettre aux jeunes g  n  rations le patrimoine culturel de l'humanit  .

Ces deux questions nous conduisent    une certaine r  serve par rapport au mod  le de la comp  tence de la linguistique, c'est    dire au couple performances (observables) / comp  tences (structures profondes), au del   de sa valeur heuristique. Tout d'abord, il g  n  re l'illusion d'une possible objectivit  , c'est-  -dire de la possibilit   d'inf  rer des comp  tences en structure profonde    partir d'observations cibl  es et de proc  dures cod  es dans des situations d  termin  es. Ensuite et surtout, il occulte la dimension communicationnelle intrins  que aux processus et syst  mes d'  valuation. Cela ne veut pas dire que ce concept soit sans int  r  t : dans sa fonction de cat  gorie, la comp  tence permet de penser globalement le processus d'  valuation (telle pertinence de la connaissance / proc  dure / attitude pour telle situation sociale ou professionnelle), et par la m  me, elle donne du sens    l'acte d'  valuation, pour les   l  ves mais aussi pour les parents. Le rapport lui m  me fait ressortir certains aspects positifs de cette conceptualisation : « clarifier l'attente institutionnelle, d  velopper et coordonner les diff  rentes fonctions d'  valuation pour favoriser la r  ussite des   l  ves, hi  rarchiser et articuler entre elles les comp  tences      valuer, articuler les diverses composantes d'une comp  tence, faire   merger des convergences entre disciplines, penser les co  herences entre   valuation et situations d'apprentissage » (idem : 51    53) Mais, pour ce faire, le concept de comp  tence doit retrouver sa juste place dans les processus d'  valuation, c'est-  -dire la position qu'il est susceptible de remplir pour enrichir nos conceptions   ducatives.

4. La comp  tence existe-t-elle ind  pendamment de la repr  sentation qu'en ont les acteurs ?

Cette question a d  j     t     voqu  e, sous diff  rentes formes, par divers auteurs, en particulier par G. Malglaive quand il critique les taxonomies d'  valuation : les apparences observables des comportements humains n'ont pas de sens en soi ; « pour   tre compris, (ceux-ci) doivent   tre construits par la pens  e de l'observateur » (1990/1993 : 119    121). Par cons  quent, la comp  tence inf  r  e    partir des observations n'a de sens qu'en r  f  rence    la repr  sentation que les diff  rents acteurs ont de la situation. A. Sch  utz (1943/1998) a insist   sur cette fonction de nos repr  sentations pour interpr  ter les comportements d'Autrui : notre fa  on d'appr  hender les actions de l'Autre ne peut l'  tre que sur le mode de la « typification », dans la mesure o   nous ne pouvons conna  tre ni ses motivations, ni son interpr  tation de la situation d'apr  s ses propres sch  mas qui sont le produit de son histoire biographique. Toute action individuelle est motiv  e par un « projet » qui n'existe pleinement que dans la pens  e de son auteur. Pour appr  hender de l'ext  rieur l'action de cet individu, l'observateur l'analyse d'apr  s la fa  on dont il con  oit le projet de cet acteur, chose qu'il ne peut faire qu'   partir d'une analyse typique de la situation, construite    partir de comparaisons avec sa propre exp  rience, et/ou de connaissances qu'il a sur les   l  ments qui constituent cette situation. Lorsqu'il s'agit de routines de la vie quotidienne, nous interpr  tons ainsi les actions de nos pairs en raison de notre connaissance de leur finalit  . Les choses deviennent plus complexes lorsque les actions sortent de nos sch  mas familiaux. L'  valuation des actions des   l  ves (ou des candidats lors des examens) implique n  cessairement une interpr  tation. Celle-ci est fonction des repr  sentations que le professeur (l'examinateur) a de la situation d'  valuation, de la comp  tence mise en jeu lors de celle-ci, mais aussi de la pertinence      valuer cette comp  tence. Son appr  ciation est orient  e par son interpr  tation des buts de la situation d'  valuation, qui sont aussi cens  s avoir   t   compris par l'  l  ve (ou le candidat) : sch  mas-types normalis  s qui font r  f  rence    cette situation et    la fa  on dont il faut s'y adapter. En d'autres termes, pour reprendre les concepts de la sociologie interactionniste (E. Goffman), les   valuateurs font r  f  rence aux « repr  sentations sociales » induites par la situation et le contexte. L'  valuation, au niveau de l'  ducation, appara  t donc bien moins une mesure d'une quelconque comp  tence (cf. la difficult   d'  laborer des   chelles de comp  tence   voqu  e par le rapport de l'IGEN) que

l'appr  ciation d'une prestation en fonction des repr  sentations que le professeur (l'examineur) a de la situation. L'  l  ve ou le candidat ont-ils la capacit   de percevoir les enjeux de la situation et de mettre en sc  ne les actions attendues, les modes de communication adapt  s au contexte. Lors d'un examen, un candidat    un dipl  me doit ainsi montrer qu'il a bien cern   les repr  sentations sociales du milieu professionnel. N'en est-il pas de m  me lors des situations d'  valuation scolaire ? P. Perrenoud (1984), en particulier, traduit ces enjeux    travers la formule « apprendre le m  tier d'  l  ve » ?

« Assimiler le curriculum, c'est devenir l'indig  ne de l'organisation scolaire, devenir capable d'y tenir son r  le d'  l  ve sans troubler l'ordre ni exiger une prise en charge particuli  re » (1984 : 251). « Ainsi r  ussir    l'  cole, c'est d'abord apprendre les r  gles du jeu » (Idem : 249).

Bien entendu, pour r  pondre    ces attentes, l'  l  ve doit d  velopper les capacit  s cognitives qui sont sollicit  es au cours des situations d'  valuation : connaissances, mais aussi adaptation aux situations, r  solution des probl  mes, positionnements pertinents, etc. Mais la question pos  e dans cet expos   n'est pas de savoir ce qu'il convient de travailler pour se pr  parer    ces   preuves, mais bien comment est-il possible, pour les enseignants et les examinateurs, d'  valuer les prestations mises en sc  ne au cours de celles-ci. Et quel est le sens de ce syst  me d'  valuation ? On peut donc faire l'hypoth  se que la comp  tence, en tant que concept g  n  rique, pr  sente une certaine pertinence pour organiser le curriculum de formation ; mais, en revanche, quand il s'agit d'  valuation, justement en raison de ce caract  re g  n  rique, elle n'est gu  re op  rationnelle. Cette hypoth  se, nous pouvons l'appr  hender par une approche macrosociologique, en comparant quelques curricula de formation, ou par une approche microsociologique, en analysant quelques processus cognitifs mis en   uvre par des examinateurs.

5. L'analyse de quelques repr  sentations sociales du m  tier :

Cette hypoth  se a   t   indirectement abord  e lors de l'  dition pr  c  dente de l'ADMEE (Paris ; 2011) au cours de laquelle on a compar   quelques repr  sentations que les professionnels avaient de leur m  tier,    partir d'une analyse discursive des fiches du RNCP (r  pertoire national de la certification professionnelle). Il a   t   ainsi possible de mettre en valeur,    partir de l'analyse de la position syntaxique des concepts les plus significatifs, que les conceptions des m  tiers divergent fortement entre l'industrie pharmaceutique, chimique ou agroalimentaire, o   s'est fortement d  velopp  e la gestion des comp  tences et les secteurs professionnels de la fonction publique am  nag  e (FPA), o   les dipl  mes sont requis pour exercer un m  tier (sport, assistant social,   ducation sp  cialis  e, psychomotricit  , ergoth  rapie). En particulier, les formes substantives des verbes (noms form  s    partir de verbes avec le suffixe « tion ») montrent que les premi  res conceptions font beaucoup appel    des proc  dures hi  rarchis  es, alors que les secondes privil  gient la « relation ». De m  me, les compl  ments circonstanciels brossent le contexte le plus significatif qui est pris en consid  ration : pour les premi  res, la « s  curit   » et le « travail d'  quipe » sont privil  gi  s, alors que pour les secondes, il s'agit du « cadre » de l'action professionnelle. Par cons  quent, la conception de la comp  tence, pour un m  tier donn  , n'est pas forc  ment g  n  ralisable    un autre. L   o   les uns observeront la construction des proc  dures en fonction de standards ou de sch  mas-type, les autres analyseront le positionnement relationnel et la fa  on de s'impliquer dans la communication au sein du groupe. L   o   les uns privil  gieront la coh  sion des   quipes et la vigilance interindividuelle, les autres seront plus sensibles aux analyses propos  es du cadre de l'action professionnelle. Les comp  tences attendues sont donc bien fonction des repr  sentations que les professionnels ont de leur m  tier. Chaque profession privil  gie certains sch  mas-type, certains modes de fonctionnement, qui ont de l'importance en raison du contexte de l'action. Nous avons vu ainsi que, pour la FPA, « conna  tre, comprendre et analyser le **cadre** des interventions professionnelles » apparaissait essentiel. Mais les cadres    ma  triser eux aussi sont diff  rents en fonction des milieux professionnels : l   o   les sportifs (DEJEPS sport) parlent de cadres r  glementaires, de la comp  tition et des objectifs de l'organisation, l'  ducateur sp  cialis   (DE) parle de politiques partenariales et d'  quipes pluridisciplinaires. Les   preuves mises en   uvre pour « v  rifier l'acquisition de ces comp  tences » sont donc bien diff  rentes, au m  me titre

que les positions, les processus cognitifs et les conceptions attendus. Certes, il est possible de regrouper toutes ces formes d'observation derri  re le vocable g  n  rique de « comp  tence », mais celui-ci traduit alors des ph  nom  nes forts divergents, et le concept n'apporte alors plus grand chose pour les   tudier. L'observation du type de comp  tences est donc bien induite par la repr  sentation qu'en ont les   valuateurs : elle est souvent d  termin  e (au sens linguistique) par leur conception du m  tier et de la fonction de l'  valuation pour le corps professionnel. Il n'existe donc pas de comp  tence ind  pendamment de la repr  sentation qu'en   laborent les acteurs qui sont d  sign  s pour l'  valuer.

6. L'analyse de quelques entretiens d'examineurs :

Pour illustrer ce questionnement, on analysera ensuite quelques   l  ments d'entretiens d'examineurs, auxquels il a   t   demand   d'expliquer la fa  on dont ils construisent leurs appr  ciations lors de certaines   preuves. Des entretiens non directifs ont   t   effectu  s aupr  s d'examineurs (nomm  s « experts ») d'un dipl  me professionnel de l'animation, le BEATEP aujourd'hui remplac   par le BPJEPS. Les   preuves (unit  s de formation technique et p  dagogique) se d  roulaient en trois temps : une mise en situation, pr  c  d  e et suivie d'un entretien avec le candidat. Un premier paradoxe avait attir   notre attention. On aurait pu s'attendre    ce que les sp  cialit  s les plus techniques (par exemple, avec une option « environnement ») se traduisent par des entretiens plus courts que les sp  cialit  s plus g  n  ralistes (par exemple, « activit   sociale et vie locale »), en raison d'une observation plus ais  e de tous les aspects techniques au cours de la s  ance. C'est l'inverse qui s'est produit. Les experts exigeaient beaucoup plus de justification des candidats qui avaient suivi les formations les plus techniques. 83¼ minutes en moyenne pour l'environnement contre 67½ pour ASVL. Il apparaissait ainsi que ce n'  tait pas tant l'observation des proc  dures et la v  rification des savoir-faire qui justifiaient la d  cision finale, car celles-ci auraient   t   plus ais  es et plus rapides avec les sp  cialit  s techniques. L'analyse des entretiens, en particulier de la construction des argumentaires, a permis de mettre en lumi  re quelques contraintes particuli  res aux processus cognitifs de la prise de d  cision. Un premier traitement par le logiciel TROPES, mis au point par R. Ghiglione et son   quipe (1998), a permis de d  gager des listes d'  nonc  s autour de certaines caract  ristiques syntaxiques, en particulier les propositions o   figuraient les verbes modaux dont on propose ici l'  tude. On qualifie de « modal » un verbe qui d  termine d'autres verbes (faire ou   tre) et qui signifie par la m  me une modalit   de ceux-ci (devoir, pouvoir, savoir, vouloir, etc.). Les verbes modaux pr  cisent ainsi un   tat subjectif du sujet de l'action (avec un faire) ou de l'  tat (avec   tre ou avoir). « Pouvoir faire » signifie que le sujet a le pouvoir n  cessaire pour effectuer l'action, « pouvoir   tre » qu'il a les capacit  s pour adopter l'  tat en question. Les verbes modaux sont donc particuli  rement pertinents pour analyser les qualit  s subjectives du candidat, attendues par l'interview  . Par ailleurs, l'enchaînement de la construction argumentative entre ces propositions modalis  es permet d'analyser la fa  on dont l'interview   structure le raisonnement qui le conduit    la d  cision. Bien entendu, d'autres formes syntaxiques induisent la m  me signification (par exemple, « il est capable de » ou « il connait »), mais les verbes modaux ont l'int  r  t de structurer la charpente de l'argumentaire. On aurait pu s'attendre, l   aussi,    avoir surtout des propositions avec les modalit  s du verbe « savoir », en raison de la question initiale : sur quoi vous appuyez vous pour justifier votre d  cision d'accorder ou non l'unit   de formation ? Or si l'on constate que les experts adoptent effectivement ce type de modalisation, parfois apr  s avoir   voqu   « les grilles de crit  res », c'est-  -dire le r  f  rentiel d'  valuation, les modalit  s du « savoir » (y compris avec le verbe « conna  tre ») restent bien modestes par rapport aux autres formes de modalisation. L'analyse pr  dicative offre un aper  u des argumentations mises en   uvre, et la faible proportion des savoirs appara  t alors intrins  que au processus d'  valuation.

Analysons les r  ponses d'un CEPJ (conseiller technique et p  dagogique du minist  re), exp  riment   dans l'  valuation de ce dipl  me, en centrant notre focale sur les verbes modaux. Apr  s avoir introduit le propos par des modalit  s g  n  rales (« d'apr  s la question, je dirai, il faut trouver la trame commune »), les premi  res modalisations sont des savoirs : « c'est savoir faire des diagnostics de situation, savoir poser des objectifs par rapport    ce diagnostic, c'est savoir partir de ces ... je dirai   tre capable de mettre en place des crit  res d'  valuation sur ces

objectifs ». D'autres modes de savoir apparaissent un peu plus loin : « *c'est quelqu'un qui a fait l'effort de conna  tre un contexte professionnel (...) qui a su reconstruire un projet d'animation (...) qui sait se situer (...) sait animer un groupe* ». Mais tr  s vite s'int  grent dans le propos deux autres modes importants, le « devoir » et le « pouvoir ». Le devoir concerne l'examineur (« *si je dois poser* »), le pouvoir concerne le candidat : « *c'est quelqu'un qui peut (...) pourra transf  rer son m  tier d'animateur* ». Cette derni  re modalisation coexiste avec une autre qui concerne l'examineur : « *ce que j'essaie de voir* ». Ces deux modes se d  veloppent ainsi de fa  on co-occurrence : il y a    la fois « ce qu'on voit » (l'examineur),    la fois « ce qu'il peut » (le candidat) : « *rappeler quels pouvaient   tre les crit  res d'  valuation (...) Est-ce qu'il pourrait en trouver d'autres* ». L'entretien   volue ensuite vers les attentes imp  ratives de l'interview  -examineur, exprim  es dans les modalisations du « devoir », essentiellement sous forme de « il faut ». La transition s'op  re    travers la proposition suivante : « *   la limite, il peut faire un choix, mais qu'il le justifie, qu'il l'argumente* », suivie un peu plus loin de : « *ce que je lui demande, c'est de me le justifier* », puis « *il faut qu'il me le justifie   galement par des crit  res* ». On retrouve cette structure plusieurs fois : « *on peut faire une s  ance, on peut perdre pied (...) mais il faut qu'il argumente, il faut qu'il le justifie, il faut qu'il comprenne (...) il faut qu'il nous propose* » ; ou encore « *une s  ance peut mal fonctionner, peut des fois d  raper (...) mais il faut que le candidat nous l'explique (...) d'une fa  on structur  e, argument  e, objective* ». Sur la fin de l'entretien, le « il faut » fait r  f  rence    l'  valuateur et non plus au candidat, il exprime des valeurs : « *il faut r  duire cette part de subjectivit  * », « *il faut prendre en consid  ration* ». Ainsi l'  valuateur fait donc r  f  rence    des savoirs ou des connaissances. Mais ceux-ci ne s'observent pas en tant que tels : l'  valuateur « essaie de voir » ce que « peut faire » le candidat, c'est-  -dire ce qu'il a la possibilit   de mettre en sc  ne. Les attentes explicites surgissent alors : « il faut que ». En bref, l'observation porte sur le « il peut », par rapport    la situation, cela induit l'inf  rence du « il sait »    partir du moment o   le candidat fait « ce qu'il faut ». L'examineur cr  e ainsi les possibilit  s et analyse la fa  on dont le candidat s'y adapte.

Cette structure appara  t    travers les argumentations des autres interview  s, mais elles commencent parfois par le « pouvoir faire » ou par le « devoir » : « *arriver un peu    cerner ce qu'on peut attendre d'un animateur* » (un professionnel / formateur occasionnel -PFO) ou « *il faut que le projet soit coh  rent dans son ensemble* » (un professionnel - P). Ainsi les ancrages dans l'argumentation s'av  rent diversifi  s : savoirs, capacit  s ou connaissances qui figurent dans le r  f  rentiel, possibilit  s de jugements et d'observation, valeurs fondamentales en termes de « devoir ». Cependant on retrouve toujours certaines constantes, certaines structures qui semblent constitutives de l'argumentation. Les connaissances du candidat : « *comprendre un public, d'apprendre    le conna  tre* », « *il conna  t tous les dispositifs* », « *il conna  t tous les m  canismes* » (P) ; « *elle ne connaissait pas les r  gles* », « *elle ne connaissait pas ce qu'  tait un 100 m  tres* », « *elle ne savait pas se servir d'un chronom  tre* » (P). Mais cela ne suffit pas, il s'agit aussi de « voir » « *s'il sait faire participer tout le monde* », « *comment il sait   couter* » (P). « *On a vu des grands techniciens mais qui ne savaient pas faire passer la technique* » (P). Il est alors fait appel largement, pour appr  cier ces savoir-faire,    la modalisation du « pouvoir ». Il y a tout d'abord ce que peut « voir » l'examineur : « *moi, de ce que j'ai pu voir* » (P), « *donc l   on a pu juger, on a pu appr  cier sa technicit  * » (P), un « pouvoir » qui a ses limites et dont ont conscience les interview  s : « *on peut trafiquer les chiffres* » (P), « *quelques fois, on peut se faire leurrer* » (P), « *il peut se pr  senter sous un tr  s mauvais jour et puis...* », « *Il pourra   tre bien aujourd'hui et tr  s mauvais demain et inversement* » (P). Mais il y a aussi ce que « peut faire le candidat ». « *Qu'est ce qu'on peut faire avec un public, avec l'outil ? Qu'est ce qu'on peut faire techniquement* » (un formateur - F). « *Elle pouvait leur montrer* », « *elle pouvait quitter son atelier    d'heure* », « *quelqu'un qui est    l'aise peut r  pondre* » (P). « *Comment avez-vous pu r  pondre et puis trouver des r  ponses    ses difficult  s* » (F). Le « pouvoir faire » du candidat est toujours situ   par rapport aux possibilit  s de la situation : « *s'il a pu g  rer la prestation du groupe, s'il a pu analyser les ph  nom  nes qui ont pu se suivre pendant la s  ance* » (PFO), « *on peut les rattacher   ventuellement    des ph  nom  nes de groupe* » (F). Le « pouvoir » est ainsi au centre de l'acte d'  valuation ; il s'appr  cie en fonction de ce qui « peut   tre », c'est-  -dire des sch  mas-type implicites    adopter dans la situation pr  sent  e, ce qu'E. Goffman appelle des « cadres de

l'expérience » (1974). Cette modalisation est complexe : elle ouvre sur des sens bien différents en référence aux divers schémas-type du contexte qui déterminent la situation d'examen. Tout d'abord, il y a la position de l'examineur à la recherche « *des éléments concrets que je peux constater pendant la séance* » (F) « *ce serait bien qu'on puisse le vérifier* » (CEPJ). Ces éléments, « *ça peut être une réflexion au détour d'une phrase, ça peut être une attitude, ça peut être un mot, ça peut être, euh, souvent c'est petit* » (P). Mais les choses ne sont pas si simples car, en soi, ces éléments concrets n'ont pas de sens. Ils n'en acquièrent qu'à partir du moment où ils s'insèrent dans une représentation plus globale de la situation, dont plusieurs modes sont exprimés à travers le discours. Cela peut se traduire par des questions pragmatiques : « *qu'est ce qu'on peut faire avec ?* » (F) ; « *en quoi je peux améliorer ma séance ?* » (CEPJ). Il y a aussi ce que « pourrait être » l'animation : « *elles pourraient être un peu plus détaillées* » (F). Un « pouvoir » au conditionnel qui fait référence à la représentation que l'examineur a de la profession : « *ce que peut être le rôle de l'animateur* » (PFO), « *là vous faites l'apprenti sorcier, vous pouvez faire beaucoup de dégâts* » (PFO), « *l'évaluation même des difficultés peut permettre...* » (P). Celle-ci peut alors renvoyer vers des connaissances évoquées plus haut : « *les dispositifs qui peuvent exister* », « *les mécanismes qui peuvent amener...* » (P) : le candidat est-il capable de les mobiliser pour résoudre le problème ? Pour cela, il a besoin d'une analyse correcte de la situation, du moins d'une représentation adaptée : « *qu'est ce que je pensais que je pouvais apporter à mon public avec ce projet là ?* » (F) « *ça peut aussi peut être se passer dans les quartiers tous les jours* » (PFO) ; « *ils peuvent être du fait de la structure* » (P). Tous ces « pouvoir être ou faire » s'inscrivent dans une conception du métier : « *ma conception, c'est de dire qu'une hiérarchie, elle ne peut pas être pyramidale* » (P). « *Je ne peux pas concevoir que quelqu'un...* » (PFO). « *On ne peut jamais tout maîtriser, ce qui est important...* » (F). Et cette conception s'enracine dans un questionnement éthique qui guide les examinateurs : « *est-ce qu'on peut pénaliser* » (P), « *là, en termes de responsabilité morale, je pense que je ne pouvais pas* » (PFO) ; ou encore, elle oriente leur façon d'agir, au moment de l'épreuve : « *on peut faire un boulot intéressant de conseil* » (P), « *s'il y a un endroit où on peut agir là-dessus* » (PFO) ; « *on peut le pousser à franchir le pallier* » (CEPJ). Cette représentation éthique du pouvoir apparaît dans la dernière partie des entretiens, de façon co-occurrence, avec les modalités du devoir, qui expriment les valeurs de l'interviewé. Celles-ci ont une double face : il y a la déontologie, « *ce que doit faire* » l'examineur, mais il y a aussi l'impératif catégorique qui s'adresse au candidat : « *il faut que...* ».

Ainsi, on voit bien apparaître d'une part les « connaissances » et « savoir-faire » dont parlent les taxonomies, concepts qui ont envahi les programmes de l'éducation nationale et les projets de formation, d'autre part les valeurs ou les « visions du monde » dont parlent J. Ardoino et G. Berger (1989). Mais entre ces deux grands modes, l'articulation des situations d'évaluation s'organisent autour du pouvoir : « pouvoir faire » du candidat, mais aussi possibilités d'observation des examinateurs. Ce « pouvoir » est co-construit avec les mêmes concepts que la notion « capable de » (« capable d'argumenter », « capable d'écouter »), structure syntaxique qui vient souvent se substituer au « pouvoir faire ». La « capacité » ne serait-elle pas, de ce fait, ce « pouvoir faire » appliqué aux situations vécues ou décrites par le discours ? Mais ce « pouvoir » ne saurait alors se construire sans faire appel aux représentations que les candidats et les examinateurs ont de la situation. Ce « pouvoir » s'exerce effectivement en fonction des possibilités de la situation, des capacités du candidat à les analyser, à y apporter les types de schémas d'action adéquats, à y intégrer des connaissances qui peuvent apporter des solutions adaptées. Mais aussi l'examineur évoque ce que le candidat « pourrait faire » au regard de sa représentation et des conceptions qui orientent sa façon d'agir. L'évaluation n'est donc pas seulement une capacité que le candidat a ou n'a pas, mais bien une confrontation entre les représentations que le candidat a de la situation, la façon dont il l'analyse et il s'y adapte, et celles d'« experts », c'est-à-dire de professionnels, formateurs ou conseillers techniques et pédagogiques plus expérimentés que lui. A travers tous ces « pouvoirs faire », évoqués au cours de ces entretiens, ne retrouve-t-on pas toutes les définitions de la compétence proposées par G. Le Boterf et par les ergonomes ?

7. Conclusion :

Le sch  ma de la « comp  tence », tel qu'il est con  u par N. Chomsky et qu'il s'est g  n  ralis   dans les milieux de l'enseignement, ne refl  te donc pas la complexit   des processus qui sont mis en place au cours de l'  valuation. Il y a l   une illusion des « experts » sur leurs propres capacit  s, au sens d  fini ci-dessus, c'est-  dire sur les possibilit  s qu'ils auraient d'inf  rer les capacit  s du candidat    partir des « faire » qu'il met en sc  ne. (Il est d'ailleurs int  ressant de remarquer que les principales limites    l'exercice, dans l'exp  rience pr  c  dente, ont   t     voqu  es par des professionnels). Illusion ne veut pas dire que cela n'est pas possible, mais cela ne peut se faire sans prendre appui, r  f  rentiellement, sur les repr  sentations de la situation et la conception du m  tier. L'  valuation est une confrontation de repr  sentations sociales, les unes affirm  es (les « experts »), les autres en construction (les candidats) et non pas une hypoth  tique recherche de comp  tences profondes chez le candidat. Si l'on constate, avec le linguiste (E. Benveniste ; 1966 : 197, 198), que les verbes modaux sont une   volution de nos langues indo-europ  ennes    partir des verbes d'  tat (savoir = j'ai le savoir ; pouvoir = j'ai le pouvoir) et que le verbe « avoir » est lui-m  me une lente transformation    partir du verbe «   tre » (« avoir » est le renversement de «   tre    »), on aper  oit    quel point ces repr  sentations sociales sont fortement ancr  es dans notre inconscient collectif. L'  valuation est un ensemble de modes sociaux qui r  gissent nos relations en soci  t   et qui ont mis des mill  naires    se constituer, et    s'instituer.

8. R  f  rences et bibliographie

- Arduino, J., Berger, G. (1989). *L'  valuation en miettes*. : Matrices Andsha.
- Benveniste, E. (1966) *Probl  me de linguistique g  n  rale. Tome 1*. Paris : Gallimard.
- Brangier, E. & Dubois, N. & Tarquinio, C. (sous la direction de) (1997) *Comp  tences et contextes professionnels*. (Metz : actes du colloque – ADRIPS)
- De Montmollin, M. (1986/1996) *L'ergonomie*. Paris : la d  couverte.
- De Montmollin, M., & Leplat, J. (2001) *Les comp  tences en ergonomie*. Toulouse : Octar  s.
- Dolz, J. & Ollagnier, E. (2002) *L'  nigme de la comp  tence en   ducation*. Bruxelles : De Boeck et Larcier.
- Dupray, A. & Guitton, C. & Monchatre, S. (sous la direction de) (2003) *R  fl  chir la comp  tence*. Toulouse : Octar  s.
- Ghiglione, R., Landre, A., Bromberg, M., Molette, P. (1998) *L'analyse automatique des contenus*. Paris : Dunod .
- Goffman, E. (1974 / 1991) *Les cadres de l'exp  rience*. Paris : Minuit
- Hymes, D.H. (1973, 1982/1991) *Vers la comp  tence de communication*. Paris : Hatier / Didier.
- I.G.E.N. (2007) Les livrets de comp  tences : nouveaux outils pour l'  valuation des acquis. *Rapport N   2007-48 de juin 2007*. Paris : Minist  re de l'  ducation Nationale.
- Le Boterf, G. (1994). *De la comp  tence, essai sur un attracteur   trange*. Paris : Les   ditions d'organisation.
- Le Boterf, G. (2000/2009) *Construire les comp  tences individuelles et collectives*. Paris : les   ditions d'organisation – groupe Eyrolles.
- Malglaive, G. (1990/1993) *Enseigner    des adultes*. Paris : PUF.
- Perrenoud, P. (1984) *La fabrication de l'excellence scolaire : du curriculum aux pratiques d'  valuation*. Gen  ve : Droz.
- Perrenoud, P. (1999) *Dix nouvelles comp  tences pour enseigner*. Paris : ESF.
- Porcher, B. & Letemplier, C. & Rak, I. et ass. (1992) *Du r  f  rentiel    l'  valuation*. Paris : Foucher.
- Sch  tz, A. (1943-1966/1998) *  l  ments de sociologie ph  nom  nologique*. Paris : L'Harmattan.

**AU CARREFOUR DU CURRICULUM PRESCRIT ET DU CURRICULUM IMPLANTE :
POLEMIQUE ET POLYSEMIE AUTOUR DU TERME DE COMPETENCE EN FEDERATION
WALLONIE-BRUXELLES**

Natacha Duroisin*, Sabine Soetewey, Marc Demeuse*****

* Université de Mons. natacha.duroisin@umons.ac.be

** Université de Mons. sabine.soetewey@umons.ac.be

*** Université de Mons. marc.demeuse@umons.ac.be

Mots-clés : Compétence – Curriculum prescrit – Curriculum implanté – Savoir – Savoir-faire.

Résumé. En Fédération Wallonie-Bruxelles, le décret définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre (1997) présente la compétence comme étant une « aptitude à mettre en œuvre un ensemble organisé de savoirs, de savoir-faire et d'attitudes permettant d'accomplir un certain nombre de tâches ». Le terme de compétence a été défini par bon nombre d'auteurs (i.e. Legendre, 2004 ; Roegiers, 2000 ; Tardif, 2006) et sa signification varie au fur et à mesure des interprétations. Par le foisonnement des acceptions, il est difficile d'aboutir à une vision commune, comprise et acceptée par tous. Ainsi, comme l'écrivent Demeuse et Strauven (2006), « la compétence est une notion complexe et nomade ». Proposée à partir des données d'une recherche portant sur l'analyse du curriculum en sciences dans les 2e et 3e degrés de l'enseignement secondaire, cette communication présente, d'une part, une analyse des référentiels de compétences terminales et des programmes d'études en sciences et, d'autre part, des données issues d'entretiens individuels et de questionnaires écrits réalisés auprès d'enseignants à propos de leur compréhension du terme de compétence et de notions qui lui sont associées. Les premiers résultats, concernant le travail d'analyse des documents cadres (curriculum prescrit), démontrent de nombreuses incohérences dans l'utilisation de la terminologie employée. Les autres résultats de la recherche rendent compte du degré de maîtrise par les enseignants sur ces notions clés du « jargon pédagogique » et décrit la place de ces dernières dans leurs pratiques pédagogiques quotidiennes (curriculum implanté).

1. De la liberté (d'enseignement) à la diversité (des programmes)

La Pacte scolaire, dont la loi a été votée en 1959, est le garant de trois grands principes fondamentaux du système éducatif belge : la liberté de choix de l'école par les parents, la fin des tensions entre les réseaux et la gratuité de l'enseignement. Si le pouvoir fédéral a la responsabilité de faire appliquer cette loi, l'organisation de l'enseignement ne fait pas partie de ces prérogatives et est décentralisée au niveau d'entités fédérées que sont les communautés. Le Pacte scolaire a permis de définir deux grands réseaux éducatifs : le réseau officiel et le réseau libre. Chacun de ces réseaux comprend des pouvoirs organisateurs différents. Ainsi, pour les réseaux officiels, le pouvoir organisateur est une personne de droit public, l'organisation de l'enseignement dit « officiel » est réalisée par le réseau de la Fédération Wallonie-Bruxelles (FWB) ou par le réseau des villes et des provinces. Pour les réseaux libres, le pouvoir organisateur est une personne de droit privé et l'organisation de l'enseignement dit « libre » repose sur les réseaux libres confessionnels ou les réseaux libres non confessionnels. L'article 24 de la Constitution permet la liberté d'enseignement. A ce titre, un troisième réseau a été reconnu : le réseau privé. Dans ce cas, le pouvoir organisateur est une personne de droit privé déterminée par l'autorité parentale. En outre, cette liberté d'enseignement réserve la possibilité, pour chaque réseau, de rédiger ses propres programmes d'études, et ce pour chaque discipline, en se reposant sur des documents cadres (tels que le Décret missions) et en veillant à atteindre, au terme des 2e et 3e degrés de l'enseignement secondaire, les exigences prescrites dans les référentiels terminaux. Cette liberté d'enseignement conduit donc inévitablement à une pluralité dans l'approche des thèmes prescrits et, en conséquence, aboutit à une grande diversité des

programmes d'études. En effet, le réseau de la FWB dispose des programmes qu'il définit, le réseau des villes et des provinces se rapporte aux programmes des Provinces et Communes¹ et les réseaux libres appliquent leurs propres programmes. Tout cela laisse entrevoir le nombre important de programmes d'études disponibles et utilisés pour l'enseignement obligatoire. Pour la présente recherche, le travail réalisé porte sur treize programmes d'études différents (en cours d'utilisation en 2010-2012) du réseau de la Fédération Wallonie-Bruxelles.

2. Contexte de la recherche

Dans le cadre d'une recherche financée par le réseau de la Fédération Wallonie-Bruxelles, il est question d'explorer le curriculum en sciences dans les 2e et 3e degrés de l'enseignement organisé par ce réseau d'enseignement, quelles que soit les filières et options choisies, afin de décrire la situation et d'identifier les éléments pouvant être améliorés. Cette exploration a été exécutée en trois phases distinctes. La première consiste en l'analyse du curriculum prescrit (Perrenoud, 1984) ou formel (Demeuse et Strauven, 2006 ; Perrenoud, 1984) à savoir, les documents écrits servant de cadres ou de supports à l'enseignement des sciences tels que les décrets, les référentiels terminaux, les programmes... Cette phase comprend, entre autres, un état de l'art, une approche théorique du curriculum prescrit et de la didactique des sciences. La deuxième phase est caractérisée par l'analyse du curriculum implanté (Audigier, Crahay et Dolz, 2006 ; Soetewey, S., Duroisin, N., Demeuse, M., 2011), c'est-à-dire la manière dont les professionnels de l'éducation traduisent et transposent le curriculum prescrit pour une utilisation concrète. La troisième phase prévoit le rapprochement du curriculum prescrit et du curriculum implanté dans l'optique d'enrichir l'état des lieux et de cibler plus précisément des pistes d'amélioration du curriculum en sciences voire une ré-écriture des programmes d'études. La présente communication se centre sur l'adéquation de la compréhension terminologique d'un nombre défini de concepts pédagogiques clés : le savoir, le savoir-faire et la compétence.

3. Méthodologie de la recherche

Basé sur une méthodologie mixte, le recueil de données s'est effectué en trois étapes. La première étape a consisté en l'analyse théorique (forme et fond) de l'ensemble des programmes d'études et des référentiels de compétences terminales et savoirs requis en sciences. Compte tenu des éléments dégagés lors de cette analyse, la deuxième étape est la réalisation de 23 entretiens individualisés accompagnés d'un questionnaire écrit dont le but est de caractériser le curriculum implanté et d'interroger les enseignants sur la traduction du curriculum prescrit en curriculum implanté. Un échantillon de convenance a permis de privilégier une représentation maximisée des différentes filières d'enseignement au dépend de la représentation des zones d'enseignement puisque le but était l'exploration de la diversité des vécus des enseignants (Duroisin & Soetewey, 2011).

Afin d'explorer, de façon approfondie, cette diversité des vécus, les questions de l'entretien portaient sur sept thématiques différentes et ont été posées de manière successive suivant un canevas en entonnoir : « non dirigé », « semi-dirigé » puis « dirigé ». Le canevas d'entretien débute donc par une question ouverte où le discours de l'interviewé est spontané (non dirigé). La question de départ, commune à l'ensemble des entretiens menés, est la suivante : « Quel est votre avis général sur les programmes de sciences ? ». Pour chacune des thématiques, des questions tantôt plus globales (questions ouvertes), tantôt plus spécifiques (questions ciblées) sont posées. Ce type de canevas d'entretien a été élaboré afin de pouvoir recentrer le dialogue sur les points les plus critiques. L'ensemble des résultats présentés ici porte sur les questions en rapport avec la compréhension des concepts-clés utilisés dans les programmes d'études : le savoir, le savoir-faire et la compétence. En fonction du déroulement des entretiens, la formulation et l'introduction des questions posées pouvaient subir des variations.

Réalisée à l'aide du logiciel NVivo9, l'analyse qualitative des informations recueillies lors des entretiens a notamment permis d'évaluer le degré de compréhension, par les enseignants, des termes issus des programmes d'études, de caractériser les besoins et difficultés des enseignants, de caractériser l'implantation des programmes en classe et de réaliser un inventaire des outils pédagogiques utilisés. En d'autres termes, cette analyse permet d'obtenir une vision plus globale de ce qui est véritablement perçu, vécu et effectué sur le terrain afin de définir des améliorations pouvant être apportées lors de la ré-écriture du curriculum prescrit afin qu'il soit davantage fonctionnel, accessible et cohérent pour ses utilisateurs.

¹ Ou à ceux de la FWB.

4. Résultats de l'analyse sémantique dans le curriculum prescrit

Les premiers résultats présentés sont issus de l'analyse sémantique réalisée sur l'ensemble des documents curriculaires (programmes d'études en sciences, référentiels de compétences pour les humanités générales et technologiques ainsi que pour les humanités professionnelles et techniques, décrets...). A ce niveau, des incohérences et contradictions importantes ont pu être mises à jour. Les seconds résultats émanent de l'analyse qualitative menée à partir des entretiens et questionnaires réalisés. Ces derniers rendent compte de la multiplicité des discours et témoignent de la difficulté qu'éprouvent les enseignants à expliquer les termes clés.

a. Les concepts-clés savoir, savoir-faire et savoir-être

Le dictionnaire encyclopédique de l'éducation et de la formation (2005) définit le savoir comme « un ensemble d'énoncés et de procédures socialement constituées et reconnues » et le savoir-faire comme étant « la capacité à modifier une partie du réel selon une intention, et ce, par les actes mentaux et gestuels appropriés ». Selon Gohier (2006), le savoir-être est quant à lui « la connaissance et l'expression de soi, le respect de l'autre, individu ou collectivité tout aussi bien que l'aspiration du Bien (...) [il] désigne au point de départ l'émergence de l'être, de ses potentialités, virtuellement présentes et ne requérant qu'un environnement stimulant pour être déployées » (p. 182). Les concepts de savoir et de savoir-faire ne sont définis à aucun endroit dans les programmes d'études, ni même décrits en référence à des auteurs. Les seules indications disponibles sont mentionnées sous la forme de tableaux et de listings (voir Figure 1). Si le concept de savoir-être n'est pas défini en tant que tel dans les programmes d'études en sciences, ces derniers mentionnent toutefois des « attitudes » et indiquent que celles-ci « déterminent la manière d'apprendre, d'utiliser ses connaissances, de penser et d'agir. Elles sont indispensables à tout citoyen appelé à prendre une place active dans la vie économique, sociale et culturelle ».

Par la

Figure 1, présentée ci-dessous, est présenté un exemple de listings d'apprentissage, issu du programme d'études de physique de 4^e transition.

B. Niveau Sciences générales	
1. Mouvements	
Prérequis - Relation entre déplacement et durée dans le MRU et dans le MRUV. - Tracés et interprétations de graphiques. - Enoncés de lois mathématiques à partir de mesures.	Exemples de questionnement - Quelle est l'accélération d'une voiture de course au démarrage ? Est-elle plus importante que celle d'une pomme qui tombe ? - Quel est le mouvement d'une pierre lancée ? Jusqu'où peut-elle aller avant de toucher le sol ? - Dans quelle direction nager pour traverser une rivière perpendiculairement aux rives ?
Compétences	Savoirs
- Analyser le mouvement d'un objet lors d'une chute libre, sur un rail incliné. - Déterminer la valeur de g. - Calculer des vitesses et des accélérations. - A partir d'un énoncé ou d'un tableau, tracer et exploiter un graphique de position, vitesse ou accélération en fonction du temps. - Donner l'ordre de grandeur de vitesses et d'accélérations dans quelques exemples de la vie quotidienne. - Associer la vitesse à la pente du graphique $x(t)$, l'accélération à la pente du graphique $v(t)$. - Interpréter des graphiques de position, vitesse et accélération en fonction du temps et retrouver l'un à partir de l'autre. - Etablir et discuter les équations des mouvements étudiés.	Mouvements rectilignes : MRU, MRUV, ... Position, vitesse moyenne et vitesse instantanée. Lois du MRU et du MRUV. Accélération moyenne et accélération instantanée dans le cas d'un mouvement rectiligne quelconque. Mouvements à deux dimensions Composition de deux mouvements uniformes Tir horizontal.
Programme de Physique de 4 ^e transition	

Figure 1 – Extrait du programme d'études en sciences physique de 4^e transition

Tout comme le terme de savoir, celui de savoir-faire apparaît à de nombreuses reprises dans les programmes d'études en sciences sans pour autant y être défini ou préalablement présenté (

Figure 2).

Activité 1 : Force d'Archimède dans les liquides			
Objectifs de l'activité			
- Rappeler la notion de forces pressantes exercées par l'eau (et tous les liquides) sur les surfaces du récipient qui la contient (surfaces pressées). - Découvrir que les forces pressantes exercées par l'eau agissent aussi sur les corps : - partiellement immergés (corps flottants) ; - complètement immergés. - Découvrir les paramètres qui influencent la force d'Archimède. - Effectuer des mesures afin d'estimer la valeur de la force d'Archimède. - Conclure en dégagant les 3 étapes conduisant à la loi d'Archimède : - un objet dans un liquide déplace un volume de liquide ; - ce volume de liquide déplacé a un poids ; - la valeur de la force d'Archimède est égale à la valeur de ce poids.			
Objectifs de savoirs			
- Découverte de la notion de corps flottants - immergés. - Découverte de la force d'Archimède. - Découverte de ses caractéristiques (à l'exception du point d'application) et des paramètres d'influence. - Applications.			
Objectifs de savoir-faire			
- Dégager les variables dépendante et contrôlée lors d'une expérience. - Tracer un graphique et dégager le coefficient de proportionnalité. - Transférer certaines notions dans d'autres contextes.			
Physique	2 ^e degré - 3 ^e année	Module 3 - niveau A	61

Figure 2 – Extrait issu du programme d'études en sciences physique de 3e année

b. Le concept-clé compétence

Si Tardif (2006) définit la compétence comme « un savoir-agir complexe prenant appui sur la mobilisation et la combinaison efficaces d'une variété de ressources internes et externes à l'intérieur d'une famille de situations », d'autres auteurs, qui mettent en évidence le caractère global, évolutif ou inédit du terme, s'entendent sur d'autres définitions (Beckers, 2002 ; Depover, C. et Noël, B., 2005 ; Legendre, 2004 ; Perrenoud, 1999 ; Roegiers, 2000). Les textes fondateurs du système éducatif belge proposent quant à eux des définitions du mot *compétence* et des notions qui lui sont associées. Dans les programmes d'études de sciences, la définition du terme de compétence est reprise de l'article 5 du décret missions qui la décrit comme une « aptitude à mettre en œuvre un ensemble organisé de savoirs, de savoir-faire et d'attitudes permettant d'accomplir un certain nombre de tâches ». Comme présenté en Figure 3 - Extrait du programme de sciences chimie du deuxième degré, cette définition est complétée par une représentation du concept sous la forme d'un organigramme. Parallèlement aux informations directement fournies dans les documents curriculaires, les conseillers pédagogiques ont pour mission (lors de formation continue, par exemple) de compléter cette définition du terme en mettant en avant le caractère inédit de la tâche ainsi que le fait qu'elle soit à réaliser sans aide pédagogique particulière.

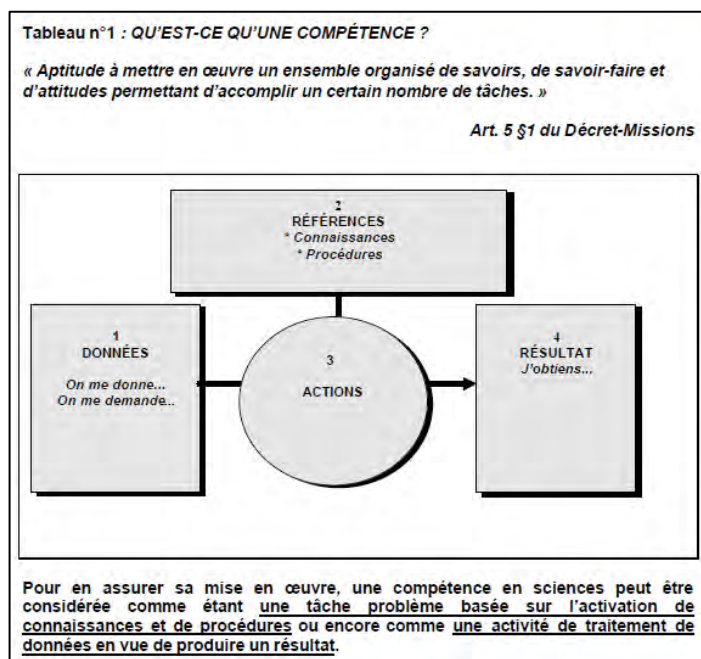


Figure 3 - Extrait du programme de sciences chimie du deuxième degré

L'analyse effectuée fait remarquer qu'à l'intérieur même des programmes d'études et des référentiels de compétences terminales et savoirs requis en sciences pour les humanités générales et technologiques, les énoncés rédigés en termes de compétence, de savoir, de savoir-faire ou encore d'attitude sont « confondus » à plusieurs reprises. Pour exemples, dans le référentiel de *compétences terminales et savoirs requis en sciences pour les humanités générales et technologiques*, les éléments mentionnés ci-dessous sont appelés « compétences spécifiques » (Tableau 1) :

Tableau 1 – Compétences spécifiques issues du référentiel de compétences terminales et savoirs requis en sciences pour les humanités générales et technologiques

Compétences spécifiques	Avis de l'équipe de recherche
➤ Utiliser une clé dichotomique ➤ Représenter dans le plan une molécule d'hydrocarbure saturée ➤ Connaître les principes de fonctionnement du corps humain ➤ Savoir que la croissance, l'émotivité, les cycles sexuels sont sous la dépendance d'hormones ➤ Utiliser un multimètre	Savoir-faire Savoir-faire Savoir Savoir Savoir-faire

Par ces exemples, on constate que des éléments repris sous l'appellation de compétences spécifiques n'en sont pas pour autant. En effet, les notions « Utiliser une clé dichotomique », « Représenter dans le plan une molécule d'hydrocarbure saturée » et « Utiliser un multimètre » s'apparentent à des savoir-faire et non à des compétences. De même, il paraît évident que « Connaître les principes de fonctionnement du corps humain » et « Savoir que la croissance, l'émotivité, les cycles sexuels sont sous la dépendance d'hormones » sont des savoirs.

5. La place des concepts-clés dans le curriculum implanté

Dans le Tableau 2 sont présentés quelques exemples de définitions fournies par les enseignants lors des entretiens individuels. Les questions sur lesquelles portent les résultats présentés ci-après concernent les termes de savoir, savoir-faire, compétence et sont reprises sous le libellé « Selon vous, qu'est-ce qu'... ? ». La liberté est laissée aux enseignants de préciser le terme en formulant une définition, en énonçant des exemples ou en proposant une définition argumentée d'exemples. La colonne de droite intitulée « degré de fiabilité » renseigne sur l'exactitude des définitions fournies par les enseignants par rapport à l'avis de l'équipe de recherche qui se base sur le corpus de définitions précédemment énoncé.

Tableau 2 – Exemples de définitions du concept « savoir » fournies par les enseignants lors des entretiens

	Définitions fournies par les enseignants	Degré de fiabilité
a.	« Un savoir? Ben, c'est une connaissance de, de, une connaissance de la matière dans le sens où, euh, c'est une loi, c'est une, euh, c'est tout ce qui est à mémoriser. »	Forte fiabilité, référence à la connaissance, à la mémorisation
b.	« Ben pour moi ça représente les connaissances scientifiques, je ne sais pas si je réponds à votre question, les connaissances scientifiques de base hum sur lesquelles on peut s'appuyer pour avancer dans une démarche vraiment de type scientifique »	Forte fiabilité, référence à la connaissance scientifique
c.	« Le savoir, ce sont les choses qu'ils doivent, euh, quand même, euh, et de un, euh, connaître au départ, et, pouvoir, éventuellement, les améliorer. »	Fiabilité moyenne
d.	« Un savoir? Ben c'est à partir du moment où il a compris, je vais dire, compris, où il a su faire le lien entre ce qu'il a appris en classe et la pratique, parce que vous savez, vous apprenez également à bien connaître le code de la route, si vous conduisez la voiture comme un sauvage, ça sert à rien... »	Fiabilité faible, confusion avec la pratique

Il apparaît que sept enseignants fournissent des définitions fiables (comme a. et b.) faisant référence aux connaissances, définitions, lois... Plus d'une dizaine d'enseignants donnent une définition approximative de ce concept (c.). Cependant, deux enseignants définissent de manière incorrecte le terme demandé (d.).

À propos du terme de « savoir-faire », des exemples d'explications fournis par les enseignants sont proposés dans le Tableau 3.

Tableau 3 – Exemples d'explications du concept « savoir-faire » donnés par les enseignants lors des entretiens

	Explications fournies par les enseignants	Degré de fiabilité
a.	« Savoir-faire c'est juste un exercice du même type qu'ils font en classe par exemple si je reprends mes exercices de nomenclature si je pose des exercices de nomenclature en classe que ça soit alcane ou alcène en 5 ^e ou alcool et acide en rhéto je refais les mêmes exercices ça reste du savoir-faire. »	Forte fiabilité, référence à l'exécution d'exercices
b.	« Et un savoir-faire c'est de la, c'est de la mise en oeuvre. C'est de la pratique, en utilisant le vocabulaire qui a dû être étudié. Ou une démarche qui a été travaillée au cours, etc. Ca, c'est un savoir-faire. Appliquer une démarche sur un, pour, pour pouvoir résoudre un problème. »	Forte fiabilité, référence à l'application de savoirs
c.	« Tandis qu'un savoir-faire, c'est plus une démarche, par exemple construire un graphique, faut placer les..., oui donc c'est plutôt connaître la démarche... »	Forte fiabilité, référence à l'acquisition d'une démarche
d.	« Euh j'ai plus facile à travailler un savoir-faire parce que c'est, pour moi c'est du conditionnement, donc je leur montre comment ça se passe, je leur donne un exercice qu'ils font avec moi, au fur et à mesure, et puis la deuxième fois, je les laisse se tromper, donc je passe entre les bancs, ça ça va, ça ça va pas et la troisième fois, je commence à vérifier et à vraiment commencer à noter... Et donc après, je les drille, devoir sur devoir, et ça passe... »	Fiabilité moyenne, référence à l'exécution d'exercices

L'analyse qualitative des discours des enseignants permet de constater que ces derniers se réfèrent à trois notions distinctes pour expliquer ce qu'est un savoir-faire. Selon neuf enseignants, le savoir-faire peut être partiellement ou totalement défini comme l'exécution d'exercices (comme a. et d.). Sept enseignants mentionnent que le savoir-faire fait appel à l'application de savoirs (b.) tandis que quatre enseignants indiquent que le savoir-faire s'apparente davantage à l'acquisition et la mise en œuvre d'une démarche dans une perspective de résolution de problèmes (c.).

Concernant le terme de compétence, les explications recueillies chez les enseignants se réfèrent à trois modèles distincts. Le premier d'entre eux est celui présenté par le décret missions et complété par le message des conseiller pédagogique qui mentionnent le caractère inédit et non-pédagogique de la tâche, le deuxième porte sur les éléments énoncés par le décret missions. Le troisième et dernier modèle évoqué est celui de l'intersection de trois cercles « savoir – savoir-faire – attitude ». Compte tenu de ces approches différentes, les réponses fournies par les enseignants ont été traitées en fonction de celles-ci.

Tableau 4 – Exemples d’explications du concept « compétence » fournis par les enseignants lors des entretiens

Explications fournies par les enseignants	Degré de fiabilité
« ce qu'on entend par compétence c'est utiliser plusieurs savoirs, associés à plusieurs savoir-faire dans une situation nouvelle. Donc, quelque chose qu'ils n'ont jamais rencontré. »	Fiabilité forte, modèle du décret missions complété par les messages des conseillers pédagogiques
« Une compétence c'est un, c'est l'application de savoir-faire et de savoirs, dans une situation inconnue, voilà. Donc ils ont acquis des savoir-faire et des savoirs, pendant le cours, et il faudrait une situation nouvelle, dans lequel ils utilisent ce qu'ils ont acquis. »	Fiabilité forte, modèle du décret missions complété par les messages des conseillers pédagogiques
« Ben c'est l'intersection entre le savoir, le savoir-faire et les attitudes, c'est ça... Oui, savoir, savoir-faire, attitudes, c'est l'intersection. »	Fiabilité forte, modèle de l'intersection des trois cercles
« ... une compétence, puisqu'une compétence est un ensemble de savoirs, savoir-faire et savoirs-être, dans une seule phrase on ne sait pas... »	Fiabilité forte, modèle de l'intersection des trois cercles
« Ben c'est, c'est, c'est la ligne de conduite principale à tenir, c'est le contenu principal de la matière, que l'élève doit comprendre, c'est le franc qui doit tomber... »	Fiabilité très faible, définition erronée
« C'est d'arriver à inculquer à un élève une matière, euh (Petite pause), et qu'il puisse la trouver, pas lui donner et lui imposer, c'est qu'il puisse, euh, après, euh, diverses choses arriver à, par lui-même à la trouver. »	Fiabilité très faible, définition erronée

Huit enseignants interrogés décrivent ce qu’est une compétence en se rapportant, d’une part, à la définition énoncée dans le décret missions et, d’autre part, au message des conseillers pédagogiques tandis que cinq autres font allusion, au moins partiellement, au décret missions. Les informations contenues dans le Tableau 5 permettent de rendre compte des éléments de définition fournis par chaque enseignant qui se sont rapportés de manière implicite au décret et/ou au discours des conseillers pédagogiques.

Tableau 5 - Eléments de définition fournis par chaque enseignant, pour ceux se référant de façon implicite au décret missions et au message des conseillers pédagogiques

	DECRET MISSIONS + MESSAGE DES CONSEILLERS PEDAGOGIQUES								DECRET MISSIONS				
ELEMENTS DE DEFINITION	Profil de chaque enseignant												
Capacité													V
Mettre en œuvre		V	V				V	V					
Savoirs	V	V	V	V	V	(v)	V	V	V	V	V	V	V
Savoir-faire	V	V	(v)	V	V	(v)	V	V	v	V	v	V	
Attitudes	V												
Tâche à accomplir				V	V	V				V			
Inédite	V	V	v	v	V	V	V	V					
Non pédagogique	V												

Le modèle de l’intersection des trois cercles permet à trois enseignants d’expliquer ce qu’est une compétence en mettant en avant le fait que la compétence désigne le croisement entre le savoir, le savoir-faire et le savoir-être.

Concernant les autres réponses des enseignants, trois d’entre eux, disposant d’un Certificat d’Aptitude Pédagogique (CAP) ou équivalent, ont fournis des explications erronées du terme de compétence. Un enseignant agrégé donne, quant à lui, un exemple inadéquat en mentionnant que la notion de valence est une compétence disciplinaire. Enfin, un enseignant met en avant une série d’exemples de compétences issue du programme

d'études du 1^{er} degré et un dernier propose une explication tellement confuse du concept-clé qu'il est difficile d'en extraire une quelconque véracité.

6. Le degré de certitude et de maîtrise relatif au concept de compétence

Les discours des enseignants interrogés permettent aussi de mettre en évidence les certitudes ou incertitudes concernant l'appropriation des termes utilisés dans les programmes d'études ou dans d'autres documents curriculaires. Force est de constater que chez 14 des enseignants interrogés, les explications du terme demandé sont empruntées de nombreuses marques d'hésitation et d'incertitude. Le Tableau 6 présenté ci-dessous comporte quelques-unes des retranscriptions des discours des enseignants lorsque ces derniers évoquent la compréhension qu'ils ont du terme de compétence.

Tableau 6 - Extraits des discours des enseignants qui permettent de mettre en évidence les incertitudes relatives à la compréhension de ce qu'est une compétence

Extraits issus des retranscriptions effectuées
« (...) [Dans] ce bouquin « socle de compétence », ce ne sont pas des compétences. Et dans le programme, on a la même chose et donc on est en désaccord, on utilise un même terme pour deux choses tout à fait différentes, [...] euh compétences en math, compétences en chimie, en physique, en bio, compétences en français ou en langues, c'est le même mot pour des choses totalement différentes... »
« Donc comment voulez-vous libeller tout ça? Nous, entre collègues de sciences, on estime que c'est impossible à libeller... Donc quand les collègues le font, (...) ils ne font que libeller un savoir-faire (...) »
« Les compétences, il n'y a aucune compétence, fin d'après ce que j'ai compris, il n'y en a aucune dans le programme hein. Les compétences c'est quelque chose qu'on crée. (...) on utilise beaucoup le terme "compétence", mais euh en tant que prof, je m'emmêle un peu les pinceaux. On me parle de compétences, de savoirs, de savoir-faire, et de motivations d'élève, au milieu c'est la compétence (...) je suis un peu perdue moi ».
« Voilà, je vais vous traduire ce qu'on m'a expliqué et que j'ai enfin je pense avoir accepté, je pense... »
(Pause). « On se dit mais 'Qu'est-ce qu'ils nous veulent?' ».
En fait, ça dépend, c'est très difficile à définir, une compétence.
« J'espère que je ne me trompe pas, parce que bon, c'est toujours difficile »

Dans le but de constater si les enseignants maîtrisent les concepts-clés issus des programmes d'études, ceux-ci ont été confrontés à une série d'énoncés qu'ils ont dû identifier comme étant des savoirs, des savoir-faire, des compétences ou appartenant à une catégorie nommée « autre/inconnu ». 16 énoncés, provenant des programmes d'études en sciences, ont été sélectionnés pour cet exercice de maîtrise comparée. Les référentiels identifient 9 de ces énoncés comme étant une compétence (ex. Connaître les principes de fonctionnement du corps humain) et 7 autres comme un savoir (ex. Notion de composé organique). L'équipe de recherche a, quant à elle, identifié ces mêmes énoncés et compte parmi eux 3 compétences, 9 savoirs, 3 savoir-faire (ex. Connaître les principes de fonctionnement du corps humain) et un « autre ». Il a été remarqué que l'identification effectuée par les enseignants est plus proche de celle réalisée par l'équipe de recherche que celle provenant des référentiels. Le degré de maîtrise comparée permet de remarquer que la confusion existante dans l'utilisation des termes a un impact important sur l'identification des énoncés par les enseignants. En effet, les énoncés identifiés par les référentiels ou l'équipe de recherche comme étant des savoirs sont reconnus comme tel par les enseignants alors que la distinction concernant les énoncés compris comme étant des compétences reste plus difficile à établir.

7. Conclusions

Au-delà du fait que la compétence soit qualifiée de « complexe, globale et intégrative, interactive ainsi qu'évolutive » (Legendre, 2004), la difficulté qu'éprouvent les enseignants à composer avec ce terme a ici été explicitement constatée. Cette observation n'est pas propre au réseau étudié ; en effet, Hirtt (2005) dresse également un constat sévère par rapport à l'un des autres réseaux d'enseignement. Les données issues des entretiens individuels et des questionnaires écrits ont permis d'apprécier le degré de concordance ou de discordance existant entre les deux types de curriculums étudiés, d'explorer la diversité des vécus des enseignants dans la transformation du curriculum prescrit en curriculum implanté et de percevoir l'éventuel impact de la compréhension de la terminologie sur la pédagogie mise en œuvre. Il faut donc constater que la compréhension et l'appropriation des concepts-clés « savoir » et « savoir-faire » ne posent pas de difficulté particulière aux enseignants interrogés alors que ces derniers ne sont, à aucun endroit dans les programmes

d'études ou autres référentiels, définis de manière claire et précise en prenant appui sur un corpus de connaissance. Par contre, il n'en va pas de même pour le terme de « compétence » qui lui est sujet à une plus forte controverse. Malgré le fait que la plupart des programmes d'études proposent la définition du décret missions voire même des organigrammes explicatifs, on constate une multitude et une variabilité des significations de ce terme ainsi qu'un degré d'incertitude assez élevé pour plus de la moitié des enseignants interrogés. Sur la base des éléments dégagés, des pistes de travail pour la ré-écriture de programme d'études devront ainsi être envisagées.

8. Bibliographie

- Audigier, F., Crahay M. et Dolz J. (2006). *Curriculum, enseignement et pilotage*. Bruxelles : De Boeck.
- Beckers, J. (2002). *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*. Bruxelles : Labor.
- Demeuse, M. & Strauven, C. (2006). *Développer un curriculum d'enseignement ou de formation*. Bruxelles : De Boeck.
- Depover, C. & Noël, B. (2005). *Le curriculum et ses logiques*. Bruxelles : De Boeck.
- Duroisin, N. & Soetewey, S. (2011). *Méthodologie de la recherche*. Université de Mons
- Gohier, C. (2006). Les sources du savoir-être comme finalité éducative : former à être, une injonction paradoxale ? *Nouvelles pratiques sociales*, vol. 18, n°2, p. 172-185.
- Hirtt, N. (2005). Pédagogie de l'incompétence. Les nouveaux programmes de physique dans l'enseignement catholique francophone belge. *L'école démocratique*. Consulté le 30 juin 2011 de <http://www.skolo.org/spip.php?article275>
- Legendre, M. – F. (2004). Cognitivism et socioconstructivisme : des fondements théoriques à leur utilisation dans l'élaboration et la mise en œuvre du nouveau programme de formation. In Philippe Jonnaert et Armand M'Batika, *Les réformes curriculaires : Regards croisés*. (p.15-47). Québec : Presses de l'Université du Québec. 2004.
- Ministère de la Communauté française de Belgique. (1997). *Décret définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre*.
- Perrenoud, Ph. (1984). *La fabrication de l'excellence scolaire : du curriculum aux pratiques d'évaluation*. Genève : Droz, (2^e édition augmentée 1995).
- Perrenoud, Ph. (1999). Transférer ou mobiliser ses connaissances ? D'une métaphore à l'autre : implications sociologiques et pédagogiques, *Raisons éducatives* (2). Récupéré le 29 juin 2011 de http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1999/1999_28.html
- Roegiers, X. (2000). *Une pédagogie de l'intégration: compétences et intégration des acquis dans l'enseignement*. Paris-Bruxelles: De Boeck Université.
- Soetewey, S., Duroisin, N., Demeuse, M. (2011). Le curriculum oublié. Analyse comparée des programmes de sciences en Belgique francophone. *Revue internationale d'Education. Sévres*, n°56, 123-134.
- Tardif, J. (2006). *L'évaluation des compétences. Documenter le parcours de développement*. Montréal : Chenelière Éducation, p.22.

COMMENT EVALUER LE DEVELOPPEMENT DE COMPETENCES ETHIQUES EN FORMATION INITIALE D'ENSEIGNANTS ?

Antoine Sautelet*

* Institut Supérieur de Pédagogie Galilée antoine.sautelet@galilee.be, Université Catholique de Louvain antoine.sautelet@student.uclouvain.be

Mots-clés : évaluation – compétences – formation initiale des enseignants – éthique – développement moral

Résumé. Cette communication présente un outil de mesure développé spécifiquement pour évaluer la compétence éthique des étudiants en formation initiale d'enseignants. Basé sur une approche développementale de la morale et de l'éthique (Kohlberg, 1981) et sur la théorie du double aspect des comportements éthiques de Lind (2008), cet outil devrait permettre aux chercheurs et aux formateurs d'approcher la compétence éthique professionnelle des enseignants et des futurs enseignants.

Introduction

De nombreux auteurs tels que Gohier (1997), Moreau (2009), Desaulniers et Jutras (2006) insistent sur la nécessité de prendre en compte la dimension éthique dans la formation des enseignants. En Communauté Française de Belgique, le décret Dupuis (12/12/2000), a instauré un référentiel pour la formation initiale des enseignants qui comporte notamment la compétence « mesurer les enjeux éthiques liés à sa pratique quotidienne ». Dans le cadre de la formation des enseignants en haute école, différents cours et modules visent le développement de cette compétence.

L'évaluation de la compétence éthique visée par ces cours pose question. Celle-ci se fait souvent au travers de rapports réflexifs sensés retracer le développement de la compétence chez l'étudiant, ses réflexions, son évolution. Une des limites de ce type de rapport écrit, est la tendance de certains étudiants à écrire ce que l'évaluateur veut lire et non pas toujours sa propre réflexion. Ce problème se révèle particulièrement présent dans les questions éthiques (Johnson, 2008).

Dans une recherche récente, Johnson (ibidem) a montré que les futurs enseignants semblent parfois faire preuve de compétences éthiques élevées en répondant à des questions types telles que « quelles sont les responsabilités éthiques des enseignants ? ». Toutefois, lorsqu'ils se retrouvent en face de dilemmes éthiques concrets, ils éprouvent beaucoup plus de difficultés à raisonner par eux-mêmes. Selon Johnson, si les étudiants peuvent facilement répéter des discours éthiques proches de ceux qu'ils entendent durant leur formation, cela ne veut pas dire qu'ils savent mesurer les enjeux éthiques ni mettre en œuvre leurs valeurs dans des réflexions à dimensions éthique. Johnson (ibidem) nous enjoint à ne pas confondre une « régurgitation éthique » avec une « réflexion éthique ».

Un des objectifs de l'outil présentés dans cette communication est de parvenir à mesurer le développement d'une compétence en raisonnement éthique qui se distinguerait clairement de la capacité des étudiants à répéter un discours éthique entendu pendant leur formation.

I. Cadre théorique

L'outil se base sur l'approche développementale de l'éthique et de la morale de Lawrence Kohlberg (1981) et sur la compétence de jugement éthique telle qu'elle a été définie par Georg Lind (2008).

1.1 Les stades du jugement moral selon Kohlberg

Kohlberg (1981) distingue trois formes d'opération caractérisant la cognition morale, trois niveaux différents qui correspondent à trois logiques de construction du sens dans le domaine moral : le niveau **préconventionnel**, le niveau **conventionnel** et le niveau **postconventionnel**. Selon Kohlberg, un individu au niveau postconventionnel est capable de prendre des décisions éthiques en tenant compte des perspectives des autres membres de la société et en se basant sur des principes universels de justice qu'il préférera à l'application pure et simple de règles juridiques ou de normes sociales. Ces 3 niveaux sont eux-mêmes chacun divisés en 2 stades. Le développement moral s'opérerait, selon Kohlberg, du premier stade au sixième stade en suivant un ordre bien déterminé :

Stade	Niveau
1er stade : orientation punitive et obéissance à l'autorité (obéissance aux personnes qui disposent du pouvoir, afin d'éviter les punitions) ;	Préconventionnel
2ème stade : morale individuelle et instrumentale (but pragmatique : maximiser la satisfaction de ses besoins et désirs et en minimiser les aspects négatifs)	
3ème stade : orientation en fonction des relations interpersonnelles (l'individu considère bien agir s'il est gentil et s'il plait à autrui)	Conventionnel
4ème stade : moralité de l'autorité et du maintien de l' ordre social (le sujet ne se conforme plus aux standards d'autres individus, mais bien à l'ordre social en tant que tel.)	
5ème stade : les droits individuels et la loi démocratiquement acceptée (La moralité se fonde sur des accords entre individus, qui désirent respecter des normes qui leur paraissent nécessaires pour maintenir l'ordre social et les droits des personnes) ;	Postconventionnel
6ème stade : orientation des principes éthiques universels (les principes universels de justice, réciprocité, égalité, respect représentent un idéal intériorisé qui guide les actions du sujet indépendamment des réactions d'autrui).	

Tableau 1 : stades de développement moral selon Kohlberg (1981)

1.2 La compétence de jugement éthique (Lind 2008)

Cette compétence de jugement éthique est liée à « la capacité de l'individu de percevoir les implications éthiques d'une situation et d'organiser et appliquer de façon cohérente des règles et des principes éthiques à des situations concrètes » (Lind, 2008). Dans cette optique, le comportement éthique est fondé sur des principes éthiques propres aux individus et non pas sur des normes sociales externes ou imposées. Cette vision des comportements éthiques intègre les aspects affectifs et comportementaux. La compétence de jugement éthique, inclut, en outre, la capacité de réflexion critique et de discours rationnel.

Un des intérêts de l'approche de Lind est qu'elle permet de distinguer clairement les aspects affectifs des aspects cognitifs de l'éthique. Les aspects affectifs étant définis comme les idéaux, les principes et les valeurs des individus ; tandis que les aspects cognitifs concernent les capacités cognitives à mettre en œuvre des idéaux ou des principes éthiques. C'est cette capacité à mettre en

en œuvre des principes éthiques ou moraux propres à chacun que Lind définit comme compétence éthique.



Figure 1 : la compétence éthique selon Lind (2008)

Pour illustrer cette conceptualisation, prenons un exemple concret issu du test MJT (Test de jugement éthique) conçu par Lind pour évaluer cette compétence en jugement éthique. Face au dilemme d'un médecin qui accepte d'euthanasier une patiente, les sujets peuvent être ou non avec le comportement du médecin, ce n'est pas cela qui déterminera une compétence en jugement éthique. Mais la façon dont ils jugent des arguments en faveur et en défaveur du comportement du médecin est un bon indicateur de cette compétence. Un sujet qui juge des arguments en fonction du niveau de raisonnement de cet argument démontre une compétence de jugement éthique élevée. Tandis que celui qui juge des arguments en fonction de sa position (pour ou contre le comportement du médecin) sans tenir compte des principes sous-jacents (du niveau de raisonnement moral modélisé par Kohlberg) révèle une compétence de jugement éthique faible.

Ainsi, à un niveau **conventionnel** de l'échelle de Kohlberg (1981), on peut être **d'accord** avec l'euthanasie parce que beaucoup de médecins l'acceptent déjà ou bien on peut être **en désaccord** parce que sa religion l'interdit. Au niveau **postconventionnel**, on peut **accepter** l'euthanasie par respect pour la dignité humaine ou bien la **refuser** en estimant que la vie est plus importante que les autres considérations. Les sujets qui sont en accord (ou en désaccord) avec les arguments cités ci-dessus, uniquement parce qu'ils sont soit « pour », soit « contre » l'euthanasie, obtiendront un score de jugement éthique faible. Alors que les sujets qui jugent les arguments en fonction de leur niveau de raisonnement (peu importe qu'ils soient pour ou contre l'euthanasie) obtiendront un score de compétence morale élevée.

II. Création d'un test de jugement éthique original

2.1. Un test basé sur le MJT (Lind, 2008)

Le test créé s'inspire du fonctionnement MJT développé par Lind (2008) mais est composé d'items originaux créés pour répondre aux spécificités de la formation initiale des enseignants. Il existe de nombreux tests pour mesurer le développement éthico-moral des mais le système du MJT imaginé par le Professeur Lind nous paraît être un des plus intéressants pour deux raisons : 1) il ne se cible pas sur les valeurs ou les idéaux des étudiants mais sur leur aptitude à mettre ces valeurs en pratique dans une réflexion. 2) il est articulé autour d'une tâche éthique à résoudre et en ce sens permet d'évaluer une compétence morale.

Le MJT a été créé par Lind afin d'évaluer la compétence en jugement éthique telle que Kohlberg l'a définie : la capacité à prendre des décisions et des jugements basés sur des principes internes et d'agir en accord avec ces jugements (Kohlberg, in Lind, 2008). Outre la compétence en jugement éthique (qui correspond à l'aspect cognitif du comportement éthique) le MJT, mesure également les préférences morales du sujet (les stades kohlbergiens de cognition morale préférés). Ces deux aspects sont évalués simultanément, mais de façon indépendante. Le MJT offre donc des mesures pour chacun de ces deux aspects. L'intérêt de mesurer ces deux aspects de façon distincte est que ceux-ci ne sont pas forcément liés. Une personne peut très bien valoriser les valeurs universelles de justice, mais être incapable de les utiliser avec cohérence lorsqu'il évalue une position morale avec laquelle il est en désaccord.

Le MJT comporte deux histoires comportant un dilemme moral. Pour chacun de ces dilemmes, le test propose six arguments qui justifient le comportement du protagoniste de l'histoire et six autres qui sont en désaccord avec ce comportement. Chacun de ces arguments est construit en référence à un stade Kohlbergien de raisonnement moral. Les sujets doivent 1° se positionner en accord ou en désaccord avec le comportement des protagonistes de l'histoire et 2° évaluer dans quelle mesure les arguments en faveur ou en défaveur du comportement des protagonistes sont acceptables ou non (entre pas du tout acceptable et tout à fait acceptable).

Le score principal obtenu par ce test, l'index-C, est compris entre 0 et 100 et mesure la consistance des jugements émis par les sujets à propos des différents arguments. Un individu qui valide les arguments en référence aux principes qu'ils sous-tendent (les stades de raisonnement moral) obtiendra un haut score de compétence en jugement moral. Ainsi un sujet qui attribue de façon consistante un score élevé aux arguments d'un niveau spécifique aura un score C élevé.

Le MJT est un outil qui a été utilisé dans de nombreuses recherches et qui, au cours de celles-ci, a acquis un degré de validité élevé, ainsi que de nombreuses données permettant des comparaisons variées. Toutefois, les dilemmes qu'il propose sont fort éloignés de la réalité professionnelle des enseignants. Le premier concerne une situation de conflit social entre une direction et des ouvriers et l'autre un cas d'euthanasie pratiqué par un médecin.

Les recherches autour des compétences professionnelles ont régulièrement mis en avant l'importance de la contextualisation des tâches qui visent à mesurer des compétences. Une compétence professionnelle se mesure au mieux dans un contexte le plus authentique possible, le plus proche de la réalité professionnelle concernée. Au vu de ces considérations, il nous semblait nécessaire de concevoir un outil basé sur le fonctionnement du MJT mais dont les dilemmes proposés aux sujets se dérouleraient dans un milieu scolaire.

2.2. Construction de l'outil

Pour construire cet outil, nous avons commencé par élaborer deux dilemmes moraux susceptibles de se dérouler dans une école primaire ou secondaire. Le premier concerne la question de la dénonciation d'un comportement non adapté d'un collègue à la direction. Et le deuxième à trait à l'évaluation d'une élève. Nous avons scrupuleusement suivi les consignes de Lind (2008) pour la création de ces dilemmes. Ce dernier estime, en effet que pour être opérationnel, un dilemme doit :

- raconter l'histoire d'une personne fictive qui est susceptible d'exister et qui doit avoir un prénom ;
- montrer que la décision prise par le sujet comporte une difficulté ;
- comporter une décision claire du sujet ;
- être court et aisément compréhensible ;
- doit être intéressant pour des raisonnements moraux élevés (stades 5 et 6 de Kohlberg) mais aussi pour des raisonnements plus bas ;
- doit permettre des arguments intéressants, aussi bien pour les sujets qui sont en accord avec la décision du protagoniste que pour les sujets qui sont en désaccord avec elle.
- Une fois les deux dilemmes rédigés, nous avons conçu les six arguments en accord avec le comportement du protagoniste et les six arguments en désaccord, et ce pour les deux dilemmes, soit 24 arguments.

Ici aussi, nous avons suivi les prescriptions du professeur Lind (ibidem), à savoir que les arguments doivent:

- être aussi courts que possible, de préférence une phrase, pas trop complexe et pas trop technique ;
- représenter le plus possible chacun des six stades Kohlbergiens ;
- être tous abordés positivement (ils ne doivent pas sembler ironiques ou dépréciatifs) ;
- être compatibles pour un même stage : deux arguments d'un même stade doivent être équivalents sans être contradictoires au point de ne pas pouvoir être acceptés en même temps.

2.3. Procédure de validation de l'outil

Une fois les deux dilemmes et les vingt-quatre arguments construits, l'outil a été soumis à un Professeur reconnu au niveau académique pour son expertise dans le domaine du développement moral et la théorie Kohlbergienne, afin de vérifier la représentativité conceptuelle des arguments conçus. Dans un second temps, l'outil a été soumis à un prétest auprès de 5 étudiants afin de détecter les problèmes de compréhension des dilemmes. Les sujets ont répondu au test en situation réelle avec la consigne supplémentaire de souligner les termes ou formulations qui semblaient ambigus, ou peu clairs. Un entretien semi-directif a également été mené afin de permettre aux étudiants d'exprimer leurs remarques. Ces multiples regards ont permis d'adapter et de clarifier certaines formulations dans les dilemmes et les arguments conçus. La version adaptée de l'outil a ensuite été administrée à 44 sujets.

La validation de cet outil ne peut se faire selon les mêmes critères que ceux en vigueur pour les questionnaires classiques, tels que les critères de consistance interne. En effet, la consistance des réponses des sujets est justement l'élément que nous voulons mesurer par ce test. Nous avons suivi les critères de validité énoncés par Lind (ibidem) pour valider notre outil. Lind (ibidem) distingue 3 critères auxquels une version adaptée du MJT devrait répondre.

Le premier de ces critères concerne les préférences morales des sujets. Celles-ci, devraient être suivre un ordre de préférence prévisible : les arguments de stade six devraient être préférés aux arguments de stade 5, eux-mêmes préférés aux arguments de stade 4 et ainsi de suite. Selon Lind, cette préférence hiérarchique se retrouve dans tous les tests étudiant le raisonnement moral, et ce, dans des cultures très différentes (Lind, 2008, p. 203). Cela n'empêche pas que certaines petites inversions se retrouvent, par exemple entre les stades 1 et 2 ou entre les stades 5 et 6. Ces petites inversions n'invalident pas la version étudiée, elles sont même fréquentes.

Le deuxième critère de validité énoncé par Lind (ibidem) est que les données doivent suivre une structure quasi-simplex, c'est-à-dire que les stades voisins (5 et 6, par exemple) doivent être davantage corrélés que les stades plus éloignés (stades 3 et 6 par exemple).

Le troisième critère concerne le parallélisme affectif-cognitif : les préférences morales pour les stades élevés doivent être clairement corrélées positivement avec le score de l'index-c, alors que les préférences morales pour les stades moraux les plus bas doivent être clairement corrélées négativement avec l'index-c.

Les résultats de l'analyse menée sur l'échantillon (n=44) montrent que le test de jugement moral professionnel que nous avons conçu répond au premier critère de validation. Les arguments du stade 5 sont ceux qui s'avèrent les plus préférés, suivis du stade 6, 4, 3, 2, et 1. Une inversion existe entre les stades 5 et 6, mais celle-ci n'est pas suffisante pour invalider le test et se retrouve fréquemment dans le MJT (Lind, 2008)

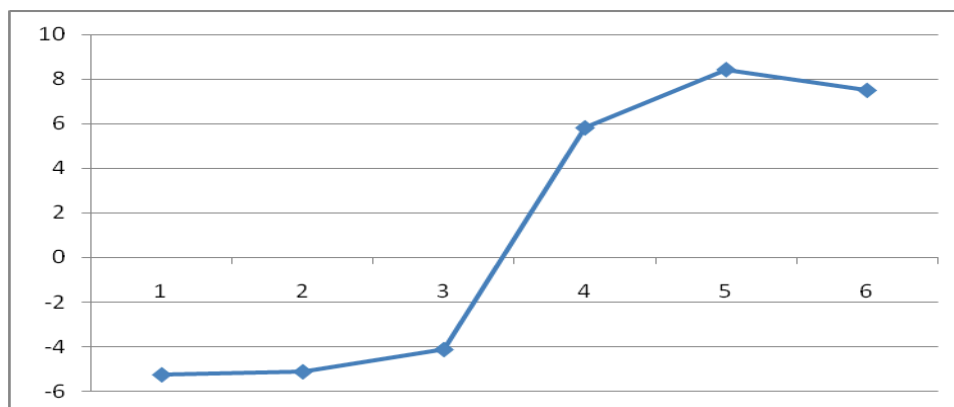


Figure 2 : somme des préférences des sujets pour les arguments des 6 stades moraux

Les résultats montrent également que le deuxième critère de validité est rencontré. Les données récoltées montrent que les stades voisins obtiennent des scores de corrélation plus élevés que les stades éloignés.

Le troisième critère de validation, celui du parallélisme affectif-cognitif est lui aussi respecté : les stades élevés sont clairement corrélés positivement au score C, les stades faibles sont clairement corrélés négativement avec celui-ci.

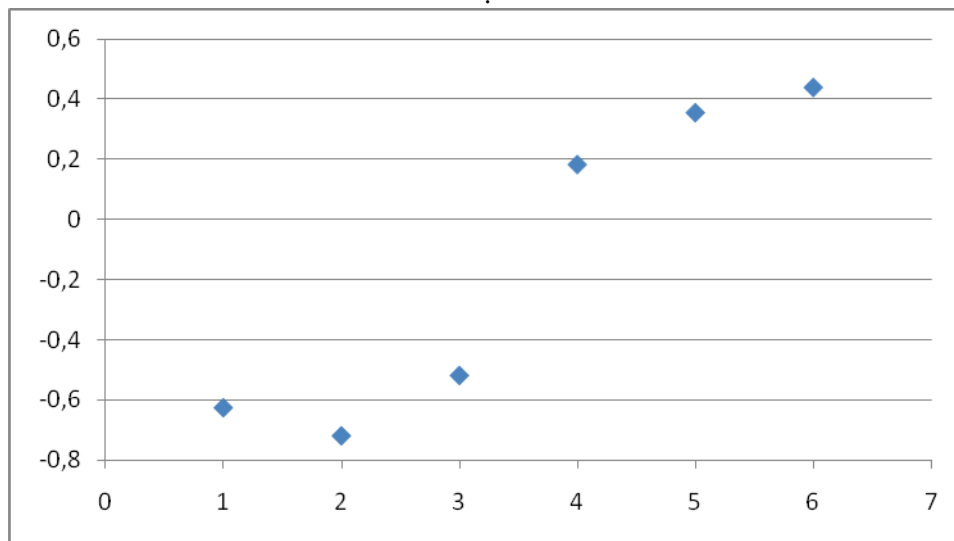


Figure 3 : corrélations entre les préférences des sujets pour les 6 stades et leur score C.

Premiers résultat et discussion

Outre les tests de validation de la procédure mise au point par Lind, nous avons également voulu comparer les données récoltées par le MJT « enseignant » développé dans cette étude et le MJT original de Lind. Les résultats de cette comparaison se révèlent particulièrement intéressants. En effet, les scores C obtenus par les sujets sont significativement plus élevés pour le dilemme enseignant (moyenne de 47,57) que pour le dilemme original du MJT (moyenne de 30,77).

Plusieurs raisons peuvent expliquer ce score plus élevé pour l'outil créé, mais il semble légitime d'interpréter celui-ci comme un indicateur d'une compétence morale professionnelle. Ces résultats tendraient à confirmer l'hypothèse que la compétence morale n'est pas une aptitude générale applicable dans toutes les situations et tous les contextes, mais qu'il s'agit bien d'une compétence dont le degré de performance dépend du contexte dans lequel elle est mise en œuvre. Le contexte professionnel enseignant dans lequel se passe le dilemme créé pour cette version du MJT est familier aux futurs enseignants qui ont répondu au test. Ces étudiants ont été amenés à réfléchir à des dilemmes enseignants durant leur formation. Il était donc prévisible que leurs résultats soient meilleurs pour un MJT enseignant que pour un MJT standard.

Des analyses comparatives plus poussées entre les résultats du MJT original et les résultats de cet outil seront réalisées, une fois que nous disposerons d'un plus grand échantillon ayant utilisé les

deux tests, afin d'approfondir la question de la contextualisation de la compétence en jugement éthique.

Perspectives et limites

Les premiers résultats obtenus tendent donc à confirmer la validité de l'outil créé et justifient sa pertinence dans le cadre de recherches autour de la compétence éthique des enseignants. Ce test semble en effet plus indiqué que le MJT original de Lind (2008) pour mesurer l'évolution de la compétence en jugement éthique des futurs enseignants.

Le MJT enseignant décrit dans cette communication devrait aider les chercheurs à appréhender la compétence en jugement éthique des futurs enseignants. Il est toutefois utile de souligner les limites de ce test. En effet celui-ci n'englobe pas toutes les dimensions de l'éthique enseignante, le cadre théorique choisi propose une interprétation forcément limitée de l'éthique. Il est également nécessaire de préciser que cet outil n'a pas pour vocation d'être utilisé comme méthode d'évaluation individuelle dans un processus d'examen ou de sélection. Ce type d'usage aurait notamment le défaut d'inciter les candidats à s'entraîner pour le réussir en apprenant son fonctionnement interne, ce qui fausserait complètement sa validité.

L'outil décrit dans cette communication ouvre également de nouvelles perspectives de recherche ; il sera en effet utilisé dans une étude en cours qui tente de comprendre les conditions de l'agir éthique dans la formation des futurs enseignants. Il permettra de mesurer la compétence éthique des étudiants en formation initiale et de mesurer son évolution.

Bibliographie

- Desaulniers, M. P. & Jutras, F. (2006) *L'éthique professionnelle en enseignement. Fondements et pratiques*. Québec : Presse de l'Université du Québec.
- Gohier, C., (1997) Éthique et déontologie : l'acte éducatif et la formation des maîtres professionnellement interpellés, dans M.P. Desaulniers, F. Jutras, P. Leblais et G.A. Legault (dir.), *Les défis éthiques de l'éducation*, 191-205. Québec : Presses de l'Université du Québec.
- Johnson, L. E. (2008) Teacher candidate disposition : moral judgement or regurgitation ? dans *Journal of moral Education*, 37 (4), 429-444
- Kohlberg, L. (1981) Essays on moral development (Vol. 1): *The philosophy of moral development: moral stages and the idea of justice*. San Francisco, CA : Harper and Row.
- Lind, G. (2008) The meaning and measurement of moral judgment competence revisited - A dual-aspect model, dans D. Fasko & W. Willis, Eds., *Contemporary Philosophical and Psychological Perspectives on Moral Development and Education*, 185 - 220. Cresskill, NJ : Hampton Press.
- Moreau, D. (2009) La question de l'éthique professionnelle des enseignants : un enjeu essentiel de la formation professionnelle et universitaire des maîtres, dans Brassart, G. et Legrand, D. (eds), *Qu'est-ce qu'une formation professionnelle universitaire des enseignants ?* Villeneuve d'Ascq : IUFM Nord-Pas de Calais

EVALUATION DE LETTRES COMMERCIALES A VISEE PERSUASIVE (LE CAS DES ETUDIANTS DE L'ECOLE NATIONALE DE COMMERCE ET DE GESTION D'AGADIR / MAROC).

Amina Saoussany, Université Ibn Zohr, Ecole Nationale de Commerce et de Gestion d'Agadir, MAROC, aminasaoussany@hotmail.com

Mots-clés : évaluation, compétence argumentative, lettres commerciales, communication spécialisée, persuasion, analyse textuelle.

Résumé. Quelles sont les principales difficultés en argumentation écrite rencontrées par les étudiants de l'Ecole Nationale de Commerce et de Gestion d'Agadir lors de leur production de lettres commerciales à visée persuasive et comment réussir leur compétence argumentative en français sur objectifs spécifiques (FOS) dans cette institution? Telle est la question centrale de cette étude. A partir de deux méthodes différentes de recueil de données: le questionnaire et les productions écrites effectives des étudiants, l'étude a abordé les questions fondamentales relatives à l'évaluation de la communication spécialisée en milieu institutionnel. Le cadre théorique porte essentiellement sur les concepts de base de la linguistique textuelle (Adam, 2005, 1999, 1990, Bronckart, 1996...) et expose les caractéristiques du FOS avec une attention particulière pour la communication épistolaire (Grassi, 1998). Il développe également les différentes théories de l'argumentation qui servent de base à la détermination des critères entrant dans la construction de notre grille d'évaluation (Anscombe & Ducrot, 1980, Grize, 1996, Amossy, 2000...). Les conclusions tirées de l'analyse des données suggèrent des réajustements pédagogiques pour améliorer l'enseignement de la communication spécialisée à l'ENCG d'Agadir.

Introduction

L'article que nous présentons se propose d'analyser les difficultés en argumentation écrite d'étudiants marocains en formation économique lors de leur production des lettres commerciales à visée persuasive. Cette étude est née de nos observations quotidiennes en tant qu'enseignante des Techniques d'Expression et de Communication à l'Ecole Nationale de Commerce et de Gestion d'Agadir et motive notre volonté de contribuer à l'amélioration de la formation en français sur objectifs spécifiques des étudiants qui envisagent soit la carrière de commercial, soit celle de gestionnaire au sein de l'entreprise.

L'enseignement d'une langue de spécialité en milieu institutionnel vise à développer chez les étudiants la maîtrise de cette langue à des fins communicatives orales et écrites liées au domaine professionnel. Sur le plan de l'écrit, entrer dans une communication spécialisée, c'est mettre en œuvre des compétences langagières bien déterminées et des savoirs linguistiques particuliers, relatifs au monde de la profession. Cependant, lorsque les apprenants sont des bacheliers arabisés n'ayant pas une solide maîtrise du français - comme c'est le cas avec les étudiants marocains - la pratique de cette langue devient un exercice souvent difficile. Dans une telle situation, le rôle de l'enseignant consiste à proposer des activités pédagogiques appropriées capables d'aider les apprenants à s'approprier les savoirs et savoir-faire nécessaires à leur réussite.

Produire des lettres commerciales à dominante argumentative, c'est savoir mobiliser un ensemble de compétences à la fois linguistiques, pragmatiques et textuelles. Il s'agit, d'une part, de répondre aux exigences d'une bonne formation textuelle et, d'autre part, de parvenir à toucher l'énonciataire consommateur potentiel en mettant en œuvre des stratégies à visée persuasive. Quelles sont les difficultés rencontrées par les étudiants en formation économique à l'ENCG d'Agadir lors de leur production des lettres de prospection de clientèle ? Comment évaluer leurs productions ? C'est à

ces questions générales que nous avons tenté d'apporter des éléments de réponses dans cette étude. A l'ENCG, nous avons en effet observé que, malgré les dispositifs mis en place et les efforts des enseignants pour développer la compétence scripturale des apprenants, nombre d'entre eux éprouvent des difficultés à produire des lettres de prospection de clientèle cohérentes et structurées sur le plan argumentatif. D'où le choix de ce thème de recherche qui touche de près à notre activité et à notre responsabilité d'enseignante.

1. Les fondements théoriques de la recherche

Sur quels indicateurs appuyer notre évaluation de lettres commerciales à visée persuasive produites par nos étudiants ? Le recours à la théorie nous a permis de repérer les traits linguistiques et discursifs qui font qu'un texte est argumentatif et persuasif et d'identifier les composantes qui entrent dans ce genre de texte. Nous avons cherché dans la théorie les outils nous permettant de décrire ce qui serait un texte modèle « un modèle prototypique », pour ensuite mieux relever les manques et les réussites des textes des étudiants. Le champ scientifique général est celui de la linguistique textuelle, discipline de base de l'analyse des textes. Nous avons fait référence aux auteurs du domaine tels que Bronckart (1996, 1997), Adam (1990, 1996, 1997), Maingueneau (1996, 1998), en leur empruntant différents concepts comme celui de genre, de séquences séquentielles et autres. Mais, c'est la nature de notre corpus qui a guidé notre travail de recherche. Les textes à évaluer sont de type argumentatif et relèvent de la correspondance commerciale. C'est sur ces deux aspects que nous allons insister en les distinguant pour la clarté de notre présentation, bien qu'ils soient interreliés.

1.1. L'argumentation

L'argumentation est la dominante de notre travail et nous avons consacré beaucoup de lectures et de temps à cette notion très complexe. Les apports des différentes théories de l'argumentation, celle d'Anscombe & Ducrot (1981, 1983), de Grize (1996) et d'Amossy (1999, 2000) nous ont été utiles pour caractériser ce qu'est l'argumentation et pour déterminer ses fondements logiques, théoriques et pragmatiques. L'analyse argumentative des textes s'inscrit dans différents courants de la pragmatique, fait intervenir la rhétorique ou l'art de persuader qui est revenu sur le devant de la scène avec les travaux de Perelman et Olbrechts-Tyteca (1970). Ces auteurs qui ont fait fructifier l'héritage de la rhétorique antique définissent la théorie de l'argumentation comme « *l'étude des techniques discursives permettant de provoquer ou d'accroître l'adhésion des esprits aux thèses que l'on présente à leur assentiment* » (1970 : 5). Cette conception nous a paru bien adaptée à la nature de notre corpus. L'argumentation développée dans les textes de nos étudiants doit avoir une visée persuasive, ce qui veut dire que les textes doivent avoir une force illocutoire afin de produire un effet sur le destinataire. Nous sommes renvoyée ici à la théorie des actes de langage mais en sachant que la force illocutoire d'un discours peut se réaliser à travers plusieurs procédés différents. Ce sont ces procédés que nous avons voulu repérer. Les deux domaines que nous avons repérés dans ce que nous avons appelé la logique argumentative sont les modes de raisonnement comme la déduction et l'explication qui enchainent des arguments à valeur de preuve. Ce sont aussi les connecteurs, en particulier les connecteurs argumentatifs.

1.2. Les procédés de l'analyse argumentative

La lettre commerciale, devant présenter un nouveau produit et vanter ses qualités, il a été nécessaire d'identifier les procédés par lesquels un objet doté d'une valeur peut persuader et pousser le destinataire à désirer l'acheter. Les travaux de Catherine Kerbrat-Orrechioni (1986) sur la subjectivité dans le langage nous ont mis sur la piste des adjectifs axiologiques, exemple *super*, *efficace*, *moderne*, qui présente le produit dans une stratégie de valorisation. Le deuxième procédé concerne le texte en tant que lettre commerciale qui s'inscrit dans un dispositif énonciatif explicite : un scripteur qui établit un contact avec un destinataire. Ce dispositif obéit à des normes sociales liées aux rôles et aux statuts des deux protagonistes : c'est un responsable de la prospection et du marketing qui s'adresse à un client potentiel qu'il ne connaît pas.

Dans l'analyse du dispositif énonciatif, nous avons isolé deux éléments. Le premier est la construction de l'image de soi par le scripteur que nous avons traité à travers la notion d'éthos évoquée par Amossy. Le deuxième est le contact établi avec le destinataire qui vise à réduire la distance avec l'autre et à renforcer l'impact de son discours. Un destinataire que l'on implique sera plus réceptif au discours d'action ou plutôt au discours d'influence qu'on lui adresse. Le texte a une fonction pragmatique qui passe par des moyens d'action sur l'autre avec l'exemple de l'usage des questions rhétoriques et de l'impératif qui incitent l'autre à faire quelque chose, ici à acheter le produit.

Nous avons accordé de l'importance à l'image du texte. Notre objectif était de vérifier s'il y avait appropriation des conventions de l'écrit épistolaire professionnel par les étudiants. Cet aspect peut être moins relié à l'argumentation, mais le découpage en paragraphes peut jouer un rôle dans la mise en valeur des arguments. De même, les mots en gras, en italique, l'usage des majuscules, ces procédés que Jean Peytard (1982) nomme entailles scripto-visuelles, sont des moyens de capter l'attention du lecteur destinataire. De façon générale, nous pouvons dire que tout le texte est traversé par une visée argumentative qui assure son homogénéité.

La relation avec le destinataire qu'on veut influencer impose cependant des limites. Ainsi, contrairement à la publicité qui se base sur l'argumentation impressive où la séduction et l'affectivité remplacent le raisonnement - on veut persuader le destinataire du message sans se soucier de la vérité -, les lettres de prospection de clientèle doivent utiliser l'argumentation de façon mesurée pour rester crédibles aux yeux du destinataire. Les lettres doivent contrôler l'argumentation afin de ne pas donner l'impression au destinataire qu'on est trop insistant. Elles doivent aussi valoriser le produit avec modération. L'usage de l'hyperbole est exclu, l'excès de valorisation est malvenu. La description de l'objet doit aussi satisfaire à la condition de vérité. Les différentes dimensions de la persuasion et de l'argumentation qui ne sont pas faciles à articuler montrent que la lettre de prospection de clientèle est un acte complexe qui implique des compétences de nature diverses que les étudiants que nous formons doivent acquérir et qu'ils doivent manifester dans la tâche qui leur a été demandée : produire une lettre de prospection de clientèle en se projetant dans un rôle professionnel futur.

1.3. La grille d'évaluation

Les écrits théoriques sur l'argumentation nous ont conduites à considérer quatre phases principales en interrelation constante qui sont à la base de la construction de la grille d'évaluation des lettres commerciales à visée persuasive qui constituent notre corpus.

Première phase : Etude de l'image du texte. Ce niveau nous permettra d'évaluer :

- la présentation de la lettre (aire scripturale normative de la lettre - zone paratextuelle et zone textuelle -, graphie) ;
- la structure typographique des paragraphes et leur structure textuelle.

L'objectif est de voir dans quelle mesure les lettres produites par les étudiants respectent les normes conventionnelles de présentation des écrits épistolaires professionnels.

Deuxième phase : Gestion du discours argumentatif. Nous analyserons la structure compositionnelle de la lettre (ouverture, exorde, corps de la lettre, péroraison et clôture) pour rendre compte du schéma textuel de la lettre et voir dans quelle mesure les écrits épistolaires produits par nos étudiants respectent le plan prototypique du genre épistolaire.

Nous nous intéresserons également à l'étude du mouvement argumentatif avec l'analyse de l'emploi des connecteurs argumentatifs intraphrastiques et interphrastiques, l'enchaînement des arguments et leur structuration dans les lettres produites par nos étudiants.

La relation énonciateur/énonciataire sera analysée afin de voir comment l'énonciataire est interpellé et impliqué dans le discours. L'usage des impératifs et des questions comme procédés pragmatiques d'action sur l'autre est à prendre en considération dans ce sens.

Troisième phase : Analyse de l'objet à promouvoir pour voir comment il est présenté et décrit par l'énonciateur :

- L'objet présenté : nous étudierons la cohésion nominale, qui ayant la fonction d'introduire des thèmes, des personnages, des objets et d'assurer leur reprise anaphorique tout au long du texte, nous aidera à étudier le système de la référence à l'objet dans le discours et à analyser les phénomènes d'anaphore qui posent des problèmes ;
- L'objet décrit : nous effectuerons une analyse de la description objective et de la description subjective du produit. La description objective nécessite l'identification des principales qualités techniques du produit et l'apport de preuves en vue de persuader l'énonciataire de l'intérêt du produit décrit. La description subjective consiste à valoriser le produit et à le décrire positivement pour favoriser sa vente. L'usage des adjectifs non axiologiques dans le premier cas et des adjectifs axiologiques dans le second est à prendre en considération dans l'analyse de ces deux types de description.

Quatrième phase : Etude du dispositif d'énonciation. Il s'agira d'étudier les mécanismes de prise en charge des énoncés par l'énonciateur et la référence au destinataire dans le discours.

Les mécanismes de prise en charge énonciative assurent la cohérence pragmatique et interactive du discours. Ils permettent de savoir quel est le mode de prise en charge énonciative des instances qui prennent en charge ce qui est énoncé dans le texte et les évaluations (modalisations) formulées à l'égard du contenu thématique. Pour notre cas, il s'agira de voir comment la référence à l'énonciateur se fait dans la lettre et quels sont les indices d'allocution utilisés à cette fin.

L'inscription de l'énonciataire dans la lettre se fait à l'aide de plusieurs procédés que nous tâcherons d'analyser dans les lettres des étudiants (pronoms d'adresse, désignations nominales, termes d'adresse...).

Le tableau suivant explicite la grille ci-dessus :

Phase 1	Etude de l'image du texte.	- Présentation de la lettre ; - Structure typographique des paragraphes et leur structure textuelle.
Phase 2	Gestion du discours argumentatif.	- Structure compositionnelle de la lettre ; - Mouvement argumentatif ; - Relation émetteur/récepteur.
Phase 3	Analyse de l'objet à promouvoir.	- L'objet présenté : cohésion nominale. - L'objet décrit : - description objective ; - description subjective.
Phase 4	Etude du dispositif d'énonciation.	-Les mécanismes de prise en charge énonciative ; - La référence à l'énonciataire dans le discours.

Tableau 1: Grille d'évaluation du corpus

Après avoir présenté le cadre théorique de notre recherche et dressé la grille à partir de laquelle il sera possible d'évaluer notre corpus, nous présenterons dans la suite de ce travail le cadre d'expérimentation dans lequel nous avons recueilli les données qui constituent notre corpus ainsi que l'analyse de ce même corpus.

2. Méthodologie de recueil de données

2.1. Sélection du public d'enquête

Le public sur lequel est axée notre étude comprend des étudiants de la deuxième année de l'ENCG d'Agadir (promotion 2007 - 2008). Ces étudiants marocains, venus de différentes régions du pays, doivent passer cinq ans à l'ENCG dans le cadre de leur formation en commerce et / ou gestion. Ils sont destinés à travailler en entreprise en tant que commerciaux, gestionnaires, chargés d'audit, publicitaires, chefs de produit...

A leur admission à l'ENCG en septembre 2007 après la réussite au TAFEM¹, les étudiants étaient au nombre de 171. A leur passage en deuxième année, ils étaient 150 répartis dans 6 groupes (groupes 1 - 6), chaque groupe étant constitué approximativement de 25 étudiants.

La répartition en groupe se fait en prenant en considération les caractéristiques des apprenants (âge, série de bac, origine géographique, sexe). Chaque groupe est représentatif de l'ensemble. Il est constitué d'étudiants originaires d'Agadir, mais aussi des différentes régions du Maroc. Ceci pour faciliter l'intégration entre eux et éviter le cloisonnement des apprenants originaires de la même région. Toutes les séries de bac sont également représentées dans chaque groupe ; ce qui fait que l'on trouve les sciences expérimentales avec les sciences économiques, les techniques de gestion comptable et les techniques de gestion administrative dans la même classe. Pour ce qui est de l'âge, les étudiants sont tous des jeunes âgés de 18 à 21 ans. Ils ont tous le même nombre d'années d'enseignement du français (11 ans) et ils ont étudié les mêmes programmes en français au fondamental et au secondaire.

Le cours de Techniques d'Expression et de Communication en français est assuré, en deuxième année, par trois professeurs. Pour notre part, nous avons dispensé ce cours dans deux groupes (les groupes 3 et 4) et nous avons eu en totalité 48 étudiants. Sur ces quarante huit étudiants, seuls 28 ont participé à notre expérimentation.

2.2. Méthodologie de l'enquête

Pour atteindre les objectifs de notre étude, nous avons adopté une approche à la fois qualitative et quantitative. Le recueil de données a été effectué principalement à travers la passation d'un questionnaire aux étudiants et la réalisation d'une activité d'expression écrite par les apprenants à partir d'une consigne. Cette activité concerne la production d'une lettre commerciale à dominante argumentative.

Le questionnaire sur lequel est basée l'enquête comprend cinq parties et totalise 38 questions. Il vise à diagnostiquer les axes suivants :

- Identification des enquêtés ;
- Le répertoire linguistique ;
- La relation à la langue française ;
- Les usages du français à l'écrit ;
- Les difficultés en français en expression écrite.

¹ Le concours d'accès à l'ENCG est appelé Test d'Admissibilité à la Formation En Management (TAFEM). Il est composé d'un test écrit et d'un entretien oral.

Ces informations ont été demandées en fonction des hypothèses suivantes :

Hypothèse 1 : les attitudes des étudiants marocains vis-à-vis de la langue française ont une influence sur leur compétence langagière et par conséquent sur leur pratique de l'écrit argumenté. En partant des études sur les représentations du français et sur la compétence langagière des apprenants marocains faites par Benzakour (2000) et Boukous (1995), nous avançons l'hypothèse que les étudiants qui considèrent que le français est une langue importante pour leurs études ont une motivation instrumentale forte pour apprendre cette langue et donc moins de difficultés dans la production des textes en français.

- *Hypothèse 2* : en considérant le fait que les types d'usage faits d'une langue non maternelle peut avoir une incidence sur le niveau de l'écrit de cette langue, nous émettons l'hypothèse que les sujets qui font des usages spécialisés du français écrit hors de l'institution de formation ont moins de difficultés à rédiger des textes argumentatifs de type professionnel.

Pour vérifier ces hypothèses, nous avons recueilli des productions écrites effectives de ces 28 étudiants pour évaluer leur compétence argumentative et voir les effets de l'enseignement que nous dispensons. Nous avons demandé aux étudiants de produire une lettre commerciale argumentative à propos d'un produit précis : le téléphone mobile Sumsung SGH-E330. La consigne de travail a été formulée comme suit :

L'entreprise MAROCTELECOM (SA au capital de 900000000 dhs, R.C. C7653, Code APE : 8763, Tel : 02387562, Fax : 02345680) vient de lancer sur le marché marocain sa nouvelle création dans le domaine des télécommunications : le mobile Sumsung SGH-E330. En tant que responsable chargé de la prospection et du marketing au sein de cette société, rédigez la lettre de prospection de clientèle que vous allez envoyer à votre client pour lui présenter les qualités de votre nouveau produit et le persuader d'effectuer son achat.

Essayez de le convaincre en l'impliquant et en lui apportant les arguments de vente les plus ciblés.

3. Les résultats de l'enquête

L'évaluation des lettres commerciales à visée persuasive du corpus a montré que les difficultés des étudiants en argumentation écrite en communication de spécialité sont d'ordre textuel et discursif. Les problèmes locaux touchant à des micro-éléments de la phrase (morphologie, orthographe d'usage, conjugaison...) ne sont pas abordés, non que ceux-ci soient mineurs, mais ce sont des problèmes auxquels nous n'avons pas accordé attention eu égard aux objectifs de recherche que nous nous sommes fixée.

Nous présentons les résultats de manière résumée. Nous passons rapidement sur l'image du texte qui ne présente pas de difficultés majeures sauf pour la structuration en paragraphes qui n'est pas toujours bien maîtrisée. En ce qui concerne l'objet à promouvoir, les stratégies de valorisation par la mise en valeur des qualités techniques de l'objet et par l'emploi d'adjectifs positifs axiologiques sont généralement bien employées. Le problème se situe plutôt au niveau des enchaînements discursifs des arguments relatifs à l'objet qui devraient être mieux mis en ordre, mieux organisés pour être plus convaincants et persuasifs.

Les mécanismes de cohésion nominale à propos de l'objet sont utilisés de manière satisfaisante même s'il y a des procédés répétitifs. Quant au discours argumentatif proprement dit, c'est principalement l'usage des connecteurs argumentatifs qui posent des problèmes aux étudiants et en particulier des connecteurs interphrastiques tels que : *ainsi, or...*

Enfin, la plupart des étudiants sont conscients de l'importance de la relation énonciative et mènent bien les indices d'allocution. Mais, l'impératif et les questions rhétoriques comme moyens

d'action sur l'autre ne sont pas bien utilisés. Ceci affaiblit, selon nous, le caractère interactionnel et argumentatif de la lettre.

Sur le plan général de la lettre, l'appropriation du schéma prototypique épistolaire présente quelques défaillances. Il manque parfois la séquence d'ouverture et de clôture, voire une phase importante comme la péroraison. Mais, les textes ont été réalisés en situation de simulation et non pas en situation réelle. On ne sait pas si les étudiants, une fois en situation professionnelle, feraient de tels oublis.

Les résultats de l'analyse confirment nos hypothèses : un rapport positif à la langue française va de pair avec une motivation instrumentale forte pour apprendre le français. Les étudiants, dans ce cas, s'impliquent davantage dans leurs études et ont de meilleurs résultats que les autres. D'autre part, les difficultés relevées se manifestent moins chez les étudiants qui font des usages spécialisés du français écrit hors de l'école. Ces usages ont renforcé leur compétence argumentative et cela s'observe dans leurs productions.

4. Conclusion et recommandations

Comme apport principal, cette étude nous a permis d'identifier les principales difficultés des apprenants marocains en formation économique au supérieur dans un exercice particulier : l'argumentation dans une lettre de prospection de clientèle. Certains problèmes relèvent de l'enseignement du français (usage des connecteurs argumentatifs, enchaînement des arguments), tandis que d'autres touchent le domaine de la formation en Techniques d'Expression et de Communication (implication du destinataire dans le discours, inscription de l'affectivité dans la lettre, appropriation de la structure compositionnelle de la lettre). Les réajustements pédagogiques que l'on pourra proposer pour améliorer l'enseignement de la communication de spécialité au sein de l'ENCG doivent tenir compte des problèmes relevés.

Deux grands axes de suggestions pédagogiques peuvent être dégagés. Le premier met l'accent sur les dispositifs à mettre en œuvre et à renforcer par les enseignants au profit des étudiants et le deuxième concerne directement les enseignants, leur formation continue et le développement de leurs pratiques d'enseignement.

D'après les résultats de l'analyse, le degré d'investissement de nos étudiants dans l'écriture spécialisée ne se révèle pas important. Alors que de façon globale, ils affichent des attitudes positives à l'égard de l'écrit en français, ils ne s'adonnent pas beaucoup à des activités écrites en français de spécialité à l'Ecole et en dehors des cours. En vue de développer leur compétence langagière en communication spécialisée en entreprise, il serait utile d'enrichir le contenu du programme des Techniques d'Expression et de Communication en deuxième année par d'autres activités en français de spécialité: production d'articles économiques, de dissertations à partir de sujets économiques d'actualité, de textes publicitaires, de lettres commerciales à dominante argumentative à partir d'un spot publicitaire diffusé à la télévision ou d'une brochure qui décrit les qualités du produit, etc.

Pour développer la compétence argumentative des apprenants dans la formation donnée à l'Ecole (cours de communication), il faudrait travailler davantage sur des extraits de magazines spécialisés, de journaux économiques et de revues de commerce. L'objectif serait de faire acquérir des savoirs et savoir-faire en communication spécialisée qui pourraient être utilisés dans l'exercice de la profession.

Au sein des classes des Techniques d'Expression et de Communication à l'ENCG et pour contribuer à résoudre les problèmes de langue rencontrés par les apprenants en formation économique, nous jugeons utile de faire travailler les étudiants sur l'analyse des textes en vue d'étudier leur structure et d'explicitier le fonctionnement des connecteurs, leur rôle au sein du

discours et l'enchaînement des arguments qui déterminent la force argumentative globale de la lettre.

En outre, et en vue de développer la compétence linguistique et langagière des étudiants marocains arabisés, nous estimons essentiel de proposer aux nouveaux bacheliers admis à l'ENCG une formation de base en français en première année pour améliorer leur niveau en langue, corriger leurs lacunes et contribuer, par conséquent, à renforcer leur compétence textuelle en français de spécialité.

Nous estimons également essentiel de former les étudiants de l'ENCG qui sont en majorité, titulaires d'un doctorat en littérature ou en linguistique en français sur objectifs spécifiques qui est une orientation pas reconnue pour eux pour leur permettre de s'ouvrir à une nouvelle voie de formation indispensable à la formation de nos étudiants. Les séminaires de formation et les stages organisés par l'Institut Français d'Agadir seraient d'un apport intéressant au projet de formation de nos enseignants.

Références et Bibliographie

- Adam, J.-M. (1997). *L'Argumentation publicitaire. Rhétorique de l'éloge et de la persuasion*. Paris : Nathan Université.
- Adam, J.-M. (1996). *Les textes, types et prototypes : récit, description, argumentation, explication, dialogue*. Paris : Nathan.
- Adam, J.-M. (1990). *Éléments de Linguistique textuelle : théorie et pratique de l'analyse textuelle*. Liège : Mardaga.
- Adam, J.-M. (1985). Quels types de textes. *Le français dans le monde*, 192(39-43).
- Amossy, R. (2000). *L'argumentation dans le discours. Discours politique, Littérature d'idées, Fiction*. Paris : Nathan.
- Amossy, R. (1999). *Images de soi dans le discours. La construction de l'ethos*. Genève : Delachaux et Niestlé.
- Anscombre, J.-C. (1995). *Théories des topoi*. Paris : Editions Kimé.
- Anscombre, J.-C. & Ducrot, O. (1983). *L'argumentation dans la langue*. Liège : Mardaga.
- Anscombre, J.-C. & Ducrot, O. (1981). Interrogation et argumentation. *Langue Française*, 52(5-21).
- Benzakour, F. (2000). *Le français au Maroc, lexique et contacts de langues*. Bruxelles : Éditions Duculot.
- Boukous, A. (1995). *Société, Langues et Cultures au Maroc. Enjeux symboliques*. Casablanca : Imprimerie Annajah Al-Jadida.
- Brassart, D.G. (1992). *Enseigner/apprendre le texte argumentatif*. CRDP de Lille.
- Breton, P. (1996). *L'argumentation dans la communication*. Paris : Editions La Découverte.
- Bronckart, J.-P. (1997). *Activité langagière, texte et discours: Pour un interactionnisme socio discursif*. Paris : Delachaux et Niestlé.
- Bronckart, J.-P. (1996). Genres de textes, types de discours et opérations psycholinguistiques. *Enjeux*, 37/38(31- 47).
- Bronckart, J.-P. (1994). Action, langage et discours. Les fondements d'une psychologie du langage. *Bulletin suisse de linguistique appliquée*, 59(7- 46).
- Charolles, M. (1995). Cohésion, cohérence et pertinence du discours. *Travaux de Linguistique*, 29(125-151).
- Charolles, M. (1976). Grammaire de texte-Théorie du discours-Narrativité. *Pratiques*, 11 /12(133-154).
- Ducrot, O. (1980). *Les échelles argumentatives*. Paris : Editions de Minuit.
- Eggs, E. (1994). *Grammaire du discours argumentatif*. Paris : Kimé.

- Everaert-Desmedt, N. (1984). *La communication publicitaire. Etude socio-pragmatique*. Louvain-La-Neuve : Cabay.
- Grize, J.-B. (1996). *Logique naturelle et communication*. Paris : PUF
- Jouve, M. (1990). *La communication publicitaire*. Paris : Seuil.
- Maingueneau, D. (1998). *Analyser les textes de communication*. Paris : Dunod.
- Maingueneau, D. (1996). *Les termes clés de l'analyse du discours*. Paris : Seuil.
- Perelman, C., Olbrechts-Tyteca. (1970). 1ère éd. (1958). *Traité de l'Argumentation. La nouvelle Rhétorique* : Editions de l'Université de Bruxelles.
- Peytard, J. (1982). *Littérature et classe de langue*. Paris : Hatier-Crédif.

L'EVALUATION DES COMPETENCES EN CONTEXTE COLLECTIF : POIDS DES RELATIONS SOCIO-AFFECTIVES SUR LES CONDUITES.

Lucie Mougenot

GEPECS équipe TEC, Université Paris Descartes. L_mougenot@yahoo.fr

Mots-clés : *compétences collectives, conduites motrices, sociométrie.*

Résumé. *Notre étude est orientée autour des compétences collectives et de leur évaluation en milieu scolaire. L'activité collective et donc les interactions nous ont incités à nous rapprocher de la sociométrie et des rapports que les relations affinitaires entretiennent avec les décisions des individus. A partir d'une observation différée nous avons recherché des corrélations pouvant exister entre les réseaux relationnels existants entre les élèves et les conduites observables au cours d'un jeu collectif avec interaction motrice. Nos résultats montrent que la composition des équipes tenant compte des réseaux affinitaires a des répercussions au niveau de la coopération entre les membres du groupe. De plus, les élèves les plus rejetés dans la classe sont aussi sujets à un type de jeu particulier.*

1. Introduction

Aider les élèves à devenir plus compétents et surtout pouvoir évaluer ce cheminement est une tâche des plus complexes pour les enseignants. Or une difficulté supplémentaire apparaît quand on s'aperçoit que si l'école a pour vocation de préparer à la vie sociale et professionnelle, 59% des salariés déclarent réaliser une partie de leur travail en groupe ou collectivement (selon une enquête du Changement Organisationnel et Information, Connaissance de l'emploi n°16, 2005). Les compétences collectives seraient essentielles dans la vie professionnelle mais ne semblent pas être au cœur des préoccupations scolaires. En effet, le socle commun de compétences définissant selon trois niveaux les connaissances, capacités et attitudes essentielles, ne fait référence au niveau social qu'à l'apprentissage des règles de vie en groupe, à la tenue d'un projet, mais sans que ne soit mentionné un lien direct avec l'acquisition de compétences collectives, c'est-à-dire en interaction avec les autres (décret du 11 juillet 2006). Les compétences sociales semblent détachées des compétences plus disciplinaires, et servent l'acquisition de savoirs individuels.

Nous nous sommes ainsi intéressés aux activités collectives présentes à l'école permettant de développer à la fois des compétences individuelles et collectives. Nous avons choisi de réaliser une expérience en Education Physique et Sportive (EPS), et ceci pour deux raisons. La première est que l'évaluation en EPS aux examens se déroule sous forme de contrôle en cours de formation la majorité du temps, ce qui amène une nécessaire centration des enseignants sur les compétences à atteindre définies dans les référentiels, et une transposition de ces niveaux en outils d'évaluation les plus fiables possibles. La seconde raison est relative à la spécificité de la discipline. Celle-ci réside dans le fait que les apprentissages se déroulent de façon collective, mais les évaluations sommatives sont individuelles comme dans les autres disciplines, pour répondre aux exigences institutionnelles de notation. Or les apprentissages sont issus de nombreuses interactions entre les individus dans un grand nombre d'activités physiques. Par conséquent, les actions des uns motivent et induisent celles des autres, demandant une constante adaptation. L'évaluation ne dépendrait donc pas seulement de compétences individuelles, mais d'une articulation entre les compétences des individus étant en interaction.

Cela pose alors problème quand on constate le foisonnement de questions touchant la notion même de compétence individuelle, scolaire, et de son évaluation. Caillot (1994) montre que

ce concept n'aurait pas fait l'objet d'une analyse suffisamment approfondie avant d'être introduit à l'école. Dans certains programmes, le concept de compétence se substitue tout simplement à celui d'objectif. Dès lors, comment évaluer les élèves quand on cherche à atteindre des compétences en activité collective, quand les conduites des uns dépendent de celles des autres ? La compétence collective serait issue de « la coopération et de la synergie existantes entre les compétences individuelles » (Bataille 1999). Plus qu'une simple addition des compétences individuelles, elle nécessite coopération, communication entre les membres du groupe. En EPS plus spécifiquement, quand les élèves sont impliqués dans des jeux collectifs nécessitant de coopérer et de s'opposer, la communication est la clé de la réussite de l'équipe (Parlebas, 1992). La notion de compétence collective a toute sa place dans ce type de situations. Dans ce contexte collectif, il nous est apparu intéressant de mesurer l'impact que peuvent avoir les relations socio-affectives entre élèves sur leurs conduites. Par exemple, la composition de l'équipe en elle-même, et de l'équipe adverse n'est pas neutre. En EPS cette gestion de groupe se fait selon les niveaux d'habileté essentiellement pour aboutir à un rapport de force équilibré et indispensable au jeu. La mixité, le nombre de joueurs, les rôles particuliers font aussi partie des critères de composition d'équipes, mais finalement, les affinités sont assez peu prises en compte. D'ailleurs, certaines dérives peuvent apparaître comme la paresse sociale (Ringelmann, 1913), ou la constitution de rapports de force déséquilibrés, ou l'isolement de certains élèves quand on les laisse choisir leurs coéquipiers. Dans tous les cas, on s'aperçoit que selon les affinités entretenues entre les personnes, les décisions prises dans l'action peuvent diverger. Au cours d'un jeu collectif, c'est l'interaction entre les individus qui permet à chacun d'orienter, de décider des actions à entreprendre. Comment percevoir l'influence que peuvent avoir les uns sur la conduite des autres ? Les décisions prises dans les situations collectives à but opératoire sont-elles tributaires des relations socio-affectives ?

Deux hypothèses majeures organisent notre recherche. Tout d'abord nous pensons qu'il existe des corrélations entre certains profils socio-affectifs et les décisions prises au cours d'un jeu. Ensuite, selon la composition des équipes faisant référence à la gestion de groupe, les conduites et progrès des élèves vont être différents. C'est à partir du questionnaire sociométrique et d'un outil d'évaluation des conduites créé pour notre expérience que nous avons pu mettre avant des corrélations entre les données socio-affectives recueillies dans une classe et les conduites motrices observées lors d'un jeu collectif. Après avoir explicité la méthodologie employée, nous discuterons nos principaux résultats mettant en avant les différences significatives apparaissant notamment entre les conduites motrices des équipes composées d'amis et les autres. Les élèves les plus rejetés dans la classe seront aussi l'objet de toute notre attention au regard de leur propre conduite sur le terrain, mais aussi de celle des autres à leur égard.

2. Méthodologie

21. Contexte d'intervention

Notre recherche a été menée dans un lycée de région parisienne avec une classe de seconde générale en EPS. Les élèves ont suivi un cycle de basket-ball que nous avons construit en fonction des compétences issues des programmes et de leur niveau de jeu observé lors de la première séance. L'enseignant de la classe est l'enseignant habituel. Les élèves ont été filmés lors d'un jeu global à 4 contre 4 en évaluation diagnostique et sommative par le chercheur.

22. Les relations sociométriques

A partir du questionnaire sociométrique de Moreno (1954) nous avons mesuré les relations socio-affectives dans la classe (voir tableau 1 infra).

Phases	1	2	3	4	5
	Pré test	Séance 1	Séances 2 à 7	Séance 8	Post test

	Questionnaire sociométrique	Evaluation diagnostique	Cycle de 7 séances	Evaluation sommativ	Questionnaire sociométrique
--	--------------------------------	----------------------------	-----------------------	------------------------	--------------------------------

Tableau 1 : plan d'expérience

Le questionnaire rempli de façon personnelle et spontanée en classe fait référence à l'organisation d'une sortie ludique en groupe, sans but opératoire car il doit permettre de mesurer uniquement les relations affinitaires entre les membres du groupe. Nous avons retenu quatre questions spécifiquement pour cette recherche :

- Avec qui aimerais-tu passer cette journée ?
- Avec qui n'aimerais-tu pas passer cette journée ?
- Selon toi qui t'a choisi ?
- Selon toi qui ne souhaite pas être avec toi ?

Les élèves ont rempli le questionnaire en heure d'Accompagnement Personnalisé avec leur professeur principal. Il a été présenté comme un outil permettant de composer des équipes affinitaires en vue d'une sortie ludique pour ne pas que les élèves puissent établir de lien avec la composition des équipes en EPS. Les données recueillies sont traduites sous forme de tableau à double entrée où toutes les dyades d'élèves apparaissent selon le type de relation affinitaire ; par exemple deux élèves qui se choisissent mutuellement, ou alors un élève qui s'attend à être choisi mais ne l'est pas par l'autre... des scores individuels et collectifs sont ensuite calculés, ce qui nous permet de dégager certains profils d'élèves particuliers : les élèves populaires, les rejetés, les oubliés, ceux qui rejettent le plus... Nous construisons ensuite sous forme de graphe, des représentations de ces relations pour voir apparaître cette fois-ci les différents sous-groupes affinitaires qui se détachent dans la classe.

Les données recueillies nous ont permis de composer six équipes selon trois catégories différentes, à savoir deux équipes composées d'élèves qui s'apprécient mutuellement, deux équipes composées d'élèves se rejetant, et deux équipes dites neutres, c'est-à-dire composées d'élèves n'ayant pas d'avis particulier à donner sur leur partenaires de jeu (ni choix, ni rejet). Les équipes ainsi créées ont été utilisées la première séance d'évaluation diagnostique, la dernière séance d'évaluation sommative et quelques fois au cours du cycle pour habituer les élèves à jouer entre eux. Cependant, par moments, d'autres configurations d'équipes ont été mises en place par l'enseignant de la classe, à des fins pédagogiques et didactiques.

23. L'observation des conduites des élèves

L'objectif de notre recherche étant de mettre en relation les données sociométriques avec les conduites des élèves, nous avons construit un outil de recueil d'informations précis pour identifier l'ensemble des décisions et actions de chaque élève en fonction de la configuration du jeu.

Nous avons ainsi réalisé une observation différée en basket-ball. Le basket-ball étant une activité sociomotrice à but opératoire, les interactions motrices entre partenaires et adversaires exigent une adaptation constante des décisions et actions de chacun aux autres joueurs. Nous nous sommes inspirés des travaux de Parlebas (2002) pour déterminer à partir des rôles et sous-rôles toutes les actions possibles du joueur à chaque moment du match. Dugas (2006) a créé et utilisé un outil permettant de repérer toutes les actions possibles et de calculer des scores individuels, collectifs pour mesurer ensuite des progrès. Notre grille a été conçue à partir de ces références puis retravaillée en fonction des conduites typiques observables dans ce jeu (tableau 2). Chaque action est donc répertoriée selon si elle est favorable ou non à l'équipe du joueur. Par exemple un dribble ne sera favorable que si le joueur n'a pas d'autre choix que de tenter le duel pour avancer, c'est-à-dire s'il n'a pas de partenaire démarqué lui permettant d'avancer plus vite vers la cible. L'action et la prise de décision sont essentielles et indissociables. Chacune action identifiée se voit attribuer

des points, allant de - 3 à + 3, selon la conséquence que cette action a sur l'ensemble du jeu : par exemple une passe interceptée équivaudra à - 2 pour le passeur et + 2 pour le défenseur qui intercepte.

Le jeu de chaque élève a ainsi été filmé et analysé au terme des séances 1 et 8, lors de 2 rencontres de 8 minutes, avec les mêmes équipes dans la mesure du possible (la gestion de groupe a parfois été légèrement modifiée en fonction des absents ou des inaptes).

Statut du joueur	Actions favorables à l'équipe	Actions défavorables à l'équipe
Le joueur est porteur de balle	Tir réussi : + 3 Passe de conservation : + 2	Tir non favorable raté : - 3 Passe interceptée : - 2
Le joueur est attaquant non porteur de balle	Se démarque : + 1 Se démarque et reçoit : + 2	Placement non judicieux : - 2 Ne monte pas en attaque : - 3
Le joueur est défenseur	Intercepte une passe adverse : + 2 Dissuade la passe : + 1	Inaction sur tir adverse : - 2 Placement non judicieux : - 2

Tableau 2 : exemples d'actions de jeu et points attribués, en fonction du statut du joueur.

24. Analyse des données

Les données sociométriques ainsi que celles relatives aux conduites des élèves ont été quantifiées et catégorisées pour pouvoir être mises en relation ensuite par le test du khi². Au niveau sociométrique, dans le cadre de cette recherche, nous avons utilisé les résultats pour composer les équipes et pour nous centrer sur un profil typique d'élèves dont les conduites se différencient de celles des autres, à savoir les élèves les plus rejetés. Au niveau des conduites motrices nous avons mis en avant des équipes ou des joueurs qui se détachent des autres en fonction de :

- Leur score global (différence entre actions positives et négatives)
- Leur volume de jeu (nombre de balles jouées)
- Leur score en attaque et en défense
- Leur niveau de coopération avec et sans ballon
- Leur progrès

3. Résultats

31. données globales de progression.

L'évolution du jeu entre S1 et S8 nous montre un jeu plus rapide, avec davantage de prises de risque. Par exemple pour le même temps de jeu, nous constatons une augmentation du nombre de passes (+ 38%), du nombre d'interceptions (+ 24%) et moins de statisme, sans que ces données soient corrélées au profil de l'équipe. Elles concernent donc indifféremment l'ensemble des joueurs. Pour contrebalancer ces progrès, l'analyse des scores globaux individuels entre les séances 1 et 8 montre que pratiquement la moitié des élèves a régressé : la somme des actions défavorables soustraite à la somme des actions favorables est plus faible en séance 8 qu'en séance 1. Par ailleurs, il est primordial de noter que le nombre total d'actions pour un même temps de jeu a augmenté. La vitesse de jeu a ainsi évolué, amenant davantage de pression temporelle, complexifiant la prise de décision et rendant plus difficile aussi l'efficacité des actions. La

régression est alors compensée par un nombre d'actions plus important et une vitesse d'exécution augmentée.

32. Résultats selon la composition des équipes

Les données sociométriques utilisées pour mettre en place des équipes à profils affinitaires différents nous apportent certains résultats intéressants. En effet, après avoir calculé les scores globaux par équipe et les écarts entre séance 1 et séance 8, nous constatons que seules les équipes affinitaires ont progressé (+ 131 points), laissant les équipes neutres et celles d'élèves qui se rejettent assez loin derrière, avec respectivement une perte de 147 et 63 points.

Plus précisément maintenant, au regard de certaines catégories d'actions motrices, nous notons des différences très significatives concernant surtout le jeu des équipes affinitaires : les équipes d'amis se distinguent des autres ($p = 0,01$) dans leurs progrès en démarquage. Tous les élèves de ces deux équipes sans exception ont progressé dans le jeu sans ballon en attaque, contrairement aux autres équipes. La coopération en attaque, l'aide au porteur de balle est une stratégie défendue par tous les membres de ces groupes d'amis. Ces équipes sont aussi les seules à avoir progressé quant au nombre de rebonds offensifs¹ récupérés, alors même qu'elles ont eu un nombre de tirs inférieur aux autres équipes. Ceci s'inscrit également dans le sens d'une plus grande coopération visant l'optimisation de la conservation de la balle. Enfin, au niveau des tirs, les équipes affinitaires sont surreprésentées parmi celles qui tentent le moins de tirs en situation défavorable² ($p = 0,01$), contrairement aux équipes d'élèves qui ne s'apprécient pas.

33. Résultats individuels

D'abord, nous avons pu montrer une corrélation très significative entre le nombre de balles jouées des élèves et le nombre d'actions positives qu'ils font : les élèves qui touchent le plus la balle sont surreprésentés parmi ceux qui font le plus grand nombre d'actions positives ($p = 0,001$). Or, évaluer les niveaux de jeu ne peut se faire seulement par comptabilisation du nombre de balles jouées ; en effet, prendre possession d'une balle peut découler de différentes circonstances : intercepter, recevoir une passe, ou jouer une touche. L'implication et le niveau de jeu requis n'est pas le même. De plus, jouer une balle n'est pas forcément prendre la bonne décision, et certains joueurs touchent moins la balle que les autres mais ont une lecture du jeu et prise de décision efficaces. Dans ce sens, nous n'obtenons d'ailleurs pas de corrélation entre les élèves qui touchent le moins de balles et ceux qui font le plus d'actions défavorables.

Maintenant, attardons-nous sur des profils sociométriques particuliers : nous constatons tout d'abord que le nombre de balles jouées n'est pas corrélé au profil socio-affectif des joueurs. Par contre, les cinq élèves les plus rejetés dans la classe sont surreprésentés ($p = 0,01$) parmi ceux qui reçoivent le moins souvent la balle provenant d'une passe comparé à l'ensemble des autres joueurs en séance 1 et 8. Précisons que le nombre de balles jouées comprend les balles reçues par passe mais aussi les balles conquises (Gréhaigne, Guillon & Roche, 1991), ce qui explique que ces élèves rejetés reçoivent moins de passes mais ne touchent pas forcément moins la balle que les autres. De plus, nous remarquons que ceux-ci ne sont pas non plus les joueurs ayant les scores globaux les plus faibles, leur niveau de jeu n'étant ainsi pas la cause d'un nombre moins important de passes reçues. Les élèves rejetés se démarquent aussi moins bien que les autres pour recevoir la balle ($p = 0,01$). Ceci peut rejoindre d'ailleurs notre précédent résultat à savoir qu'ils reçoivent moins de passes que les autres. En revanche, ils ont un jeu plus efficace en attaque avec ballon ce

1 Récupérer un rebond offensif c'est prendre possession de la balle suite à un tir manqué d'un partenaire ou de soi-même, le plus souvent avant que celle-ci ne touche le sol.

2 Une situation défavorable de tir se définit par une probabilité faible de réussite au tir, par exemple quand le tireur est placé trop loin du panier, selon un angle trop fermé, quand il est marqué par un défenseur, ou en déséquilibre... le basket étant aussi un jeu d'adresse, la situation favorable de tir est largement recherchée dans l'apprentissage.

qui nous montre une certaine contradiction : un bon joueur devrait recevoir plus souvent la balle de la part de ses partenaires, et se démarquer efficacement.

En défense, ces élèves sont aussi surreprésentés parmi ceux qui ont le plus grand nombre d'actions défavorables (statisme, placement ou déplacement non favorable, retour en défense très tardif...) et le plus petit score d'actions positives (interception, placement et déplacement permettant de freiner l'attaque adverse, dissuasion de la passe adverse, contestation du tir). Enfin, ces élèves sont surreprésentés parmi ceux ayant un score total d'actions négatives élevé, ce qui est différent du niveau de jeu global car ici nous ne prenons pas en compte les actions favorables pouvant contrebalancer le score global.

4. Discussion

Au terme de cette recherche nous observons que les relations socio-affectives peuvent interférer avec les conduites décisionnelles dans une situation collective de coopération et d'opposition à but opératoire. Cette situation scolaire nécessite l'implication de chacun pour optimiser le niveau de compétence collective de l'équipe. Dans le monde du travail, la notion de compétence collective suscite un engouement non négligeable en valorisant la capacité des équipes à résoudre des problèmes et à s'adapter aux aléas (Chétodel, Pujol, 2009). Bataille (op.cit.) a montré qu'il ne suffit pas de faire rencontrer les personnes les plus compétentes pour qu'elles deviennent un collectif compétent. C'est dans ce sens que nous nous sommes intéressés à la variable affective définissant chaque dyade d'individus et pouvant selon nous avoir des répercussions sur les décisions individuelles et donc collectives prises dans une situation avec interaction.

En effet, dans notre expérience en situation collective nous avons constaté que les données sociométriques sont parfois corrélées de façon très significative aux décisions prises au cours du jeu. Tout d'abord la composition des équipes tenant compte des relations socio-affectives au sein de la classe nous montre que même en situation à but opératoire, le facteur relationnel reste présent et peut amener plus ou moins de coopération entre les membres du groupe. Ainsi, les deux équipes formées d'élèves s'appréciant ont mis en place un jeu significativement plus coopératif. Leur niveau de performance n'est pas supérieur à celui des autres, mais ces élèves se conduisent d'une telle façon dans le jeu qu'ils collent davantage aux exigences des grilles de compétences attendues définies par les programmes officiels de la discipline (bulletin officiel spécial n° 4 du 29 avril 2010 et fiches d'accompagnement de février 2011).

Ces documents préconisent en attaque collective comme capacités à acquérir pour le niveau 3 en basket, de progresser collectivement vers la cible adverse, d'utiliser un relais, d'exploiter la profondeur et d'utiliser les couloirs latéraux. Or on s'aperçoit à l'issue de notre recherche que les équipes d'amis se conduisent davantage dans cette direction que les autres, sans pour autant gagner le match. Les référentiels d'évaluation aux examens en EPS tiennent compte du niveau de maîtrise, c'est-à-dire des moyens mis en œuvre, beaucoup plus que de la victoire qui représente ici le but du jeu. Les deux équipes d'amis ont ainsi devancé les autres équipes au niveau de leur capacité à se démarquer pour aider le porteur de balle, de leur récupération des rebonds offensifs, et de leur recherche de solutions pour éviter les tirs en situation défavorable. La recherche de conservation de la balle de façon collective est privilégiée pour ces deux équipes, mais surtout elle concerne l'ensemble des joueurs. Tous ont progressé dans le jeu sans ballon en attaque. L'atteinte de la compétence relative à l'organisation collective en attaque paraît facilitée par la bonne entente des joueurs de l'équipe.

Chétodel et Pujol (op. cit.) ont mis en avant le processus d'identification sociale pour mieux comprendre l'articulation entre compétence individuelle et collective. Dans la sphère professionnelle, il a été montré que la diversité de compétences individuelles permet d'augmenter le répertoire de ressources potentielles pour faire face à la situation rencontrée en milieu collectif (Thomas et Ely, 2001). Mais ceci peut être potentiellement source de tensions car crée des

discriminations par exemple entre groupes sociaux, au point parfois de créer un rejet du principe même de l'équipe (dis-identification) chez certains sujets : celui-ci ne s'identifie pas à son groupe, la communication se révèle donc difficile et les conflits apparaissent. Notre expérimentation va dans ce sens en mettant en lumière que les groupes d'amis, même s'ils sont au final pas plus efficaces que les autres (au moins pour la période donnée et non à long terme), s'organisent mieux collectivement et cherchent à offrir des solutions aux membres de l'équipe pour atteindre leur but commun. La communication entre eux (non verbale) leur permet de mieux s'organiser et de s'entraider davantage que dans les autres équipes.

Nos résultats mettent aussi en avant l'importance du jeu sans ballon en jeux collectifs. Les actions répertoriées sont bien plus nombreuses sans ballon, qu'avec ballon, le temps de possession de balle par joueur au cours d'un jeu étant très court. Evaluer les actions d'un joueur c'est prendre en compte ses actions à chaque moment du match en fonction de la situation de jeu. La centration que peut avoir un spectateur sur le porteur de balle ne rend pas compte du niveau global de ce joueur, ni du travail de ses partenaires pour optimiser cette possession de balle. D'ailleurs dans le cadre scolaire les programmes et référentiels d'évaluation traduisent bien ce constat en accentuant les capacités à développer autour du jeu sans ballon et de l'entraide entre joueur (BO 31 du 06/09/2007). Jouer sans ballon c'est aider, soutenir le porteur de balle, mais aussi oser se démarquer au risque de ne pas avoir la balle. S'investir dans le démarquage et les actions d'entraide est un acte de pure coopération, d'autant plus qu'il n'aboutit évidemment pas à chaque fois à une gratification individuelle se traduisant par la récupération de la balle.

Individuellement maintenant, certains élèves dispersés dans les équipes ont aussi un profil de jeu significativement différent de celui des autres joueurs. C'est notamment le cas des élèves rejetés. Bien qu'ils ne soient pas d'un niveau de jeu inférieur à celui des autres, ils reçoivent moins la balle de la part de leurs partenaires et ont un score en démarquage significativement inférieur à celui des autres : leur placement et déplacement pour aider le porteur de balle de leur équipe est moins efficace. Comme ceci n'est pas corrélé à leur niveau de jeu global, nous pouvons nous demander si cette faiblesse ne résulte pas d'un choix, d'un moindre effort pour coopérer, tout en sachant qu'ils reçoivent aussi moins souvent de passes que les autres ? En effet, être rejeté par les autres et obtenir moins souvent de passe ne motive pas à s'investir et à faire des efforts pour aider. Selon Moreno (Mucchielli, 1973, p. 9) « la dimension sociale de l'individu est l'essentiel de la personnalité ». Et tout groupe humain a une structure affective informelle qui détermine les comportements des individus du groupe, les uns par rapport aux autres. Apprendre en classe, c'est interagir avec les autres, c'est aussi s'adapter aux autres, et donc communiquer (Abric, 1996). L'enjeu est fondamental au niveau des apprentissages sociaux mais aussi au niveau des apprentissages cognitifs, moteurs et affectifs. En EPS la dimension collective des apprentissages est reconnue. Une nouvelle étape a été franchie en déterminant des niveaux de compétences collectives, notamment au baccalauréat dans les activités sociomotrices où une partie de l'évaluation est collective. Par contre de nouveaux obstacles apparaissent en contexte collectif, contrairement aux activités sans interaction et ne sollicitant que des ressources individuelles. La présence interactive des autres a pour conséquence que les relations socio-affectives font partie de la situation collective et ne peuvent être écartées, sous prétexte de la poursuite d'un but commun, ou d'une évaluation.

5. Conclusion

L'accent mis sur la dimension socio-affective nous paraît importante car corrélée à la façon de jouer des élèves, et à leurs progrès. Que ce soit au niveau de la composition des équipes, ou des caractéristiques relationnelles plus individuelles nous notons des corrélations très significatives qui amènent les élèves à se centrer davantage sur des actions de coopération ou de non-coopération avec les joueurs de leur équipe. Or les référentiels d'évaluations scolaires prennent en compte ces notions d'adversité et de coopération (stratégie collective en attaque et défense, Bulletin Officiel spécial n° 4 du 29 avril 2010). Les joueurs, selon les affinités entretenues avec leurs pairs sont donc dans des situations ne leur permettant pas toujours d'exceller dans tous

les domaines. C'est par exemple le cas des élèves rejetés dans notre recherche, qui reçoivent moins la balle que les autres de la part de leurs partenaires.

L'activité collective entraîne une sollicitation particulière des compétences individuelles, du fait même de la présence d'autrui, et donc des relations affinitaires entre les joueurs. Or, cet appui sur les relations affectives n'est que rarement mis en relation avec les conduites motrices observées en EPS (sauf par exemple Parlebas, 1986 ou Bordes, 2002). Plus souvent, la gestion de groupe est appréhendée par rapport à l'apprentissage moteur, par la réalisation de groupes de niveaux, de tutorat, de groupes homogènes, mixtes...les critères utilisés par les enseignants pour composer les groupes étant majoritairement de l'ordre des habilités motrices. Or, nous constatons que la variable sociométrique implique aussi des conduites motrices spécifiques, des choix particuliers relatifs notamment à la coopération. La compréhension de l'importance des facteurs affectifs et psychologiques au sein des groupes a été au cœur de la problématique de dynamique des groupes impulsée par Lewin dès les années 1940. Mais dans le cadre scolaire, la dimension sociale des apprentissages reste peu exploitée. La logique institutionnelle hiérarchise les élèves dans un but d'orientation, évince les relations affectives ce qui revient à nier les sentiments, émotions et affects des apprentissages qui pourtant en font partie (Wallon, 1934). L'activité collective qui paraît être déterminante à l'école pour former les futurs citoyens à leur vie sociale et professionnelle reste peu exploitée, mais surtout est indifférente aux données socio-affectives que l'on tente d'évincer à l'école, pour ne pas qu'elles interfèrent avec la réalisation des tâches en groupe. Or, même la poursuite d'un but commun n'est pas garante d'une disparition des affects existants entre les personnes, surtout à l'école, à l'âge où la construction identitaire passe par la reconnaissance sociale, les liens d'amitié et la recherche d'appartenance et d'identification au groupe. D'ailleurs, les enjeux dans le cadre scolaire sont différents de ceux que l'on pourrait trouver dans d'autres contextes comme par exemple la sphère professionnelle. Nos résultats obtenus en EPS auraient certainement été différents en contexte sportif dans un cadre comportant d'autres enjeux, c'est-à-dire une logique externe différente, dont les caractéristiques auraient été répercutées d'une autre façon sur les conduites motrices (Dugas, 2011).

Bibliographie :

- Abric, J-C. (1996). *Psychologie de la communication : méthodes et théories*. Paris : A. Colin.
- Bataille, F. (1999). *Compétence collective et management des équipes opérationnelles. Une étude longitudinale de Philips Consumer Communications*. Thèse de doctorat en sciences de gestion, Université de Caen.
- Bordes, P. (2002). Les effets des modes de regroupement des élèves en EPS. In *Hyper EPS*. 218, 10-14.
- Caillot, M. (1994). Des objectifs aux compétences dans l'enseignement scientifique : une évolution de vingt ans. In F. Ropé & L. Tanguy, *Savoirs et compétences. De l'usage social des notions à leur problématisation*. Paris : L'Harmattan. 95-117.
- Chétodel, F. & Pujol, L. (2009). De la compétence individuelle à la compétence collective : l'influence des processus d'identification sociale. In D. Retour, T. Picq & C. Defélix (Eds.) *Gestion des compétences - Nouvelles relations nouvelles dimensions*. Paris : Vuibert coll. AGRH.
- Dugas, E. (2006). La evaluación de las conductas motrices en los juegos colectivos : presentación de un instrumento científico aplicado a la educación física. In *Revue internationale Apunts*, Espagne. 83, 61-69.
- Dugas, E. (2011). *L'homme systémique. Pour comprendre les pratiquants de jeux sportifs*. Nancy : Presses Universitaires de Nancy. Collection épistémologie du corps.
- Gréhaigne, J-F., Guillon, & R. Roche, J. (1991). Contribution à une réflexion sur les savoirs en sports collectifs à l'école. In G. Bui-Xuan & J. Gleyse (Eds.). *Enseigner l'Education Physique et Sportive*. Clermont-Ferrand : AFRAPS. 27-38.
- Moreno, J. L. (1954). *Fondements de la sociométrie*. Paris : PUF.
- Mucchielli, R. (1973). *La dynamique des groupes*. Paris : ESF.

Parlebas, P. (1986). *Eléments de sociologie du sport*. Paris : PUF.

Parlebas, P. (1992). *Sociométrie, réseaux et communication*. Paris : PUF.

Parlebas, P. (2002). Réseaux dans les jeux et sports. In *l'année sociologique*. 2002/2, 52, 315-349.

Ringelmann, M. (1913). Recherches sur les moteurs animés : travail de l'homme. *Annales de l'Institut National Agronomique*, 2nd series, 12, 1-40.

Wallon, H. (1934/1976). Les origines du caractère chez l'enfant. Paris : Boivin. 6ème éd., Paris : PUF, 1976.

DETERMINANTS PERSONNELS DES COMPETENCES DE L'ECOLIER : LE CAS DE SA PERFORMANCE SCOLAIRE ET DE SES REACTIONS ENVERS LA CLASSE

Halim Bennacer

IRÉDU-CNRS,¹ Université de Bourgogne
halim.bennacer@u-bourgogne.fr

Mots-clés : *compétences, caractéristiques personnelles, caractéristiques affectives de départ, école élémentaire, analyse en pistes causales*

Résumé. *Cette étude analyse les déterminants personnels des compétences cognitives de l'élève (soit sa performance scolaire) et de celles relevant du savoir-être (ou ses réactions envers la classe). Nous testons un modèle théorique, par l'analyse en communautés (n = 554), quant à son applicabilité aux principes d'association et de non-artificialité. Les caractéristiques externes et cognitives antérieures de la personnalité (âge, sexe, CSP, passé scolaire) agissent directement sur les critères, et indirectement à travers les mécanismes dynamiques de la personnalité relatifs à cinq attitudes envers l'école (Bennacer, 2008). À l'aide de l'analyse en pistes causales, nous aboutissons à un modèle structural qui explique 39 % de la variance dans la performance scolaire, et de 19 à 39 % de la variance dans les réactions envers la classe. Nous déterminons, finalement, le profil de l'élève ayant de mauvais résultats scolaires et des réactions négatives envers la classe.*

1. Introduction

Tous les auteurs s'accordent à dire que la réussite de l'élève et ses compétences scolaires dépendent de sa personnalité, ou de ses caractéristiques personnelles (Chamarro-Premuzic & Furnham, 2003 ; Potvin & Paradis, 2000). Cette étude porte sur les déterminants personnels de deux types de compétences chez l'élève : celles d'ordre cognitif et les compétences relevant du savoir-être. Les premières concernent sa performance scolaire, les secondes ses réactions envers la classe relatives à son bien-être, à sa satisfaction envers l'apprentissage, le maître et la classe.

La quasi-totalité des chercheurs a tenté d'expliquer les compétences et le comportement scolaires de l'élève par ses caractéristiques externes, d'ordre statique, qu'il importe de retenir. Il s'agit de l'âge, du sexe et de la classe socioprofessionnelle des parents. Il a été prouvé que le développement des individus diffère selon l'âge (voir Guidetti, 2002). L'augmentation dans l'âge s'avère, par exemple, positivement liée aussi bien à l'absentéisme et à la performance scolaire des élèves qu'à leurs attitudes négatives et leur insatisfaction envers le collège (Bennacer, 1994). Concernant le sexe, des travaux montrent que les filles et les garçons présentent, dès la rentrée à l'école, des différences comportementales typiques à leur genre liées aux stéréotypes caractérisant l'environnement socio-familial (Duru-Bellat, 2004 ; Lajoie, 2003). Sous l'effet de ces stéréotypes, les filles ont souvent appris l'écoute, le respect, la souplesse et certaines valeurs sociales qui facilitent l'acceptation des exigences scolaires. Alors que les garçons se caractérisent plus par l'individualisme, la quête d'autonomie, le non-respect des règles, la turbulence, le chahut scolaire et des conduites asociales (Deschesnes *et al.*, 2003). Les filles développent des attitudes plus favorables envers l'école, consacrent plus de temps à leurs devoirs et ont plus d'aspirations scolaires (Cloutier, 2003). Par rapport aux garçons, elles créent ainsi les conditions de leur meilleure réussite en adoptant des conduites qui expriment une forte adhésion aux normes et aux

¹ Institut sur la Recherche en Éducation (IRÉDU), UMR.

valeurs de l'école (Lajoie, 2004). Au niveau de la classe socioprofessionnelle des parents, on a constaté que les attitudes des élèves et leur performance scolaire sont d'autant meilleures que l'on augmente dans son importance. Ce sont les élèves des classes moyennes et aisées qui montrent plus d'affection envers l'école (Francis, 1992) et obtiennent une bonne réussite scolaire (Jeantheau & Murat, 1998). Ce constat a fait l'objet de nombreuses controverses scientifiques et politiques en sociologie de l'éducation (Forquin, 1996). Les recherches insistent également sur l'importance des caractéristiques cognitives antérieures, relatives à l'histoire ou au passé scolaire de l'élève (Heller *et al.*, 1978). Celle-ci s'est avérée un déterminant majeur autant de ses attitudes envers lui-même et envers l'école (Gladney, 1998) que de sa réussite ultérieure et de l'abandon des études (Janosz *et al.*, 2000).

Les traits généraux de la personnalité n'ayant pas permis d'expliquer la performance et les compétences scolaires de l'élève (voir Kneip, 1979), les auteurs se sont tournés vers l'étude des dimensions affectives spécifiques de type situationnel, dites « dynamiques ou motivationnelles », qui se développent dans l'école. À l'issue de l'examen d'une cinquantaine d'études, Bloom (1979) propose pour désigner ces dimensions le concept des « caractéristiques affectives de départ ». Celles-ci se constituent selon lui des attitudes et intérêts envers les matières scolaires, des attitudes vis-à-vis de l'école et de la perception de soi scolaire. L'auteur constate que ces caractéristiques permettent d'expliquer jusqu'à 25% de la variance dans la performance scolaire. Des études plus récentes soutiennent leur importance dans la détermination du comportement et du développement de l'élève (voir Potvin & Paradis, 2000). Il ressort, par exemple, que les attitudes positives de l'élève au sein de l'établissement scolaire telles que sa satisfaction générale envers l'école, son engagement ou intérêt scolaire et ses aspirations éducationnelles découragent chez lui les problèmes comportementaux comme la délinquance juvénile (Mak *et al.*, 2003), le chahut scolaire, l'usage fréquent de la cigarette et de l'alcool (Perry *et al.*, 1993) ainsi que l'abandon des études (CRSC, 1999). De plus, elles s'accompagnent souvent d'autonomie, de motivation (Cock & Halvari, 1999), de contrôle interne et d'une bonne estime de soi surtout scolaire (Saint-Laurent, 2000). Les attitudes favorables à l'égard de l'école constituent aussi de bons prédicteurs des résultats scolaires de l'élève (Abu-Hilal, 2000 ; Bragard, 2000). Elles déterminent favorablement ses bonnes réactions vis-à-vis de la classe comme le bien-être, sa satisfaction envers le professeur et l'enseignement (Bennacer, 2003).

De la revue de la littérature, force est de constater que les recherches ont souvent porté sur *de simples associations* entre certaines variables personnelles et la performance ou les compétences scolaires de l'élève. Elles n'ont cependant pas tenté de saisir, dans un modèle théorique de base, les mécanismes dynamiques par lesquels la personnalité de l'élève influe sur sa performance ou ses compétences scolaires. Cette approche conduit à une compréhension superficielle du phénomène étudié. Force est de constater par ailleurs qu'aucune recherche n'a étudié, au sein d'un même modèle théorique, les relations entre les compétences scolaires d'ordre cognitif et celles relevant du savoir-être, relatives respectivement à la performance de l'élève et ses réactions envers la classe telles que son bien-être, sa satisfaction envers l'apprentissage, le maître et la classe. Cette étude a pour intérêt principal d'identifier chez l'élève le rôle de certains facteurs personnels et mécanismes dynamiques de la personnalité dans l'explication de ces deux types de compétences. Nous y identifions aussi les relations et l'articulation entre ces derniers.

2. Modèle théorique de base et hypothèses

L'examen des travaux a permis de développer un modèle de base, qui récapitule les considérations théoriques de l'étude (voir Figure 1). Il est considéré que la personnalité de l'élève détermine sa performance scolaire et ses réactions envers la classe directement par des variables *distales*, et indirectement à travers des variables *proximales*. Les premières ne sont pas proches des phénomènes expliqués, et sont – relativement – stables dans les situations. Elles se constituent des facteurs d'entrées relatifs aux caractéristiques externes (*âge, sexe, classe socioprofessionnelle des parents*) et cognitives antérieures de l'élève (*telles que histoire ou passé scolaire*). Les secondes, qui font comprendre le pourquoi du comportement (Cattell, 1965), renvoient aux déterminants les plus

directs des critères et se rapportent aux attitudes affectives de départ dites motivationnelles (Bloom, 1979). Elles sont dynamiques, car elles évoluent et se structurent dans l'école selon les situations, et jouent un rôle intermédiaire dans le mécanisme d'influence en dépendant des variables distales et expliquant en même temps les critères.

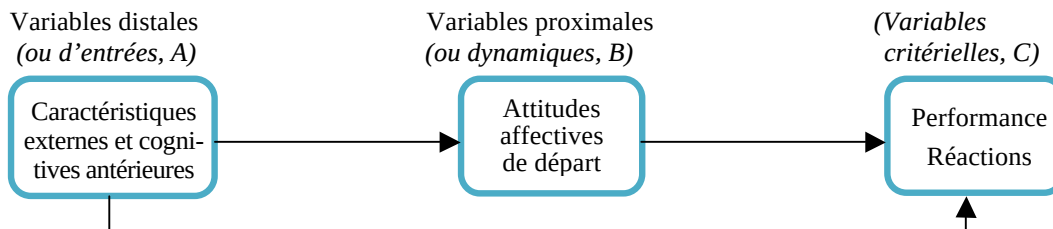


Figure 1. Modèle théorique de base sur les déterminants personnels de la performance scolaire de l'élève et de ses réactions envers la classe

Suivant l'ordre causal adopté, ce modèle clarifie les relations entre les blocs de variables retenus. Il récapitule ainsi les principales hypothèses de l'étude :

1. Les attitudes affectives de départ (bloc B) sont liées aux caractéristiques externes et cognitives antérieures de l'élève (bloc A), qui les précèdent dans l'ordre causal.
2. Ces deux groupes de variables exercent, individuellement et ensemble, des effets spécifiques dans la prédiction de la performance de l'élève et de ses réactions envers la classe (bloc C).
3. Les attitudes affectives de départ (bloc B) apportent, indépendamment des caractéristiques externes et cognitives antérieures (bloc A), des effets – ou explications – supplémentaires qui leurs sont propres dans l'explication des variables critérielles de l'étude (bloc C).

3. Méthode

3.1 Participants

Ils concernent 554 élèves, âgés de 6 à 12 ans (moyenne = 9.22, 51,5 % de filles). Ces élèves fréquentent 25 classes réparties équitablement, et choisies selon la méthode d'échantillonnage par grappes, en fonction des cinq niveaux scolaires de l'école élémentaire.

3.2 Mesures

Trois caractéristiques externes de l'élève sont retenues à savoir son sexe (1 = garçon, 2 = fille), son âge (ou années vécues) et l'origine socioprofessionnelle de ses parents (1 = famille favorisée, 2 = moyenne, 3 = défavorisée). Les caractéristiques cognitives antérieures ont porté sur le passé scolaire de l'élève, se rapportant au nombre d'années d'échecs qu'il a connu dans sa scolarité. Les variables dynamiques de la personnalité furent représentées par les cinq dimensions de l'*Échelle des attitudes des élèves envers l'école élémentaire* (ÉAÉE, Bennacer, 2008). Il s'agit de l'affectivité négative envers l'école, du désintérêt scolaire, des aspirations pour les études futures ainsi que des facteurs : « anxiété, souci et stress scolaires » et « école drôle, lieu de distraction » (voir Tableau 1).

Dans le questionnaire, nous avons aussi introduit les variables critérielles (dites aussi critères) relatives à la performance scolaire de l'élève, ou la note globale accordée par le maître, et à ses réactions envers la classe estimées par quatre items (voir Tableau 2). Celles-ci concernent le bien-être, la satisfaction envers le maître, l'apprentissage et la classe. Nous avons ensuite constitué le facteur : « réaction globale envers la classe », qui correspond au score composite à ces quatre variables et comporte un coefficient alpha de 0.76.

Tableau 1. Description des dimensions de l'  chelle des attitudes des   l  ves envers l'  cole   l  mentaire (  A  E, Bennacer, 2008)

Dimension	Alpha (N = 336)	Exemples d'items
Affectivit�� n��gative envers l'��cole (7 items)	.89	L'��cole, c'est formidable (F). J'aime aller �� l'��cole (F).
D��sint��r��t scolaire (6 items)	.72	En classe, j'��coute souvent mon ma��tre (F). Je m'int��resse beaucoup �� mon travail (F).
Aspirations pour les ��tudes futures (4 items)	.81	Quitter l'��cole, ��a me ferait beaucoup plaisir (V). Je r��ve de ne plus aller �� l'��cole (V).
Anxi��t��, souci et stress scolaires (9 items)	.76	J'ai souvent peur quand je pense �� mon travail (V). �� force de penser �� l'��cole, je m'��nerve souvent (V).
��cole dr��le, lieu de distraction (6 items)	.66	L'��cole, c'est dr��le (V). L'��cole n'est pas un endroit o�� l'on s'amuse (F).

Note. Les r  ponses aux items sont dichotomiques

Tableau 2. Variables crit  rielles de l'  tude et leurs indicateurs

Variable	Indicateur ou item
Performance scolaire	Note globale accord��e �� l'��l��ve par le ma��tre
A. Bien-��tre	Comment te sens-tu dans ta classe ? (4 r��ponses : de "Tr��s mal �� Tr��s bien")
B. Satisfaction envers le professeur ¹	Es-tu content (e) de ton ma��tre ?
C. Satisfaction envers l'apprentissage ¹	Es-tu content (e) de ce que tu apprends en classe ?
D. Satisfaction envers la classe ¹	En g��n��ral, es-tu content (e) de ta classe ?
R��action globale envers la classe	A + B + C + D (alpha = .76, n = 554)

Note. ¹ Quatre r  ponses : "Pas content du tout"    "Tr  s content".

3.3 Proc  dures

La passation du questionnaire a eu lieu au troisi  me trimestre de l'ann  e. Pour   tudier les relations entre les blocs de variables et v  rifier les hypoth  ses, nous avons utilis   l'*analyse en communaut  s* (voir Nimon, 2010). En plus du principe de causalit   respect   dans le cadre th  orique, cette m  thode permet de tester le mod  le g  n  ral quant    son applicabilit   aux principes d'association entre les blocs et de non-artificialit  ². Ceux-ci sont indispensables pour l'explication du comportement, si l'on croit Herschi et Selvin (1975, cit  s in Dicks, 1978). Nous avons ensuite appliqu   une *analyse en pistes causales* (voir Kuha & Goldthorpe, 2010), qui permet une compr  hension plus fine du r  le des variables et de s  lectionner le meilleur mod  le structural.

4. R  sultats

4.1 Test du mod  le de base

4.1.1 Pr  diction des attitudes affectives de d  part

Toutes les corr  lations multiples sont significatives ($p < .001$), et au moins   gales    .15 (voir Tableau 3). Le principe d'association entre les blocs A et B est ainsi respect  . Les caract  ristiques

² La non-artificialit   exige que l'association entre A et B ne disparaisse pas quand on supprime les effets des variables qui les pr  c  dent dans l'ordre causal.

externes et cognitive ant  rieure de l'  l  ve sont li  es    ses attitudes affectives de d  part. Elles permettent d'expliquer entre 2.3 et 7.5 % de la variance dans les attitudes g  n  rales envers l'  cole (soit   cole dr  le, Aspirations pour les   tudes et Affectivit   n  gative). Les parts de variance s'  l  vent dans l'explication des attitudes directement li  es au travail scolaire, et atteignent 11.3 et 12.4 % respectivement pour le d  sint  r  t scolaire et le facteur « Anxi  t  , souci et stress scolaires ».

Tableau 3. R  gressions multiples du bloc relatif aux caract  ristiques externes et cognitive ant  rieure (A) sur les attitudes affectives de d  part (B)

Attitudes affectives de d��part	Rm	Rm ² %
Affectivit�� n��gative envers l'��cole	.336*	07.5
D��sint��r��t scolaire	.274*	11.3
Anxi��t��, souci ou stress scolaires	.352*	12.4
Aspirations pour les ��tudes futures	.259*	06.7
��cole dr��le, lieu de distraction	.152*	02.3

Note. Caract  ristiques externes et cognitive ant  rieure (sexe,   ge, CSP des parents, pass   scolaire). Variance expliqu  e dans une attitude affective (Rm² %). N = 554   l  ves. * p <.001

4.1.2 Pr  diction des variables crit  rielles

Les r  sultats confirment l'hypoth  se selon laquelle les blocs A et B exercent, individuellement et ensemble, des effets sp  cifiques sur la performance scolaire et la r  action globale envers la classe. Toutes les corr  lations multiples sont, en effet, significatives    .001 (voir Tableau 4). Le principe d'association est donc respect  .

Tableau 4. Pr  diction des comp  tences scolaires de l'  l  ve (C). Analyse en communaut  s des blocs de variables relatifs aux caract  ristiques externes et cognitive ant  rieure (A) et aux attitudes affectives de d  part (B)

Bloc de Variables	Performance scolaire (C)				R��action globale envers la classe (C)			
	Rm	Rm ² %	A	B	Rm	Rm ² %	A	B
A	.582*	33.9	14.1		.484*	23.4	08.5	
B	.551*	30.4		10.6	.641*	41.1		26.2
AB	.667*	44.5	19.8	19.8	.704*	49.6	14.9	14.9
Total			33.9	30.4	Total		23.4	41.1

Note. Pour les variables des blocs, voir Tableau 3. Part de variance expliqu  e (Rm² %, *effet sp  cifique*). Les colonnes A et B indiquent 2 parts de variance unique (significatives    .01) se rapportant    un bloc ou    l'influence commune de AB (*effet commun*). N = 554   l  ves. * p <.001.

Les caract  ristiques externes et cognitive ant  rieure expliquent dans la performance scolaire une part de variance sp  cifique, l  g  rement, plus   lev  e que celle des attitudes affectives de d  part (33.9 contre 30.4 %). Elles rendent compte, cependant, de parts moins importantes dans la r  action globale envers la classe que celles relatives aux attitudes affectives,    savoir 23.4 contre 41.1 %. Ce constat soutient le bilan des   tudes fait par Bloom (1979), qui montre que les caract  ristiques affectives de d  part expliquent jusqu'   25 % de la variance dans la r  ussite scolaire. Avec l'ensemble des variables des blocs A et B, nous obtenons des valeurs plus importantes que celles se rapportant    l'un ou    l'autre bloc. La validit   pr  dictive du crit  re correspond    44.5 et 49.6 % respectivement pour la performance scolaire et la r  action globale envers la classe.

L'analyse des variances uniques, significatives   .01, montre que chacun des blocs A et B apporte des contributions propres, dans l'explication de la performance scolaire et de la r  action globale envers la classe. La variance unique est de 10.6 et 26.2 %, respectivement. Nous confirmons l'hypoth  se selon laquelle les attitudes affectives de d  part exercent, sur les crit  res, un effet suppl  mentaire   celui des caract  ristiques externes et cognitive ant  rieure. Le principe de non-artificialit   est respect  . L'examen des variances communes, correspondant   19.8 et 14.9 %, permet de constater que l'influence des blocs A et B sur les crit  res se fait aussi, de mani  re conjointe et consid  rable, avec les caract  ristiques affectives de d  part. Les r  sultats relatifs au facteur global (r  action globale) s'appliquent   toutes les r  actions de l'  l  ve envers la classe (voir Tableau 5).

Tableau 5. Pr  diction des r  actions de l'  l  ve envers la classe (C). Analyse en communaut  s des blocs de variables relatifs aux caract  ristiques externes et cognitive ant  rieure (A) et aux attitudes affectives de d  part (B)

Bloc de variables	R��actions de l'��l��ve envers la classe (C)											
	Bien-��tre			Satisfaction envers le ma��tre			Satisfaction envers l'apprentissage			Satisfaction envers la classe		
	Rm	Rm ² %	A B	Rm	Rm ² %	A B	Rm	Rm ² %	A B	Rm	Rm ² %	A B
A	.333*	11.1	02.8	.438*	19.2	08.3	.351*	12.3	03.8	.323*	10.4	05.1
B	.513*	26.3	18.0	.507*	25.7	14.8	.523*	27.4	18.9	.383*	14.7	09.4
AB	.539*	29.1	08.3 08.3	.583*	34.0	10.9 10.9	.559*	31.2	08.5 08.5	.445*	19.8	05.3 05.3
Total			11.1 26.3	Total		19.2 25.7	Total		12.3 27.4	Total		10.4 14.7

Note. Pour les variables des blocs, voir Tableau 3. Part de variance expliqu  e (Rm² %, *effet sp  cifique*). Les colonnes A et B indiquent 2 parts de variance unique (significatives   .01) se rapportant   un bloc ou   l'influence commune de AB (*effet commun*). N = 554   l  ves. * p < .001.

4.2. Affinement du mod  le de base

4.2.1 Premier mod  le structural

L'application   nos donn  es d'une analyse en pistes causales a permis d'aboutir   un premier mod  le structural, qui s'ajuste exactement aux donn  es (voir Figure 2). Le chi-carr   est non significatif   .05 [soit $\chi^2(23) = 27.65$, $p = .229$]. Le RMSEA est seulement de .009. Les mesures d'ajustement GFI, AGFI et CFI correspondent   des valeurs respectives aussi   lev  es que .991, .975 et .992.

4.2.1.1 Relations intra-blocs

Le pass   scolaire est positivement li     l'  ge de l'  l  ve (.23) et   la classe socioprofessionnelle de ses parents (.23). Les enfants ayant connu un   chec scolaire, plus   g  s, proviennent probablement de familles d  favoris  es socialement. En ce qui concerne les attitudes affectives de d  part, il importe de constater que l'affectivit   n  gative envers l'  cole s'accompagne autant d'un degr   tr  s faible en aspirations pour les   tudes futures (-.51) que du d  sint  r  t scolaire (.18), qui renforce l'anxi  t   de l'  l  ve (.26). Celle-ci engendre la d  saffectation envers l'  cole (.34), et n'encourage pas les aspirations pour les   tudes futures (-.19). Le niveau   lev   dans ces derni  res s'accompagne d'une bonne perception de l'  cole comme dr  le et amusante (.19), qui inhibe la d  saffectation envers l'  cole (-.20). Une meilleure performance de l'  l  ve encourage, plus probablement, le d  veloppement d'une bonne r  action globale envers la classe (.13).

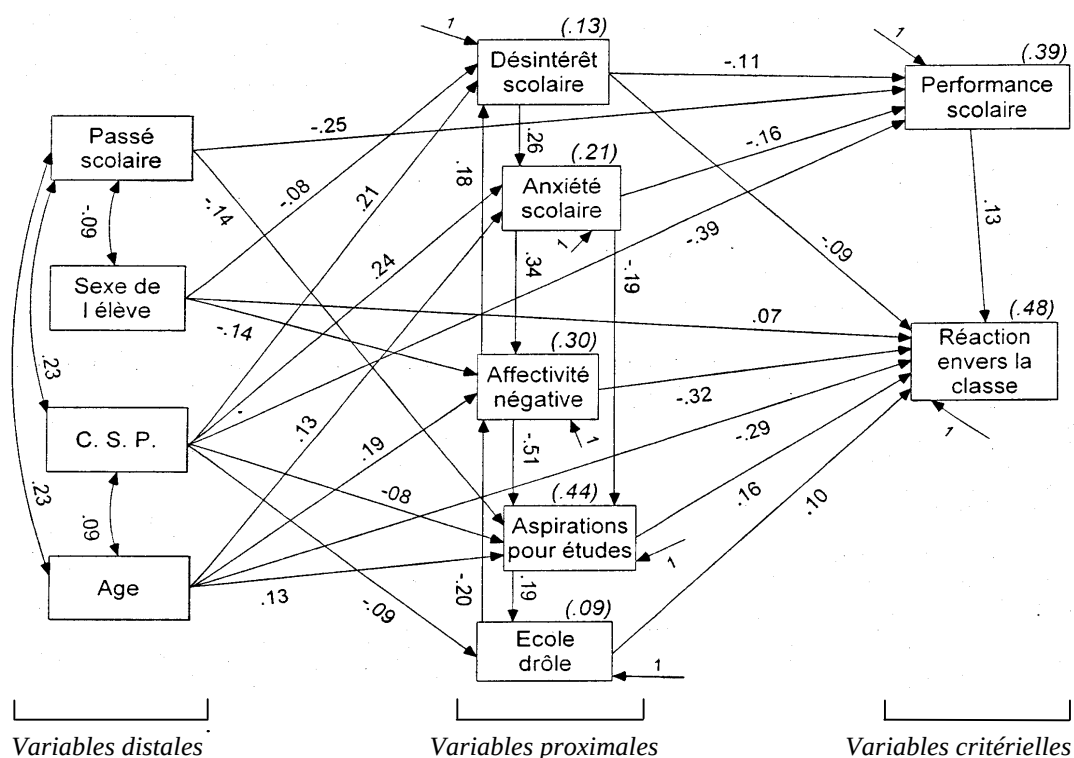


Figure 2. Modèle structural sur les déterminants personnels de la performance scolaire de l'élève et de sa réaction globale envers la classe (Les estimations des paramètres, standardisés, sont significatives à $p < .05$. La part de variance expliquée est indiquée entre parenthèses, pour les variables endogènes. CSP = classe socioprofessionnelle des parents. N = 554 élèves)

4.2.1.2 Relations des variables des blocs A avec les attitudes affectives

Avec le modèle obtenu, nous expliquons entre 9 et 44 % de la variance dans les attitudes affectives de départ. Celles-ci dépendent, en effet, des caractéristiques externes de l'élève. À mesure que l'on augmente dans l'âge, l'élève éprouve une désaffection envers l'école (.19) et une anxiété scolaire (.13) en souhaitant faire de longues études futures (.13). Par rapport aux garçons, les filles montrent moins probablement une affectivité négative envers l'école (-.14) et un désintérêt scolaire (-.08). Les enfants de familles défavorisées développent plus des attitudes affectives négatives, plus particulièrement un désintérêt (.21), une anxiété scolaire (.24) et une perception de l'école comme un lieu ennuyeux (-.09 avec *École drôle*). Le passé scolaire a un effet négatif sur les aspirations pour les études futures (-.14). Ce sont, en effet, les redoublants qui n'espèrent pas faire d'études.

4.2.1.3 Relations des variables des blocs A et B avec les critères

Le modèle permet d'expliquer des variances aussi importantes que 39 et 48 % dans la performance scolaire et la réaction envers la classe. Les caractéristiques externes et cognitive antérieure ont chacune un effet sur, au moins, un critère. L'augmentation dans l'âge engendre plus une mauvaise réaction globale (-.29). Les filles développent légèrement une bonne réaction (.07). À mesure que le niveau de la classe sociale des parents décroît, les élèves obtiennent une moins bonne performance (-.39) comme les redoublants (-.25 avec *Passé scolaire*). Ce qui soutient le constat selon lequel les enfants de familles défavorisées et ceux ayant connu au moins un échec scolaire sont moins performants à l'école que les autres (Forquin, 1996 ; Jeantheau & Murat, 1998). Les attitudes affectives de départ sont liées chacune, au moins, à un critère ; plaidant ainsi en faveur de leur importance (Bloom, 1979 ; Gladney, 1998). Ce sont les attitudes liées au travail en classe, relatives à l'intérêt et à l'anxiété scolaires, qui déterminent de manière négative la performance de l'élève

(-.11, -.16). Alors que les attitudes générales envers l'école influencent la réaction globale vis-à-vis de la classe. Cette dernière est d'autant meilleure que l'élève manifeste une faible désaffection envers l'école (-.32), des aspirations pour les études (.16) et perçoit l'école comme drôle (.10).

4.2.2 Deuxième modèle structural

Nous a obtenu un deuxième modèle portant sur les différentes réactions envers l'école (voir Figure 3), qui présente un bon ajustement aux données. Le chi-carré est non significatif à .05 [soit $\chi^2(31) = 46.32, p = .44$]. Le RMSEA est égal seulement à .010. Les mesures d'ajustement GFI, AGFI et CFI correspondent à des valeurs respectives aussi élevées que .988, .978 et .999. Le modèle permet d'expliquer des variances qui varient entre 19 et 39 % dans les réactions envers la classe.

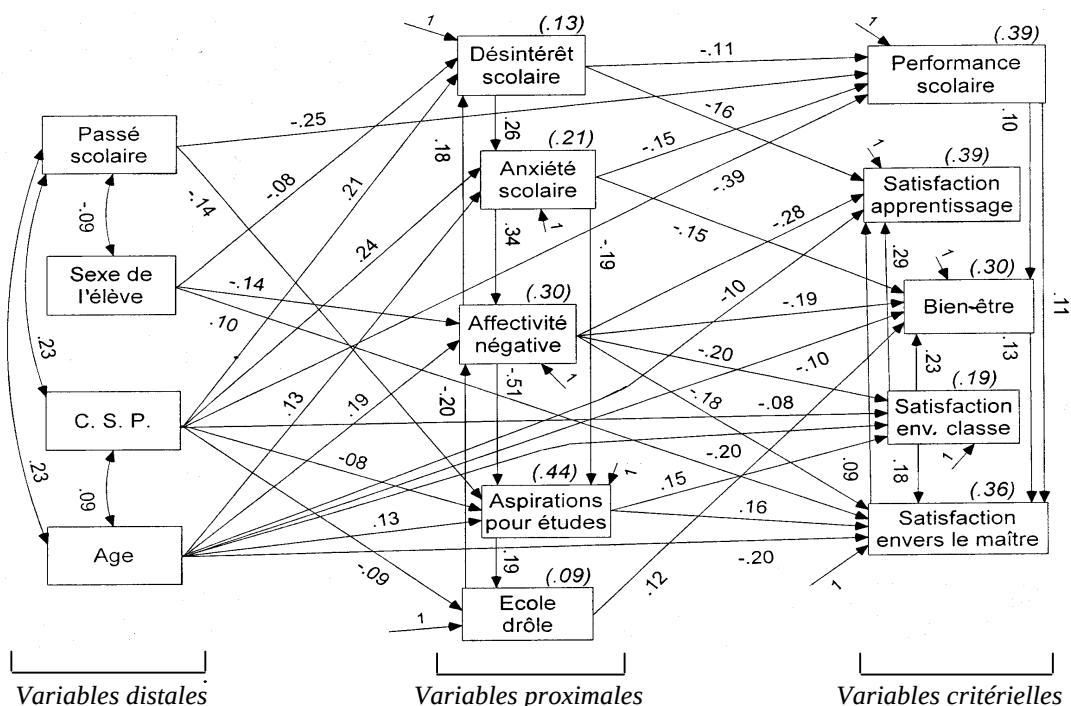


Figure 3. Modèle structural sur les déterminants personnels de la performance scolaire de l'élève et de ses réactions envers la classe (Les estimations des paramètres, standardisés, sont significatives à $p < .05$. La part de variance expliquée est indiquée entre parenthèses, pour les variables endogènes. CSP = classe socioprofessionnelle des parents. N = 554 élèves)

4.2.2.1 Relations au sein du bloc C

Au niveau des relations entre les critères, il ressort d'abord que la satisfaction envers la classe exerce autant une influence directe sur la satisfaction vis-à-vis de l'apprentissage (.29), qu'un effet indirect à travers le bien-être (.23) et la satisfaction envers le maître (.18). Il s'avère aussi que la performance scolaire a un effet positif sur le bien-être de l'élève (.10) et sa satisfaction envers le maître (.11). Cependant, elle ne fait l'objet d'aucune influence de la part des réactions de l'écopier.

4.2.2.2 Relations avec les critères

Les attitudes générales de l'élève envers l'école déterminent ses quatre réactions envers la classe. C'est l'affectivité négative qui engendre le plus des réactions défavorables (-.18 à -.28), suivie des aspirations pour les études futures qui favorisent la satisfaction envers la classe et le maître (.15, .16). Vient, en dernier lieu, le facteur : « École drôle » qui encourage chez l'élève son bien-être (.12). Celui-ci dépend aussi, négativement, de l'anxiété scolaire (-.15). L'augmentation dans l'âge

engendre toujours de mauvaises r  actions envers la classe (-.10    -.20). Par rapport aux gar  ons, les filles sont probablement plus satisfaites de leur ma  tre (.10 avec Sexe). Les enfants de familles d  favoris  es s'av  rent l  g  rement moins satisfaits de leur classe (.08 avec CSP).

5. Conclusion et discussion

Dans ce travail, nous confirmons le mod  le de base en respectant les principes indispensables    l'explication du comportement requis par Herschi et Selvin (1975; in Dickes, 1978, voir plus haut). Il ressort que la personnalit   de l'  l  ve agit sur sa performance et ses r  actions envers la classe de mani  re directe par les caract  ristiques externes et cognitive ant  rieure, ou indirectement    travers les facteurs dynamiques relatifs aux attitudes affectives de d  part. Celles-ci jouent un r  le interm  diaire dans le m  canisme d'influence, en d  pendant des derni  res caract  ristiques et pr  disant les divers crit  res. Elles expliquent d'importantes parts de variances sp  cifiques (14.7    41.1 %, soutenant ainsi le bilan d'  tudes selon lequel ces attitudes expliquent jusqu'   25 % de la variance dans la r  ussite scolaire (Bloom, 1979). De plus, elles apportent d'importantes contributions suppl  mentaires dans l'explication des crit  res (9.4    26.2 %), ce qui plaide en faveur du respect du principe de non-artificialit  . Les caract  ristiques externes et cognitive ant  rieure ont cependant un poids moins   lev   dans la pr  diction des diverses r  actions envers la classe (variances uniques = 2.8    8.5 % uniquement).

Cette   tude sugg  re d'agir sur les caract  ristiques (ou attitudes) affectives de d  part pour favoriser le d  veloppement des r  actions favorables envers la classe,    l'instar des zones d'enseignement prioritaires (ZEP). Dans ces derni  res, les enfants pr  sentent une meilleure estime de soi que les   l  ves hors ZEP (Maintier, 2004), gr  ce surtout aux moyens mis    la disposition de leur   cole. Il serait cependant difficile d'am  liorer consid  rablement la performance scolaire de l'  l  ve, qui d  pend plus des effets des caract  ristiques externes et cognitives ant  rieures (voir Bennacer, 2005).

Nous avons affin   le mod  le g  n  ral en aboutissant    deux mod  les structuraux, qui d  terminent le r  le de chacune des variables retenues. Ce sont les attitudes n  gatives directement li  es au travail en classe,    savoir l'int  r  t et l'anxi  t   scolaires, qui expliquent la faible performance de l'  l  ve. Alors que les mauvaises attitudes g  n  rales envers l'  cole, notamment l'affectivit   n  gative, pr  disent les r  actions d  favorables envers la classe. Cette   tude montre que la performance scolaire agit sur le bien-  tre de l'  l  ve et sa satisfaction envers le ma  tre, sans subir aucune influence la part des r  actions de l'  l  ve envers la classe. Elle permet, aussi, de d  terminer les caract  ristiques personnelles de l'  colier qui manifeste des attitudes n  gatives envers l'  cole, une faible performance scolaire et/ou des r  actions d  favorables vis-  -vis de la classe.

6. R  f  rences

- Abu-Hilal, M.-M. (2000). A structural model of attitudes towards school subjects, academic aspiration and achievement. *Educational Psychology*, 20, 75-84.
- Bennacer, H. (1994).   valuation des attitudes des   l  ves envers l'  cole :   laboration d'un instrument et   tude des mesures comme variables d  pendantes et ind  pendantes. *Psychologie et Psychom  trie*, 15(1), 44-64.
- Bennacer, H. (2003). Validation d'une   chelle des attitudes des   l  ves envers l'  cole. *Revue Canadienne des Sciences du Comportement*, 35(1), 50-55.
- Bennacer, H. (2005).   tude des d  terminants personnels du comportement scolaire de l'  l  ve : test d'un mod  le structural. *Psychologie Fran  aise*, 50(4), 451-469.
- Bennacer, H. (2008). Les attitudes des   l  ves envers l'  cole   l  mentaire et leur   valuation. *Revue Europ  enne de Psychologie Appliqu  e*, 58(2), 75-87.
- Bloom, B.-S. (1979). *Caract  ristiques individuelles et apprentissage scolaire* (1976). Bruxelles : Labor.
- Bragard, J.-L. (2000). *Concept de soi, performances scolaires et apprentissages fondamentaux*. Th  se de doctorat, Universit   Nancy 2.

- Chamarro-Premuzic, T. & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality*, 17, 237-250.
- Cloutier, R. (2003). La réussite scolaire des garçons : un défi à multiples facettes. *Vie Pédagogique*, 127, 9-12.
- Cock, D. & Halvari, H. (1999). Relations among achievement motives, autonomy, performance in mathematics, and satisfaction of pupils in elementary school. *Psychological Reports*, 84, 983-997.
- CRSC (centre de recherche sur les services communautaires) (1999). L'abandon scolaire. *Bulletin du CRSC* (Québec), 19 (www.crsc.ulaval.ca).
- Deschesnes, M., Demers, S. & Finès, P. (2003). *Styles de vie des jeunes du secondaire en Outaouais, 1991-1996-2002*. Direction de santé publique, RRSS de l'Outaouais (www.santepublique-outaouais.qc.ca).
- Dickes, P. (1978). *Le savoir-lire de l'enfant en première année de l'école primaire : évaluation et prodromes*. Thèse de doctorat, Université de Liège.
- Duru-Bellat, M. (1995). Filles et garçons : Approches sociologiques et psychosociales, 2^{ème} partie (La construction scolaire des différences entre les sexes). *Revue Française de Pédagogie*, 110, 75-109.
- Francis, L.-J. (1992). The influence of religion, gender, and social class on attitudes toward school among 11-year-olds in England. *Journal of Experimental Education*, 60, 339-348.
- Forquin, J.-C. (1996). *École et culture : le point de vue des sociologues britanniques*. Bruxelles : De Boeck Univ.
- Gladney, L.-S. (1998). Fictive Kinship, racial identity, peer influence, attitudes toward school, and goals: Relationships with achievement for African American high school students. *Dissertation Abstracts International Section*, 59A, 1066.
- Guidetti, M. (2002). *Les étapes du développement psychologique*. Paris : Armand Colin.
- Heller, K., Rosemann, B. & Steffens, K.-H. (1978). *Prognose des schulerfolgs*. Beltz: Weinheim.
- Jeantheau, J.-P. & Murat, F. (1998). *Observation à l'entrée au CP du "panel 1997"*. Ministère de l'éducation nationale, de la recherche et de la technologie, Note d'information n° 98-40.
- Janosz, M., Fallu, J.-S. & Deniger, M.-A. (2000). La prévention du décrochage scolaire. Facteurs de risque et efficacité des programmes d'intervention. In F. Vitaro & C. Gagnon (Eds.), *Prévention des problèmes d'adaptation chez les enfants et les adolescents* (pp. 117-163). Québec : P U du Québec.
- Jeantheau, J.-P. & Murat, F. (1998). *Observation à l'entrée au CP du "panel 1997"*. Ministère de l'éducation nationale, de la recherche et de la technologie, Note d'information n° 98-40.
- Kneip, E. (1979). *Les déterminants de la performance scolaire*. Thèse de doctorat, Université Nancy II.
- Kuha, J. & Goldthorpe, J.-H. (2010). Path Analysis for Discrete Variables: The role of Education in Social Mobility. *Journal of the Royal Statistical Society*, 173(2), 351-369.
- Lajoie, G., 2003. *L'École au Masculin*. Québec : Septembre Éditeur, Sainte-Foy.
- Lajoie, G. (2004). *Les garçons, les filles et l'école*. Consultable sur le site : www.petitmonde.com.
- Maintier, C. (2004). L'estime de soi des élèves au sein d'écoles élémentaires françaises. In ADRIPS 5^{ème} *Congrès International de Psychologie Sociale en Langue Française* (pp. 431-432). Grenoble : PUG.
- Mak, A., Heaven, P. C.-L. & Rummery, A., 2003. The role of group identity and personality domains as indicators of self-reported delinquency. *Psychology Crime and Law*, 9, 9-18.
- Nimon, K. (2010). Regression commonality Analysis: Demonstration of an SPSS Solution. *Multiple Linear Regression Viewpoints*, 36, 1, 10-17.
- Perry, C.-L., Kelder, S.-H. & Komro, K.-A., 1993. The social world of adolescents: family, peers, schools, and the community. In S.-G. Millstein, A.-C. Peterson & E.-O. Nigthingale, (Eds.), *Promoting the Health of Adolescents: New Directions for the Twenty-First Century* (pp. 73-96). New York: Oxford University Press.
- Potvin, P., & Paradis, L. (2000). Facteurs de réussite dès le début de l'éducation préscolaire et du primaire. Rapport de recherche, CRIRES, Université Laval, *Études et recherches*, 5, 133 pages.
- Saint Laurent, L. (2000). Les programmes de prévention de l'échec scolaire. In F. Vitaro & C. Gagnon (Eds.), *Prévention des problèmes d'adaptation chez les enfants et les adolescents, Tome 2 : les problèmes externalisés* (pp. 5-54). Québec : Presses Universitaires du Québec.

EVALUER DES COMPETENCES LATENTES PAR L'ANALYSE DES ERREURS : L'EXEMPLE D'EPREUVES DE RAISONNEMENT

Philippe Chartier*

* INETOP/CNAM/CRTD. philippe.chartier@cnam.fr

Mots-clés : erreur, raisonnement, tests, compétence latente, potentiel

Résumé. Dans l'évaluation des compétences l'observation est le plus souvent centrée uniquement sur le niveau de réussite, alors même que la prise en compte d'autres types d'information, comme ici l'analyse des erreurs, est négligée. Dans l'exemple que nous développerons ici -l'évaluation des compétences de raisonnement logique par un test- nous proposons de différencier des types d'erreurs en fonction du degré de proximité de celles-ci avec la bonne réponse attendue. Dans ce cadre nous calculons un score supplémentaire qui vise à rendre compte de compétences latentes, c'est-à-dire de compétences qui ne sont pas obligatoirement prises en compte par le score classique de l'épreuve. Les résultats indiquent bien que des élèves de niveau comparable peuvent être différenciés sur ce score et qu'il est donc possible de compléter l'indicateur classique du niveau de réussite à l'épreuve par cet indice issu de l'analyse des erreurs. L'utilité et la validité de cette approche fait l'objet de la discussion.

Dans l'évaluation des compétences l'observation est trop souvent centrée sur la performance d'un sujet, placé dans une situation donnée, à partir de critères de réussite définis. Pourtant l'analyse des erreurs mériterait une attention particulière car, sous certaines conditions, elle permet d'apporter des éléments d'information sur le fonctionnement cognitif des personnes mais également, et c'est l'objet de notre communication, de rendre compte de capacités latentes qui ne sont pas obligatoirement prises en compte dans le calcul du score à l'épreuve. Nous présenterons quelques références théoriques concernant l'utilité de cette approche avant de l'illustrer à partir de propositions de principes de cotation des réponses dans une épreuve de raisonnement logique.

1. Les approches de l'analyse des erreurs

Cette approche a été surtout développée dans le cadre général de l'évaluation formative / diagnostique et de la didactique des disciplines qui ont donné un nouveau statut à l'erreur : « *on est globalement passé d'une conception négative donnant lieu une sanction à une autre, où les erreurs se présentent plutôt comme des indices pour comprendre le processus d'apprentissage et comme témoins pour repérer les difficultés des élèves* » (Astolfi, 1997, couverture). A partir de cette rupture épistémologique (Favre, 1995) l'erreur n'est plus une faute (Favre, 2010) mais devient un indicateur, une information pertinente, un objet possible d'analyse. Il s'agit par exemple de rendre compte des procédures de résolution utilisées par un élève qui l'ont amené à fournir une réponse fautive, de retrouver les opérations intellectuelles dont elles sont la trace ou encore de renseigner sur le type de représentation mentale de la situation élaboré par une personne (Astolfi, 1997). Cette variable est également pertinente dans une perspective méta-cognitive comme moyen pour faire réfléchir les élèves sur leur propre fonctionnement intellectuel « *l'objectif est alors que l'élève maîtrise ses procédures, qu'il sache faire et dire, qu'il s'interroge sur ce qu'il a fait, comment il a fait pour faire mieux, pour trouver des réponses (domaine d'analyse des stratégies, de la gestion des procédures, réflexivité, métacognition...)*. » (Bayet et Seknadje-Askénazi, 1998, p. 136). L'erreur est ainsi une bonne occasion de «réfléchir, de rectifier

ou d'approfondir ses connaissances, de mieux se connaître soi-même » (Baruk, 1985, p.73). Analyser les erreurs, tenter de les comprendre et de les modéliser, c'est aussi se donner les moyens pour développer des situations de remédiation adaptées (Morissette, 2009).

Sans pouvoir ici être exhaustif sur les auteurs et les approches qui accordent à l'erreur un certain statut, signalons que l'intérêt pour l'analyse des erreurs se retrouve également dans le cadre plus général de la psychologie cognitive (par exemple Fayol, 1995) ou encore en ergonomie (Leplat, 1999).

2. L'analyse des erreurs dans les tests de raisonnement

Reste que cet intérêt semble moins développé pour ce qui relève de l'évaluation de compétences cognitives. En effet dans ce domaine, et tout particulièrement dans celui de l'évaluation des compétences de raisonnement à partir de tests, la pratique la plus fréquente consiste à identifier une (et une seule) bonne réponse pour chaque item, puis à sommer celles-ci afin d'obtenir le score total à l'épreuve. L'évaluation ne repose donc que le nombre d'items réussis sans tenir compte des autres informations présentes dans le protocole. On considère alors, au moins de manière implicite, qu'il n'existe pas de différences interindividuelles, ou du moins de différences interindividuelles méritant d'être prises en compte, autre que le score observé. Pourtant on peut observer que :

1° des sujets obtenant un même score à une épreuve peuvent être différenciés sur les items qu'ils ont réussi (leur profil des bonnes réponses)¹,

2° des sujets échouant à un même item ne fournissent pas obligatoirement la même réponse erronée (le type d'erreur).

Nous ne développerons ici que la seconde forme de différence liée aux erreurs commises (pour une illustration de la première voir Chartier, 2010a). Pour pouvoir interpréter des différences éventuelles à ce niveau il est nécessaire de disposer d'une typologie des réponses. Par exemple, et c'est l'un des rares cas d'analyse de ce type, dans le manuel des matrices de Raven APM² une typologie des erreurs est proposée. Quatre type d'erreurs ont été identifiées (Raven, Court et Raven, 1998, p. 15-19) :

1° *Solution incomplète* (type A) : identification incomplète des règles à prendre en compte,

2° *Mode de raisonnement arbitraire* (type B) : principe de raisonnement qualitativement différent de celui que requiert l'item,

3° *Choix surdéterminé par des éléments intrus* (type C) : solutions « fourre-tout » qui combine souvent le maximum d'éléments disponibles,

4° *Répétitions* (type D) : sélection d'une figure identique à l'une des figures du problème.

Ces type d'erreurs se différencient au niveau de leur qualité : certaines étant proches de la bonne réponse (type A), d'autres plus éloignées (type B ou D). Il peut être alors pertinent de repérer le type d'erreur le plus fréquent chez un sujet donné afin, par exemple, de tenter de comprendre ses logiques de raisonnement, d'identifier les difficultés spécifiques rencontrées. On observe ainsi que la proportion de ces différentes erreurs varie selon le niveau de compétence : les sujets de faible niveau commettent plus d'erreurs de type B alors que ce sont les erreurs de type A qui sont les plus fréquentes chez les sujets de niveau de compétence plus élevé.

Bien que l'intérêt de l'analyse des erreurs sur cette épreuve fasse encore l'objet de débats (voir par exemple Grégoire, 2004) cette approche nous semble intéressante car elle permet de prendre en compte des informations pertinentes concernant les conditions de réalisation d'une performance et met en évidence l'existence de différences individuelles autres que le score à l'épreuve. Il est possible également de mettre en place des pratiques d'évaluation qui complètent une évaluation traditionnelle par un retour réflexif, avec le sujet, sur les items échoués (Barbot, 2010).

¹ Par exemple un total de 20 points est considéré comme équivalent à tout autre total de 20 même si ces points ne sont pas obtenus sur les mêmes items.

² Cette épreuve très connue visant à évaluer le raisonnement logique (de type aptitude générale ou facteur g) consiste à choisir une matrice qui vient compléter une configuration proposée.

Nous trouvons également une analyse des erreurs dans un autre test de raisonnement : le test BLS4 de Bonnardel (Thiébaud, 2000). Sans élaborer de manière explicite une réelle typologie des erreurs l'auteur relève les différentes réponses données par les sujets pour les 16 premiers items de l'épreuve. Mais l'utilisation possible de ces données dans une pratique d'évaluation n'est pas abordée, l'analyse reste assez descriptive et développe de façon rapide la problématique de la variabilité interindividuelle.

Enfin rappelons que la cotation de certains subtests des échelles de Wechsler (2005) peut être interprétée en terme d'analyse de l'erreur avec une approche différente de la notation classique juste (1)/faux (0) en distinguant trois niveaux de qualité de la réponse : la note 1 correspondant à un niveau intermédiaire se situant entre la mauvaise réponse (cotée 0) et la bonne réponse attendue (cotée 2).

On retrouve également cette approche dans certains items de l'enquête PISA pour lesquels il est prévu une cotation selon une modélisation à crédit partiel à trois niveaux : crédit complet, crédit partiel et absence de crédit (Lafontaine & Raïche, 2011).

3. Partie expérimentale

3.1 Présentation du test RCC³

Ce test vise à évaluer les capacités de raisonnement logique à partir d'un support original : des cartes à jouer (Chartier, 2008 et 2010a). La tâche est inspirée du test MGM de Pire (1957) dans lequel le sujet doit découvrir les caractéristiques (famille et valeur) d'une carte retournée qui complète une série proposée. Pour les familles il existe quatre possibilités de réponse (carreau, cœur, trèfle et pique), pour les valeurs dix possibilités (seules les cartes de 1 à 10 sont utilisées). Nous avons repris le principe de cotation du test original : la réponse est considérée comme juste si la personne fournit à la fois la bonne famille et la bonne valeur. En additionnant les items réussis on obtient le score RCC.

3.2 Principe général d'analyse des erreurs au RCC

Ce principe général de cotation peut être discuté car il revient à attribuer la note 0 à toute réponse différente de la bonne réponse attendue, alors même qu'il peut exister des différences de qualité entre les réponses fausses. Ainsi il nous semble pertinent de repérer ici au moins deux types d'erreur :

Type 1 (ou erreur 00) : des réponses entièrement fausses qui ne comportent aucun élément de la bonne réponse (valeur et famille fausses) ;

Type 2 : des réponses fausses qui comportent une partie de la bonne réponse, c'est-à-dire un élément correct parmi les deux attendus (valeur fausse ou famille fausse).

Nous proposons de nommer les réponses de type 2 *Bonnes Réponses Partielles* (B. R. P.) car elles sont plus proches des bonnes réponses attendues que les réponses de type 1 et comportent de fait une facette du raisonnement que l'on cherche à estimer par le test RCC. De plus des études exploratoires ont montré qu'il existait des différences interindividuelles non négligeables sur ces réponses fausses (Chartier, 2010b).

Deux types de B.R.P peuvent être distingués ici :

1/ des réponses qui comportent la bonne valeur, mais avec erreur pour la famille, nous nommerons ces réponses des *Bonnes Réponses Partielles valeur* (B.R.P valeur),

2/ le cas inverse : des réponses qui comportent la bonne famille, mais avec erreur sur la valeur, nous nommerons ces réponses des *Bonnes Réponses Partielles famille* (B.R.P famille).

Dans une approche descriptive nous pouvons repérer, par exemple, quelles sont les types d'erreurs les plus fréquentes pour un individu donné. Mais l'analyse peut se poursuivre par une approche quantitative et le calcul de deux indicateurs :

³ Raisonnement sur Cartes de Chartier.

- le score *Réussite Partielle* (R.P) qui consiste à accorder $\frac{1}{2}$ point pour toute réponse de type B.R.P,
 - le score *RCC Potentiel* qui consiste à ajouter le score RP au score RCC initial.
- Ce score *Potentiel* vise à prendre en compte des capacités/compétences latentes de raisonnement qui n'apparaissent pas dans le calcul de l'indicateur classique RCC.

3.3 Méthodologie

Le test RCC a été passé auprès de 123 élèves de 6^{ème}, scolarisés en France (âge moyen d'environ 11 ans).

3.4 Résultats

Les données descriptives figurent dans le tableau 1. Les valeurs nous indiquent que le test présente des qualités métriques convenables et que son niveau de difficulté convient aux élèves de notre échantillon. Comme attendu les élèves se différencient sur leur score RCC qui témoigne de leur compétence générale de raisonnement logique. Rappelons que ce score RCC ne prend en compte que les réponses complètement correctes (famille *et* valeur justes).

	Nb d'items	moyenne	écart type	mini	maxi	alpha
RCC	40	18,6	6,6	1	33	.87

Tableau 1 : résultats descriptifs du RCC

En reprenant notre cadre d'analyse des erreurs nous pouvons également calculer les taux relatifs des différentes réponses fausses. Les données sont présentées dans le tableau suivant.

	moyenne	mini	maxi	écart type
T1 : erreur 00	82%	50	100	11
T2 : BRP valeur	6%	0	27	5
T2 : BRP famille	12%	0	40	9

Tableau 2 : répartition des réponses fausses

On peut observer que les réponses de type 1 (erreur 00) sont les plus fréquentes, les réponses de type 2 ne représentant que moins de 20% des réponses. Parmi celles-ci les BRP *famille* sont deux fois plus nombreuses que les BRP *valeur*, ce qui n'est pas étonnant car il est statistiquement plus probable de trouver la bonne famille (4 possibilités seulement) que de trouver la bonne valeur (10 possibilités). Mais ce que ce tableau nous apporte c'est l'illustration d'une variabilité importante entre les élèves : certains ne fournissant aucune réponse de type 2 alors que pour d'autres élèves ce type de réponse est très fréquent. On retrouve ici des différences interindividuelles concernant la qualité des réponses erronées.

L'étape suivante consiste à la cotation du score *Réussite Partielle* (RP) en attribuant $\frac{1}{2}$ points à chaque réponse de type 2. En cumulant ce score RP avec le score RCC on obtient alors le score *RCC Potentiel* (voir tableau 3).

	moyenne	écart type	mini	maxi
RP	3,6	2,5	1	14,5
RCC Potentiel	22,2	6,3	6,5	36,5

Tableau 3 : valeurs descriptives des scores RP et RCC *Potentiel*

Les différences interindividuelles apparaissent ici de manière flagrante sur le score RP : les élèves obtiennent en moyenne 3,6 points mais les écarts sont importants et varient de 1 à 14,5 points. En comparant les scores RCC et RCC *Potentiel* (tableaux 1 et 3) on observe que l'augmentation est plus importante pour les scores RCC faibles (passage du score minimum de 1 à 6,5) que pour les scores élevés (passage de 33 à 36,5). Effectivement il est probablement plus facile d'obtenir des demi points supplémentaires lorsque le protocole comporte beaucoup d'items échoués que lorsque le score est proche du maximum (pour un score très élevé au RCC le nombre d'erreur est faible, donc la marge possible de progression est beaucoup plus réduite que dans le cas d'un score très faible). C'est pour cette raison qu'il est utile de comparer les élèves qui ont obtenu le même niveau global de raisonnement. Pour cela nous avons sélectionné les scores RCC les plus fréquents et avons relevé les scores RCC *Potentiel* des élèves concernés (voir tableau 4). La question est la suivante : à niveau équivalent les élèves peuvent-ils être différenciés sur le score *Potentiel* ?

		Scores	RCC	
	14	16	22	23
Effectifs	11	9	13	8
moyenne RP	3,1	4,6	3,3	4,4
score RP mini	.50	1	.50	1
score RP maxi	8	10	7,5	6,5
scores RCC <i>Potentiel</i>	14,5 à 22	17 à 26	22,5 à 29,5	24 à 29,5

Tableau 4 : variabilité des scores RCC *Potentiel* pour quelques scores RCC

Quel que soit le score RCC d'origine (ici 14, 16, 22 ou 23) on observe d'importantes différences inter-élèves sur les scores RCC *Potentiel*. Par exemple pour un score RCC de 16 points le score RCC *Potentiel* varie de 17 à 26 points. Ainsi certains élèves, lorsqu'ils donnent une réponse fausse, elle est souvent de type 1 (ce qui aboutit alors à un score RP faible et aux scores *Potentiel* les plus faibles du groupe) alors que d'autres élèves, de même niveau⁴, fournissent plus fréquemment des réponses fausses de type 2, plus proches des bonnes réponses attendues (d'où un score RP élevé, qui aboutit aux scores RCC *Potentiel* les plus élevés du groupe).

Et ces différences se retrouvent à tous les niveaux de réussite. Autrement dit des élèves considérés comme comparables vis-à-vis de leur compétence générale de raisonnement (scores RCC identiques) peuvent présenter des différences parfois importantes lorsque l'on prend en compte la qualité de leurs réponses fausses.

L'une des questions qui se pose alors est la suivante : pour rendre compte des capacités/compétences des élèves quel est le meilleur indicateur : le score RCC ou le score RCC *Potentiel* ? Cette question se pose de manière encore plus vive dans le cas des élèves qui présentent un score RCC *Potentiel* très supérieur à leur score RCC.

Conclusion

L'analyse des erreurs, largement négligée dans l'évaluation des compétences de raisonnement à l'aide de tests, mérite pourtant toute notre attention car elle peut fournir différentes informations aussi bien sur les conditions de réalisation d'une performance que sur les réelles capacités des sujets. Si, comme nous l'avons illustré, à la suite d'une typologie des erreurs il est possible de différencier celles-ci sur leur degré de proximité avec les bonnes réponses attendues, les différences interindividuelles observées à ce niveau pourraient alors être interprétées comme des

⁴ Même score RCC.

différences de compétences *latentes*, c'est-à-dire de compétences qui ne permettent pas (ou qui ne permettent pas encore) de donner la bonne réponse mais qui permettent au sujet de fournir une réponse de qualité supérieure aux autres erreurs possibles. Ces compétences ne sont pas obligatoirement prises en compte dans l'indicateur classique des tests : le score total. Pour cette raison nous avons proposé le principe de calcul de deux indicateurs 1° le score de Réussite Partielle (RP) qui consiste à accorder ½ points aux réponses fausses considérées comme étant de qualité supérieures aux autres possibilités de réponse, 2° le score RCC *Potentiel*, qui consiste à rajouter ce score RP au score initial RCC.

Ces scores sont complémentaires au score RCC et visent à fournir des informations spécifiques concernant l'existence de compétence latente chez les sujets. Nous pourrions émettre l'hypothèse suivante : le score classique RCC évaluerait la *performance* tandis que le score RCC *Potentiel* évaluerait plutôt la *compétence* (référence ici à la distinction compétence/performance proposée par Chomsky dans le cadre de la linguistique). L'interprétation des résultats d'une personne pourrait alors reposer bien entendu sur le niveau global de performance (le score RCC) mais aussi sur la différence entre ces deux indicateurs à partir de références normées (de type étalonnage) pour ces deux scores.

Nous pouvons rapprocher notre approche du courant de l'évaluation dynamique des aptitudes qui vise également, mais à partir d'un cadre théorique différent et de pratiques spécifiques, à rendre compte d'une même réalité, qui n'apparaît pas obligatoirement totalement dans les procédures classiques d'évaluation : les potentialités des personnes (Chartier et Loarer, 2008).

Même si les recherches sont à poursuivre pour conforter la signification, la validité et les utilisations possibles de ce type d'indicateur, nous considérons que l'analyse des erreurs représente des perspectives pertinentes aussi bien pour la recherche que pour la pratique.

Enfin, l'approche que nous proposons peut être généralisée à tout test existant, comme par exemple les matrices de Raven, les tests de type dominos etc., mais elle permet également de proposer un cadre de référence pour l'élaboration de nouvelles épreuves à visée diagnostique permettant de recueillir une pluralité d'informations sur la réalisation d'une performance et ainsi pouvoir en inférer de manière plus précise la compétence.

Références

- Astolfi, J-P. (1997). *L'erreur, un outil pour enseigner*. Paris : ESF.
- Barbot, H. (2010). *Evaluation psychométrique des élèves en grande difficulté scolaire pour un accompagnement et un conseil en orientation*. Communication au Colloque International L'accompagnement à l'orientation aux différents âges de la vie, INETOP/CNAM, Paris, 17-19 mars.
- Baruk, S. (1985). *L'âge du capitaine. De l'erreur en mathématiques*. Paris : Seuil.
- Bayet, B. & Seknadje-Askénazi, J. (1998). Les enjeux de l'évaluation. *La nouvelle revue de l'AIS*, 4, 131-136.
- Chartier, P. (2010a). Prise en compte de différentes formes de variabilité interindividuelle dans des tests de type facteur g. In A. De Ribaupierre, P. Ghisletta, T. Lecerf et J-L. Roulin (Ed.), *Identité et spécificité de la psychologie différentielle*, (pp. 81-86). Rennes : Presses Universitaires de Rennes.
- Chartier, P. (2010b). Analyse des profils des réponses et des erreurs dans les tests d'aptitudes : principes et illustrations. Colloque International *L'accompagnement à l'orientation aux différents âges de la vie*, CNAM/INETOP Paris, 17-19 mars.
- Chartier, P. (2008). Expérimentation d'une épreuve de facteur g utilisant comme support des cartes à jouer. In E. Loarer, J-L. Mogenet, F. Cuisinier, H. Gottesdiener, P. Mallet et P. Vignaud (Ed.), *Perspectives différentielles en psychologie* (pp. 39-42). Rennes : Presses Universitaires de Rennes.
- Chartier, P. & Loarer, E. (2008). *Evaluer l'intelligence logique*. Paris : DUNOD.
- Favre, D. (2010). *Cessons de démotiver les élèves*. Paris : DUNOD.

- Favre, D. (1995). Conception de l'erreur et rupture épistémologique. *Revue Française de Pédagogie*, 111, 85-94.
- Fayol, M. (1995). La notion d'erreur, éléments pour une approche cognitive. In G. Blanchet, J. Raffier et R. Voyazopoulos (Ed.), *Intelligence, scolarité et réussite* (pp. 137-152). Paris : La pensée sauvage.
- Grégoire, J. (2004). *L'examen clinique de l'intelligence de l'adulte*. Sprimont : Mardaga.
- Lafontaine, D. & Raïche, G. (2011). Principes méthodologiques et techniques des enquêtes internationales, *Mesure et évaluation en éducation*, 34, 2, 25-55.
- Leplat, J. (1999). Analyse cognitive de l'erreur. *Revue Européenne de Psychologie Appliquée*, 49, 1, 31-41.
- Morissette, J. (2009). La portée d'une perspective socioculturelle de l'évaluation formative : vers l'élargissement d'une conceptualisation. *Mesure et évaluation en éducation*, 32, 2, 1-27.
- Pire, G. (1957). *Test MGM. Manuel d'instructions*. Paris : Editions scientifiques et psychotechniques.
- Raven, J-C., Court, J-H. & Raven, J. (1998). *Manuel des Raven : section 4*. Paris : EAP
- Thiébaud, E. (2000). *Manuel du test BLS4*. Paris : EAP.
- Wechsler, D. (2005). *WISC-IV. Manuel d'administration et de cotation*. Paris : ECPA.

La face cachée des compétences en lecture

Patricia Schillings, Université de Liège – aSPe, patricia.schillings@ulg.ac.be

France Neuberg, Université de Liège – aSPe, france.neuberg@ulg.ac.be

Ariane Baye, Université de Liège – aSPe, ariane.baye@ulg.ac.be

Mots-clés : autoévaluation– lecture- compétence.

Résumé. *Un modèle de compétence en lecture intégrant des aspects motivationnels (perception de ses difficultés et progrès) et une mesure d'attitudes envers la lecture a été testé dans le cadre des évaluations externes en Belgique francophone. Validé auprès de deux échantillons contrastés d'adolescents, ce modèle confirme le rôle déterminant des perceptions de soi comme lecteur sur le développement des attitudes envers la lecture. Ces dernières constituent quant à elles une charnière en lien direct avec les performances. Ces résultats amènent à proposer des pistes et outils didactiques incluant un travail sur l'efficacité de ses processus de lecteur et une analyse réflexive de ce qui rend un texte difficile pour un lecteur particulier.*

Pour la plupart des auteurs, la compétence suppose un agir efficace en situation de résolution de problème. Identifier et mobiliser les ressources nécessaires à la résolution d'une tâche complexe et inédite suppose chez l'apprenant une capacité de réflexion sur l'action qui prend pour objet à la fois les caractéristiques de la tâche et l'évaluation de la qualité de ses propres démarches. La définition de Beckers insiste bien sur le caractère subjectif de la difficulté de la tâche « la capacité d'un sujet à mobiliser, de manière intégrée, des ressources internes (savoirs, savoir-faire et attitudes) et externes pour faire face efficacement à une famille de tâches complexes pour lui (2002, p. 57).

Les modèles d'expertise en lecture soulignent également le lien incontournable entre la perception de la difficulté de la tâche et la mise en œuvre de stratégies de lecture : un lecteur compétent serait capable d'adapter sa lecture à la difficulté du texte en passant d'un mode de lecture automatisé (caractérisé par un usage non conscient des procédures mises en œuvre) à un mode contrôlé, caractérisé par un usage conscient et volontaire de stratégies de lecture (Afflerbach, Pearson, & Paris, 2008).

Mais la perception de la difficulté de la tâche suffit-elle à déclencher chez l'élève la mobilisation de ses ressources cognitives ?

Dans les approches théoriques de la motivation pour la lecture, de même que dans les modèles d'apprentissage autorégulé, l'efficacité perçue joue un rôle central dans le développement de la compétence. Wigfield et Eccles (1992, 2000) insistent sur le rôle de variables telles que le sentiment d'efficacité personnelle par rapport à une tâche précise (Suis-je capable de réussir cette tâche ?) ou la valeur accordée à cette même tâche. Cette seconde variable serait affectée par l'intérêt porté par l'apprenant pour la tâche, l'importance qu'il accorde au fait de la mener à bien, la mesure dans laquelle il perçoit l'utilité de la tâche et le coût estimé en termes d'efforts cognitifs à fournir. Pour Guthrie et Wigfield (2000), ces variables motivationnelles sont d'autant plus importantes à prendre en compte qu'elles conditionnent chez le lecteur la mise en œuvre de ses stratégies de lecture.

Comment dès lors envisager l'évaluation à large échelle des compétences des élèves ? Quelles variables évaluer si l'on se place dans une logique diagnostique ? Quels aspects de la compétence mettre en évidence dans les documents didactiques envoyés aux enseignants dans le but de réguler des apprentissages ?

Les évaluations externes non certificatives en lecture et production d'écrit menées en novembre 2010 auprès des élèves de 5^e secondaire de la Communauté française de Belgique ont permis le recueil de données ciblées sur les aspects cognitifs et motivationnels des adolescents. Deux échantillons représentatifs sont constitués au départ des deux formes d'enseignement (transition et qualification).

Trois dimensions ont été évaluées au sein d'un questionnaire adressé aux élèves après la passation de l'épreuve de lecture :

- *Les attitudes envers la lecture* : cette dimension composée de dix items présente un indice de consistance (alpha de Cronbach) de .90.
- *Les difficultés perçues* en lien avec la lecture : cette dimension amène l'élève à se positionner face à des obstacles tels que la déconcentration, la mémoire ou la confusion entre les personnages. Elle est composée de cinq items et présente un indice de consistance interne de .76.
- *Les progrès perçus* visent le sentiment d'avoir effectué des progrès en lecture depuis la deuxième année du secondaire. Cette dimension est composée 5 items. L'alpha de Cronbach s'élève à .70.

La mise en lien de ces trois variables avec les performances des élèves permet de souligner l'importance de mettre au point dans les classes des situations d'évaluation susceptibles de donner à voir aux élèves des indices de leur progression. Les résultats confortent également le rôle prépondérant des activités métacognitives axées sur l'autoévaluation de la maîtrise des stratégies de lecture y compris au sein des filières d'enseignement dites qualifiantes. Enfin, les résultats montrent une fois de plus l'importance de maintenir effectif jusqu'en fin d'enseignement obligatoire un curriculum d'enseignement des stratégies de lecture.

Afflerbach, P., Pearson, D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The reading teacher*, 61,(5), 364-373.

Beckers, J. (2002). Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité. Bruxelles : Labor.

Guthrie, J. T.; & Wigfield, A. (2000). Engagement and motivation in reading. In Kamil, M.L. and Mosenthal, P.B. (Eds), *Handbook of reading research*, Volume 3 (pp. 403-422),. Mahwah : Erlbaum.

Wigfield A., & Eccles, J.(2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68-81.

Wigfield, A., & Eccles, J.(1992). The development of achievement task value : A theoretical analysis. *Developmental review*, 12, 265-310.

QUELLE ARTICULATION ENTRE LES COMPETENCES PAR DOMAINE ET PROFILS DE COMPETENCES : L'EXEMPLE DE PISA EN SUISSE

Jean Moreau*, Christian Nidegger**

* Unité de recherche pour le pilotage des systèmes pédagogiques, Jean.moreau@unil.ch

** Service de la recherche en éducation, christian.nidegger@etat.ge.ch

Mots-clés : *profils de compétences, évaluation, comparaison*

Résumé. *Les enquêtes à large échelle PISA ont comme objectif d'évaluer les élèves de 15 ans et de mettre en relation les performances ainsi observées avec les caractéristiques des élèves ainsi qu'un certain de variables permettant de cerner le contexte scolaire. Les compétences évaluées relèvent de trois domaines: la lecture, les mathématiques et les sciences. Cependant on trouve peu d'analyses qui prennent en compte les trois domaines de compétences en même temps en établissant ainsi des profils de compétences. La prise en compte de tels profils d'élèves permet une meilleure comparaison de l'efficacité des différents systèmes scolaires. Ainsi on tentera de répondre à un certain nombre de questions. Les élèves ont-ils le même niveau de compétences dans chaque domaine ? L'influence des caractéristiques des élèves est-il le même si on prend en compte les trois domaines de compétences ? Les profils de compétences sont-ils influencés par des dimensions culturelles ou les caractéristiques des systèmes scolaires ?*

1. Introduction

Les enquêtes à large échelle PISA, initiées par l'OCDE, sont devenues un élément incontournable du paysage éducatif suisse. Ces enquêtes ont comme objectif d'évaluer les élèves de 15 ans et de mettre en relation les performances ainsi observées avec les caractéristiques des élèves par exemple le niveau socioéconomique, l'origine de la famille, la langue parlée à la maison, le genre de l'élève ainsi qu'un certain de variables permettant de cerner le contexte scolaire. Cette évaluation vise à cerner les compétences des élèves qui leur permettent de maîtriser les situations de la vie quotidienne et de répondre aux défis de leur vie future (OCDE 2010). Les compétences évaluées relèvent de trois domaines qui sont le plus souvent testés dans les épreuves à large échelle : la lecture, les mathématiques et les sciences. Les publications de l'OCDE fournissent de nombreuses comparaisons et analyses sur le lien entre les compétences d'un des trois domaines et les différentes variables disponibles. Cependant on trouve peu d'analyses qui prennent en compte les trois domaines de compétences en même temps en établissant ainsi des profils de compétences. La prise en compte de tels profils d'élèves permet une meilleure comparaison de l'efficacité des différents systèmes scolaires. Le concept de littératie décliné dans chacun des trois domaines ne permet pas de réduire la notion de compétence à une seule dimension qui serait transversale et pourrait être mesurée indépendamment d'un domaine de compétence. En effet certains élèves peuvent avoir une certaine « maîtrise » des trois domaines alors que d'autres élèves ne maîtrisent que deux domaines seulement, voire un seul ou aucun. Par ailleurs on peut se poser un certain nombre de questions. Les élèves ont-ils le même niveau de compétences dans chaque domaine ? L'influence des caractéristiques des élèves est-il le même si on prend en compte les trois domaines de compétences ? Les profils de compétences sont-ils influencés par des dimensions culturelles ou les caractéristiques des systèmes scolaires ?

2. Résultats par domaine

Les comparaisons internationales ont porté sur les élèves de 15 ans, mais la Suisse et certains cantons (dont l'ensemble des cantons romands) ont étendu leur échantillon afin de pouvoir mener des analyses sur les élèves de 9^e année, c'est-à-dire auprès des élèves qui sont à la fin de leur scolarité obligatoire, indépendamment de leur âge. Les résultats qui suivent concernent les élèves de 9^e année.

Les résultats globaux ont été le plus souvent centrés sur la moyenne des performances des élèves, domaine par domaine. Cette position centrale ne donne que peu d'indications sur l'ensemble de la distribution des scores des élèves. Une autre option est de considérer les taux d'élèves les plus performants et d'élèves les moins performants. En effet, le fait de savoir si un système scolaire génère, par exemple, peu d'élèves en difficulté ou encore peu d'élèves brillants permet de mieux apprécier son efficacité. On constate que ces taux varient beaucoup entre les cantons (figure 1). Par exemple, le pourcentage d'élèves faibles en mathématiques est d'environ 5% à Fribourg (FR), alors qu'il atteint près de 20% à Zurich (ZH). A l'intérieur d'un même canton, on peut également relever des différences importantes entre domaines. Par exemple, dans le canton du Jura (JU), le taux d'élèves faibles en mathématiques est de 4.2%, alors qu'il est de 11.9% en littérature.

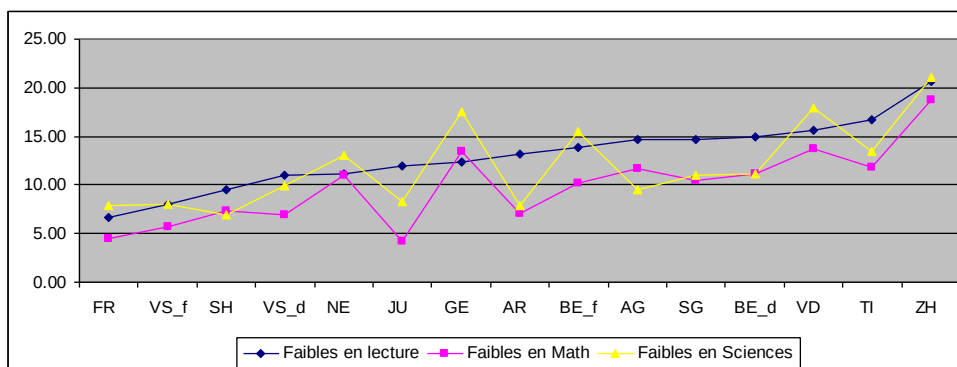


Figure 1: Proportion d'élèves faibles (niveaux <2) dans chacun des trois domaines pour les cantons Suisses

3. Profils de compétences

Les taux d'élèves faibles ou performants dans un seul domaine ne donnent qu'une vision incomplète de l'efficacité d'un système scolaire. Les différences sont, en effet, importantes entre domaines de compétence. Pour obtenir une appréciation plus synthétique de la performance des systèmes, il convient de considérer simultanément les différents domaines. Pour ce faire, on cherche à établir des profils de compétence. Pour déterminer ces profils, on considère les niveaux de compétence des élèves dans chacun des domaines. On peut attribuer à chaque élève le profil de ses niveaux en mathématiques, lecture et sciences. Il y a 294 profils possibles correspondant aux combinaisons des 7 niveaux de mathématiques et sciences (de <1 à 6) et 8 niveaux de lecture (<1, 1b, 1a jusqu'à 6). En fait, on associe, les niveaux plausibles de compétence correspondant pour obtenir cinq profils de compétence plausibles par élève.

Il est intéressant de considérer d'une part les élèves ayant des performances très faibles dans les trois domaines (niveau inférieur ou égal à 1) et d'autre part les meilleurs élèves obtenant dans chaque domaine les meilleurs scores (niveau 5 et 6), les proportions de ces différents types d'élèves étant révélatrices de la qualité du système scolaire dans son ensemble. On constate que l'évaluation de la performance par la considération de 3 domaines (3 niveaux faibles ou 3 niveaux

forts) se distingue nettement d'une évaluation déterminée par un seul domaine (figure 3). Par exemple le Jura présente 2% d'élèves faibles dans les 3 domaines et 11.9% en lecture ; Le Tessin (TI), 5.1% d'élèves faibles dans les 3 domaines et 16.7% en lecture.

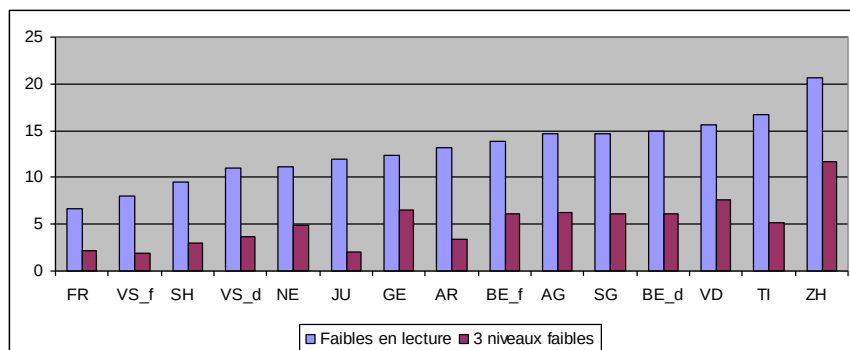


Figure 2 : Proportion d'élèves faibles (niveaux <2) en lecture et dans les trois domaines pour les cantons Suisses

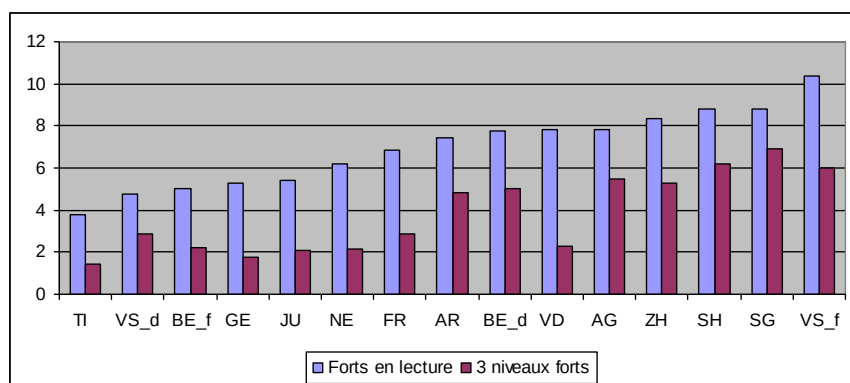


Figure 3 : Proportion d'élèves forts (niveaux 5 et 6) en lecture et dans les trois domaines pour les cantons Suisses

En Suisse, les résultats de PISA 2009 mettent en évidence des différences d'une certaine ampleur entre les cantons (figure 4). Les taux d'élèves très faibles dans les trois domaines peuvent atteindre plus de 11% (Zurich). Celui des très bons élèves (dans les trois domaines) près de 7% de à Saint-Gall (SG). Les cantons romands se signalent par des taux plus faibles de très bons élèves (moins de 3%). Ce taux pouvant atteindre 1.4% au Tessin. Ces cantons se distinguent entre eux, avant tout, par leurs taux d'élèves très faibles. Le canton de Zurich occupe toujours une position particulière avec à la fois beaucoup d'élèves très faibles et beaucoup d'élèves très brillants. Les cantons comportant des grandes villes (Genève, Zurich) recèlent un grand nombre d'élèves faibles (migration, niveau socioéconomique faible) mais elles devraient aussi être le lieu d'un grand nombre d'élèves forts (cf. Zurich) en raison de la concentration des classes aisées. On pourrait donc faire l'hypothèse que les cantons se distinguent aussi bien en fonction de leur appartenance linguistique que des caractéristiques de leurs populations notamment en termes de degré d'urbanisation.

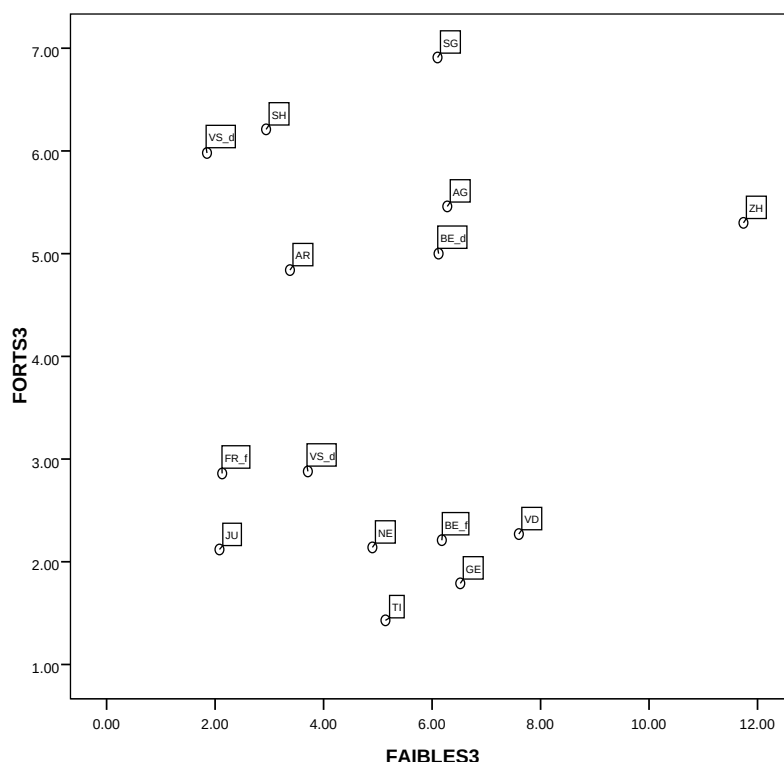


Figure 4 : Proportion d'élèves faibles (3 niveaux <2) et d'élèves forts (3 niveaux 5 ou 6) dans les trois domaines pour les cantons Suisses

3.1 Evolution des profils

La définition des niveaux de compétence étant identique pour les enquêtes PISA 2006 et 2009, nous sommes en mesure de comparer les profils de performances des cantons suisses pour ces deux évaluations. La figure 5 présente la position des cantons Suisses en 2006 et 2009 en fonction de leurs taux d'élèves faibles et forts dans les trois domaines.

On constate, dans la plupart des cantons alémaniques une tendance à l'augmentation des élèves forts et faibles sauf à Schaffhouse (SH) où on observe une diminution des élèves forts. En Suisse romande en général, on relève une évolution moins importante se caractérisant surtout par la diminution des élèves faibles sauf dans le canton de Vaud (VD) où on observe une légère augmentation des élèves faibles et dans le Valais francophone (VS-f) une augmentation des élèves forts. Globalement, il est difficile d'observer des tendances générales entre les cantons. On pourrait faire l'hypothèse que les différents contextes locaux ont un impact non négligeable sur l'évolution des profils entre 2006 et 2009.

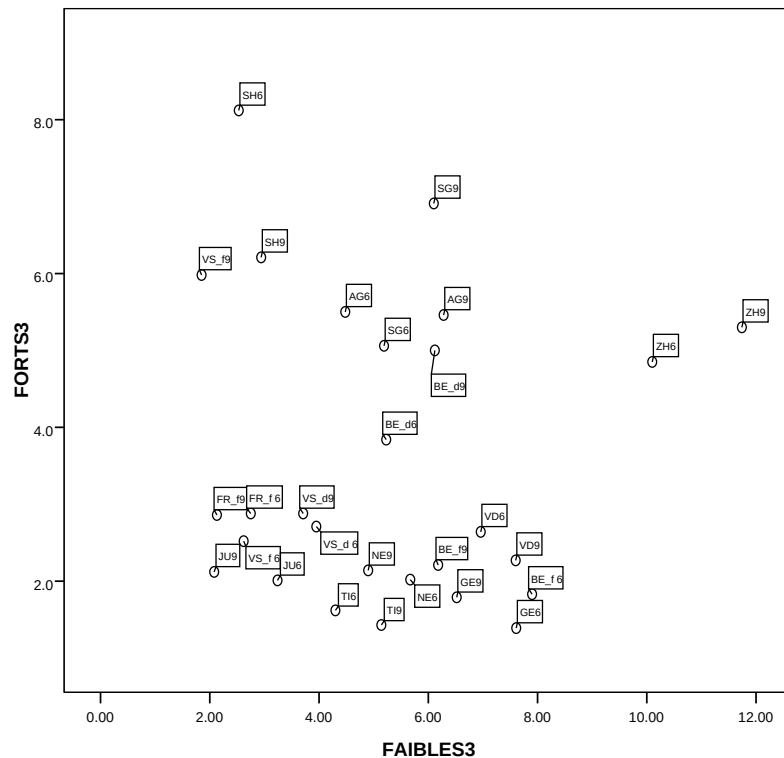
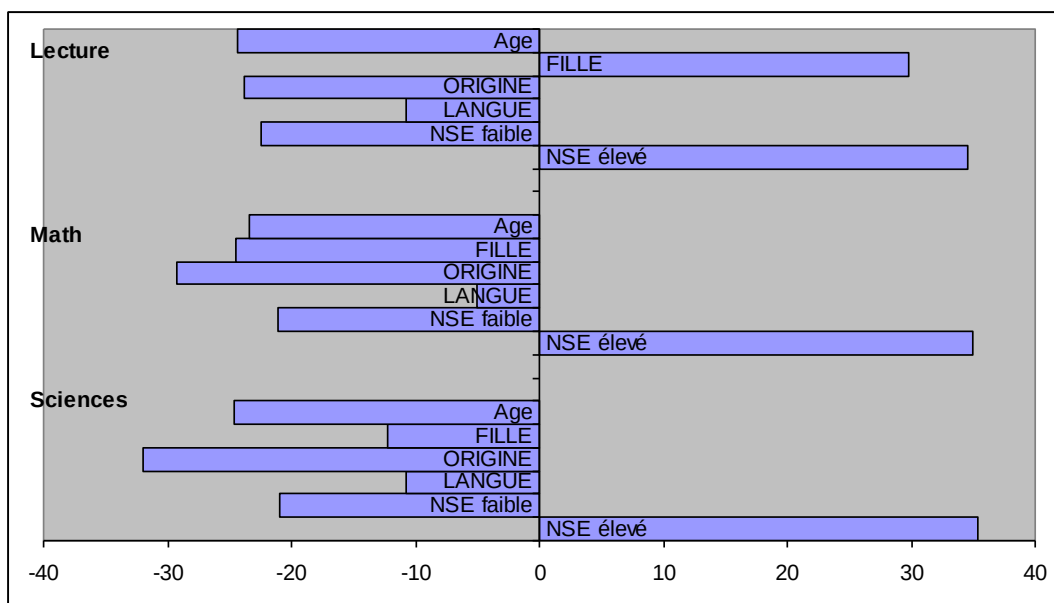


Figure 5 : Proportion d'élèves faibles (3 niveaux <2) et d'élèves forts (3 niveaux 5 ou 6) dans les trois domaines pour les cantons Suisses pour les enquêtes PISA 2006 et 2009

4. Effets des caractéristiques individuelles sur les performances

Il est possible d'évaluer l'effet des caractéristiques individuelles (genre, origine, niveau socioéconomique) sur les résultats des élèves dans chacun des domaines investigués. Il est intéressant de préciser en quoi l'introduction du concept de profil de performance permet d'apprécier ces effets sur l'ensemble des performances des élèves.

On rappelle (figure 6), que ces différentes caractéristiques ont un effet marqué sur les performances dans tous les domaines (Nidegger, 2011). Le niveau socioculturel de la famille a l'influence la plus importante sur les performances. En ce qui concerne le genre des élèves : Les garçons obtiennent de meilleures performances en sciences et surtout en mathématiques et les filles en lecture. Les élèves plus âgés ont de moins bonnes performances que les autres dans tous les domaines. Les habitudes linguistiques (la langue parlée à la maison) ont une influence importante sur les performances en sciences et en lecture mais faible sur les compétences en mathématiques. L'origine de la famille (ne pas être né en suisse) peut également préjudicier les élèves dans tous les domaines et particulièrement en sciences.



Les barres du graphique indiquent, pour chaque domaine, la différence moyenne de performance selon les caractéristiques personnelles suivantes: l'âge (une année de plus), le genre (fille), l'origine de l'élève (pas né en Suisse), la langue parlée à la maison (ne parle pas la langue du test), le niveau socioculturel (faible et élevé). Ces différences moyennes sont déterminées par rapport au score moyen d'un garçon d'âge moyen, né en Suisse, de niveau socioculturel médian (second et troisième quartile) et parlant la langue du test à la maison.

Figure 6 : effet des caractéristiques individuelles des élèves sur les performances des élèves dans chacun des domaines

5. Profils et caractéristiques individuelles

On constate que les effets de ces différentes caractéristiques ne sont pas les mêmes selon les domaines. C'est particulièrement le cas pour le genre, Les filles étant plus performantes en lecture et les garçons en mathématiques. Il est intéressant de considérer simultanément les domaines pour apprécier globalement l'influence des différentes caractéristiques des élèves. Les effets peuvent également différer selon les cantons (Nidegger, 2011), on précisera donc chacune des situations cantonales.

5.1 Le genre

On détermine, pour chaque canton, le pourcentage de filles et de garçons présentant des niveaux faibles (<2) ou des niveaux forts (5 et 6). La figure 7 présente les diverses situations cantonales. Chaque canton est représenté par le profil des filles et celui des garçons. On peut apprécier à l'intérieur d'un même canton, les différences entre les profils des filles et des garçons. Certains cantons présentent peu de différence entre filles et garçons sur la proportion d'élèves faibles et forts (FR, VD). Lorsque les différences sont plus importantes, On constate que les profils des filles présentent souvent des taux plus importants de niveaux forts que ceux des garçons (ZH et VS-f). A l'opposé, à GE on relève des taux plus faibles de filles avec des niveaux faibles.

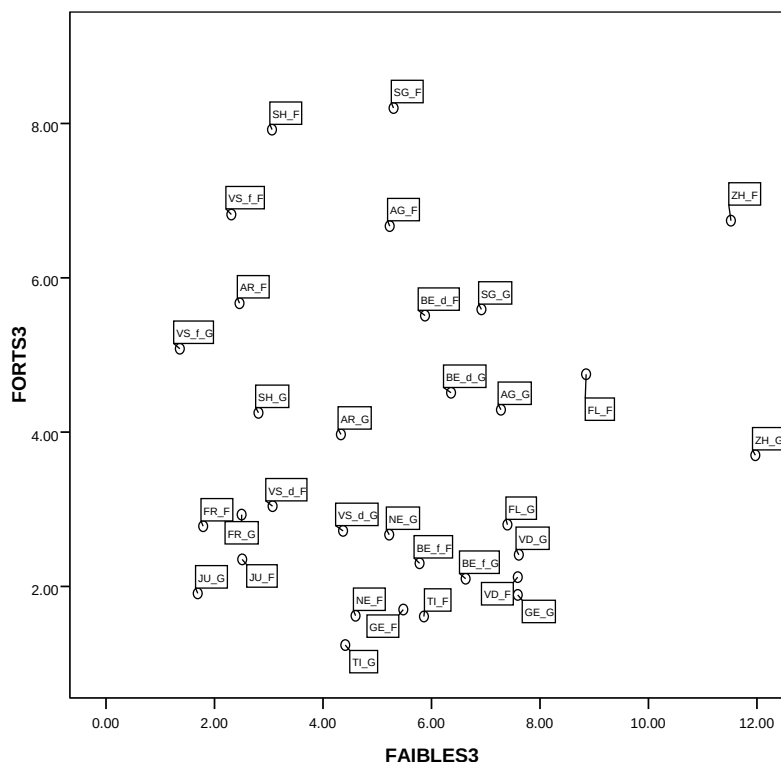


Figure 7 : Proportion d'élèves faibles (3 niveaux <2) et d'élèves forts (3 niveaux 5 ou 6) dans les trois domaines pour les cantons Suisses selon le genre

5.2 L'origine

On détermine, pour chaque canton, le pourcentage d'élèves nés en Suisse et celui d'élèves nés à l'étranger, présentant des niveaux faibles (<2) ou des niveaux forts (5 et 6). La figure 8 présente pour chaque canton les profils de ces différentes populations d'élèves : Natifs (N) et non Natifs (NN). Cette représentation permet d'apprécier d'une part, à l'intérieur d'un même canton les différences de profils entre ces deux catégories d'élèves et d'autre part les différences entre les situations cantonales. A l'intérieur d'un canton, les différences peuvent concerner à la fois le taux d'élèves faibles et le taux d'élèves forts (ZH) ou plutôt le taux d'élèves forts (FR). Les profils des élèves natifs romands apparaissent comme relativement proches.

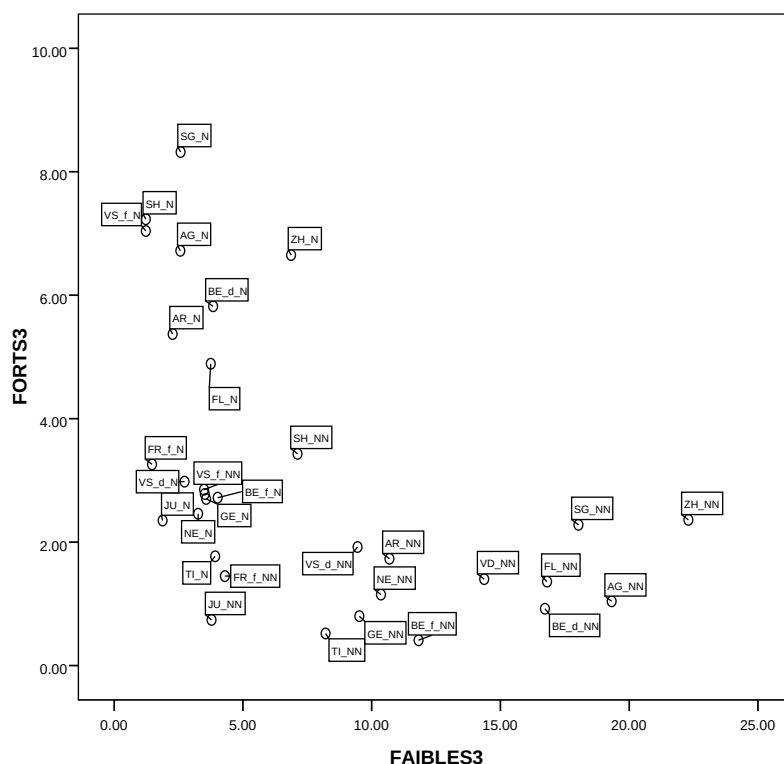


Figure 8 : Proportion d'élèves faibles (3 niveaux <2) et d'élèves forts (3 niveaux 5 ou 6) dans les trois domaines pour les cantons Suisses selon l'origine

5.3 Le niveau socioéconomique de la famille

On considère dans chaque canton deux populations d'élèves selon le niveau socioéconomique de leur famille. Les élèves les plus défavorisés (quartile inférieur de l'indice socioéconomique NSE1) et les autres élèves. La figure 9 présente pour chaque canton les profils de ces différentes populations d'élèves selon les pourcentages de niveaux faibles et de niveaux forts. Les différences sont marquées entre les cantons et à l'intérieur des cantons. Cette analyse va dans le même sens que l'analyse précédente et indique que les différences peuvent également se retrouver à l'intérieur d'une même région linguistique. Toutefois la tendance générale observée va d'une augmentation de la proportion des élèves forts et d'une diminution de la proportion des élèves faibles chez les élèves qui ne sont pas issus des milieux socioéconomiques défavorisés. Cette tendance se retrouve également pour la comparaison entre élèves natifs et non-natifs.

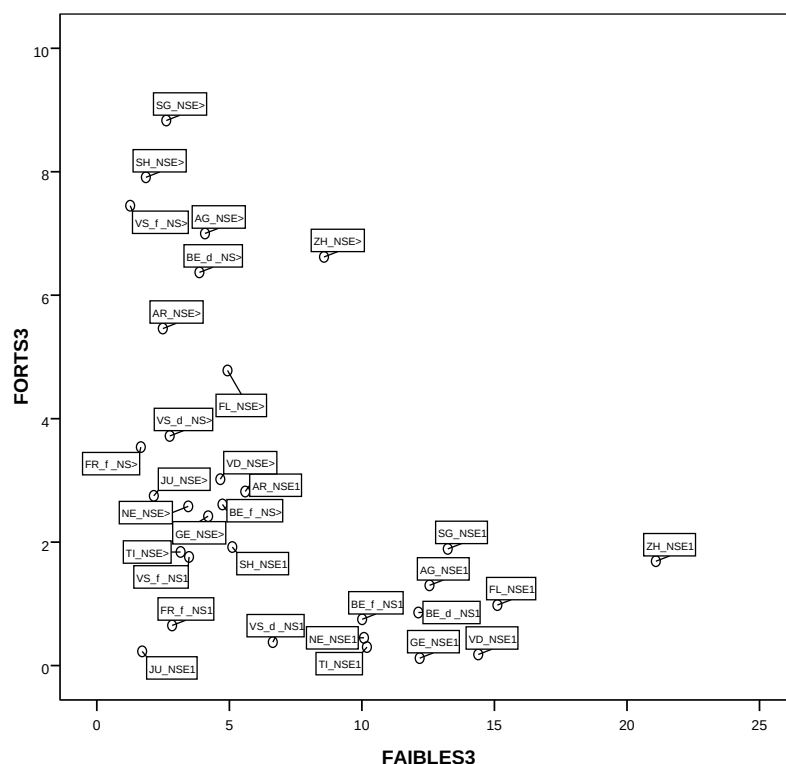


Figure 9 : Proportion d'élèves faibles (niveaux <2) et d'élèves forts (niveaux 5 et 6) dans les trois domaines pour les cantons Suisses selon le niveau socioéconomique de la famille

6. Discussion

La prise en compte simultanée des trois domaines de compétences met en évidence des situations différentes de celles obtenues par des analyses séparées par domaine. De plus, cette vision globale correspond plus à la situation que l'on rencontre dans les différents systèmes où l'on attend des compétences des élèves sur les différents aspects qui font l'objet d'enseignement. Ceci se manifeste par le fait que la réussite des élèves est jugée sur la maîtrise de compétences relevant de plusieurs domaines même si, en s'appuyant sur les données PISA, on ne peut prendre en considération que les trois domaines qui font l'objet de l'évaluation alors que les systèmes scolaires visent l'acquisition d'une palette plus grande de compétences.

Par ailleurs, la mise en relation des profils avec les caractéristiques individuelles des élèves donne une image un peu différente de l'effet de ces caractéristiques. Par exemple la différence de proportion de filles et de garçons dans les niveaux forts ou faible varient selon les cantons alors la variation des élèves forts ou faibles en fonction du niveau socioéconomique va dans le même sens, mais avec une intensité différente, dans tous les cantons: plus élèves sont favorisés plus ils ont tendance à être nombreux dans les niveaux forts et moins représentés dans les niveaux faibles.

7. Références bibliographiques

Moreau, J. (2007). Compétence et contexte des élèves vaudois lors de l'enquête PISA 2003, comparaisons entre cantons, filières et types d'élèves. Lausanne : URSP

Moreau, J. (à paraître). PISA 2006 : Vaud à la mesure des cantons suisses. Lausanne : URSP

Nidegger, Ch. (coord.) (2011). PISA 2006: Compétences des jeunes romands: Résultats de la quatrième enquête PISA auprès des élèves de 9e année. Neuchâtel : IRDP.

Nidegger, Ch. (coord.) (2008). PISA 2006: Compétences des jeunes romands: Résultats de la troisième enquête PISA auprès des élèves de 9e année. Neuchâtel : IRDP.

OECD (2010). Learning for tomorrow's world. First results from PISA 2009. Paris : OECD.

DES OUTILS D'ÉVALUATION DES COMPÉTENCES EN LANGUE FRANÇAISE DES ENSEIGNANTS DU PRIMAIRE

Wajiha Smaili

Université Libanaise, faculté de pédagogie wsmaili@cnam.fr

Mots-clés : *Évaluation en langue, FOS, dispositif de formation, tâche, compétence.*

Résumé. Dans cet article, nous présentons des outils d'évaluation préparés dans le cadre d'un module de formation pour enseigner la langue française à des enseignants de mathématiques, sciences et français au primaire à l'université libanaise, faculté de pédagogie. Les outils et instruments existants pour évaluer les compétences en langue de nos étudiants tel que les examens du DELF/DALF sont inadaptés à notre dispositif. Le programme de formation était préparé en se basant sur un référentiel conçu pour les besoins spécifiques langagiers des enseignants, il était nécessaire d'adopter une évaluation qui tient en compte des compétences professionnelles de notre public. Pour préparer les outils en question, nous sommes partis des tâches fixées dans le programme, nous avons déterminé les composantes langagières : pragmatique, linguistique et sociolinguistique de chacune, préparer des activités, des critères et des éléments observables ainsi qu'un barème.

1. Contexte du projet

Au Liban le français a le statut de langue seconde, il est enseigné dès l'école maternelle en parallèle avec la langue maternelle dans plus de la moitié des institutions scolaires, il est aussi la langue d'enseignement des disciplines scientifiques (mathématiques, sciences de la vie, physique et chimie) dans le primaire, le collège et le lycée.

Dans ce contexte, le bilinguisme instauré à l'entrée de l'école réussit parfaitement à former des enfants bilingues quand les enseignants sont capables d'utiliser la langue française dans leurs pratiques langagières professionnelles.

Cependant, beaucoup d'enseignants utilisent l'arabe comme langue d'enseignement au lieu du français alors que les manuels et les examens sont en français. Ces pratiques langagières affectent la qualité de l'enseignement : les élèves par faute de compréhension des consignes, faillent aux examens officiels. Parfois, cet échec est dû au manque d'outils linguistiques pour s'exprimer et comprendre les consignes et les énoncés.

La faculté de pédagogie de l'université libanaise est la première institution du pays pour assurer des formations initiales et continues aux maîtres des écoles. Alertée par le problème du niveau en langue des enseignants, les responsables ont introduit dans le cursus des formations un module de français intitulé « français de l'enseignement ». Ce module vise à compléter la formation linguistique pour amener les apprenants (étudiants ou enseignants en exercice) à maîtriser la langue française utilisée pour enseigner.

Dans ce cadre, un programme et un contenu ont été préparés pour ce module. La préparation s'est basée sur une analyse des besoins langagiers des apprenants. Les compétences et tâches langagières à utiliser par le public en question ont été répertoriées, ce qui a permis de constituer un référentiel de formation. Ce référentiel a servi de base pour concevoir des unités didactiques sur

objectifs sp  cifiques pour d  velopper les comp  tences langagi  res n  cessaires dans la formation universitaire comme celles importantes pour l'exercice de la profession de l'enseignement (Smaili : 2012).

Suite    l'application des contenus pr  par  s du module, la question de l'  valuation des acquis s'est impos  e. Il n'existe pas d'outils d'  valuation adapt  s au programme en question, ceux-ci sont    concevoir. C'est le travail qui a   t   men   et qui sera pr  sent   dans le cadre de cette   tude.

Notre travail se situe alors dans un domaine pr  cis : l'  valuation des comp  tences en langue. Il s'appuie sur un cadre conceptuel et r  f  rentiel d'o   il tire ses principes, les d  marches de sa conception.

2. Cadre conceptuel

Une   valuation des acquis des apprenants dans une discipline donn  e part toujours d'une conception de la discipline et des m  thodologies d'enseignement/apprentissage adopt  es. De nos jours, Le langage est consid  r   comme un outil de communication en r  f  rence    des comportements d'  changes observables entre divers intervenants (Hymes : 1972), c'est un moyen de communication permettant de v  hiculer des id  es et d'interagir avec les autres. Dans cette perspective, la didactique des langues vise en priorit   le d  veloppement de la communication qui met l'accent dans l'enseignement sur les situations authentiques de communication en d  veloppant les habilit  s linguistiques dans un contexte socioculturel (Lussier : 1992).

Une r  f  rence de base est incontournable actuellement en didactique et en   valuation des langues, c'est le cadre europ  en commun de r  f  rence (CECR, Conseil de l'Europe : 2000) qui d  finit six niveaux de comp  tences en langues (A1, A2, B1, B2, C1, C2). C'est un instrument pratique qui permet aux diff  rents acteurs en formation de langues de pouvoir   tablir les objectifs    atteindre, il d  crit les capacit  s langagi  res, les savoirs mobilis  s pour les d  velopper, les situations et domaines dans lesquels on peut   tre amen      utiliser une langue   trang  re pour communiquer.

Le CECR adopte les recherches des linguistes Canale et Swaine (1981) ainsi que celles de Bachman (1990) et propose pour l'enseignement un mod  le dit «actionnel», La perspective actionnelle consid  re l'usager et l'apprenant d'une langue comme des acteurs sociaux ayant    accomplir des t  ches dans des circonstances et un environnement donn  es,    l'int  rieur d'un domaine d'action particulier.

Une t  che est d  finie par le CECR comme « toute vis  e actionnelle que l'acteur se repr  sente comme devant parvenir    un r  sultat donn   en fonction d'un probl  me    r  soudre, d'une obligation    remplir, d'un but qu'il s'est fix   ». Une t  che requiert un plan de travail pour atteindre un objectif qui est extralinguistique qui passe par l'utilisation de la langue et la r  flexion, elle incite alors    la communication entre les apprenants et aboutit    un produit communicatif.

Notre programme de formation se situe dans ce cadre m  thodologique, il se base aussi sur le domaine de l'enseignement du fran  ais sur objectifs sp  cifiques (FOS) qui « est une d  marche fonctionnelle d'enseignement et d'apprentissage : l'objectif de la formation linguistique n'est pas la ma  trise de la langue en soi mais l'acc  s    des savoir-faire langagiers dans des situations dument identifi  es de communication professionnelles ou acad  miques » (Cuq : 2003).

Pr  parer une formation et   valuer dans une perspective FOS « consiste    proc  der    l'analyse des besoins langagiers des apprenants, c'est-  -dire    recenser les situations de communications cibles et les discours qu'elles comportent en collectant des donn  es langagi  res authentiques, s  lectionn  es, trait  es et transform  es en support de formation linguistique » (Cuq : 2003). Ainsi la pr  paration aboutit    un programme sur mesure qui fixe des objectifs d'apprentissage tr  s pr  cis et des d  lais de mise en   uvre limit  s (Mangiante et Parpette : 2004). L'  valuation dans une telle

formation suit la m  me logique dans le sens o   elle va cibler les objectifs sp  cifiques du programme pour   tre une partie int  grante de la d  marche d'enseignement/apprentissage.

Nos cours ont pr  par   les   tudiants de niveau B1 et B2 (CECR)    r  aliser des t  ches dans des situations de communication li  s    leurs besoins langagiers dans le cadre de leur formation ou dans des situations professionnelles d'enseignants.

3. Programme du fran  ais sur objectifs sp  cifiques et probl  matique de l'  valuation

Suite    l'analyse des besoins langagiers de nos   tudiants, nous avons constat   que la priorit   est    accorder    l'utilisation du fran  ais dans le cadre des activit  s de l'  tudiant au cours de sa formation    l'universit   et celles de sa profession comme enseignant    l'  cole. Le tableau 1 pr  sente les principaux objectifs d  termin  s pour notre module de formation:

Objectifs li��s aux activit��s de la profession d'enseignants	Objectifs li��s �� la formation �� l'universit��
• Pr��parer une le��on (lecture/ math. /sciences). • Animer un cours (sciences, math��matiques, fran��ais) • Pr��senter aux ��l��ves une exp��rience au laboratoire. • Pr��senter un livre aux enfants pour motiver �� lire. • Conter / raconter aux ��l��ves. • R��diger un compte rendu suite �� une visite lors d'une sortie scolaire. • Pr��parer et g��rer les ��lections des d��l��gu��s de classe. • Animer une heure de vie en classe. • R��aliser un panneau d'information. • R��aliser un livre documentaire. • Participer aux r��unions �� l'��cole (r��union de coordination, conseil de classe). • Animer une r��union avec les parents d'��l��ves. • R��diger une fiche p��dagogique d'une le��on. 	• Suivre un cours et prendre notes. • Pr��parer et pr��senter un expos�� oral. • R��diger ou pr��senter oralement un compte rendu d'une visite �� l'��cole. • Rendre compte/ r��sum�� par ��crit ou oralement un document lu. • Rendre compte �� l'oral d'un projet r��alis�� ou d'une s��ance anim��e lors du stage. • Pr��senter un compte rendu oral d'activit��s pr��par��es ou r��alis��es dans le cadre des travaux pratiques �� l'universit��. • R��diger une synth��se de documents. • R��diger un rapport d'observation de classe. • R��diger un rapport de stage. • R��diger un article d'opinion �� publier dans le journal de l'universit��. • Participer �� un forum de discussion sur internet. • Participer �� un entretien d'embauche.

Tableau 1 : Objectifs du module de formation

Lors de la premi  re application de ce programme de fran  ais sur objectifs sp  cifiques, la question de l'  valuation des acquis s'est impos  e. Il n'existe pas d'outils d'  valuation adapt  s    notre dispositif. Dans un premier temps, une   valuation suivant les modalit  s des examens de certification DELF/DALF¹ a   t   appliqu  e, elle nous a permis de bien   valuer le niveau linguistique des apprenants. Elle avait   galement l'avantage de mettre    notre port  e des outils et instruments d'  valuation bien pr  cis (crit  res, modalit  s, bar  me...). Cependant, l'inconv  nient de telle modalit   est qu'elle ne s'articule pas convenablement avec le contenu enseign   et les comp  tences et savoirs faire du domaine professionnel de nos apprenants. En pr  sentant les t  ches

¹ DELF = Dipl  me d'  tude en langue fran  aise. DALF = Dipl  me approfondi de langue fran  aise

langagi  res de notre programme en face de celles du DELF, le tableau 2 illustre cette inad  quation :

T��ches langagi��res de notre dispositif	T��ches langagi��res des ��preuves du DELF
• Pr��parer une le��on (lecture/ math. /sciences). PE² • Animer un cours (sciences, math��matiques, fran��ais). PO³ • Pr��senter aux ��l��ves une exp��rience au laboratoire. PO • Animer une r��union avec les parents d'��l��ves. PO • Suivre un cours et prendre notes. CO⁴, PE • Pr��parer et pr��senter un expos�� oral. PO • R��diger ou pr��senter oralement un compte rendu d'une visite �� l'��cole. PE ou PO	• ��crire une lettre formelle/un article critique pour prendre une position argument��e. PE (B2) • ��crire un essaie/courrier/article pour exprimer une attitude personnelle. PE (B1) • pr��senter et d��fendre son point de vue �� partir d'un document. PO (B2) • Participer �� un entretien dirig�� par l'examinateur. PO (B1) • Exprimer un point de vue �� partir d'un document d��clencheur. PO (B1)

Tableau 2 : Diff  rence entre les t  ches langagi  res du programme et celles du DELF

Nos form  s ont besoin par exemple de pr  parer et animer un cours, de participer    une r  union de travail (conseil de classe, r  union de coordination...) or les examens du DELF   valuent leurs niveaux en leur proposant d'  crire un courrier, un article, un essai... les savoirs et savoirs faire requis dans ces deux types de production sont diff  rents. Ainsi l'  valuation n'est pas adapt  e aux contenus du cours et surtout elle n'est pas li  e aux comp  tences scolaires et professionnelles de notre public. Evaluer suivant le DELF nous a contraint    consacrer une partie de la formation pour la pr  paration aux modalit  s et types d'  preuves du DELF. Il nous a conduit par la suite    la recherche et    la conception des   valuations plus coh  rentes avec notre programme.

4. D  marche dans la conception d'outils

Conform  ment aux principes du CECR relatifs    l'  valuation, l'apprenant est consid  r   comme un acteur social qui mobilise ses comp  tences pour r  aliser des t  ches. Les activit  s d'  valuation sont con  ues en fonction des t  ches fix  es dans le programme de formation.

La premi  re   tape de la d  marche est la « d  limitation » de chaque t  che, cela consiste    pr  senter les comp  tences n  cessaires    sa r  alisation et permet de d  terminer les   l  ments    prendre en consid  ration dans l'  valuation. Ce travail est appel   « sch  matisation de la t  che » (CIEP : 2012) ou « table de sp  cification » (Tagliante : 2005).

En vue des   l  ments du sch  ma de la t  che, une ou plusieurs activit  s d'  valuation sont pr  par  es avec des consignes claires et   ventuellement des supports adapt  s. Ces activit  s sont choisies en fonction du niveau des apprenants      valuer et de leurs centres d'int  r  t.

Par la suite et pour mettre en   uvre une   valuation crit  ri  e qui permet d'  valuer la comp  tence langagi  re, comme le pr  conise le CECR, nous retenons comme cat  gories de base les 3 composantes de la comp  tence de communication en langue : pragmatique, linguistique et sociolinguistique/ socioculturelle. La composante pragmatique renvoie    la ma  trise fonctionnelle de la langue, celle linguistique renvoie    la connaissance de la « norme » de la langue. Quant    la

² PE = Production   crite

³ PO = Production orale

⁴ CO = Compr  hension orale

composante sociolinguistique, elle repr  sente les outils n  cessaires pour l'utilisation de la langue dans sa dimension sociale.

Pour chaque composante un ensemble de crit  res sont fix  s, pour chacun nous avons d  termin   « les indices de performance » qui sont les   l  ments observables dans la r  alisation de l'activit   d'  valuation. Un indice de performance est appel   « indicateur, c'est la trace concr  te de ce qu'  crit ou dit l'apprenant. Un m  me crit  re peut se traduire par une vari  t   d'indicateurs » (Barb   : 2005).

Enfin il faut mesurer le degr   de performance de l'apprenant en   tablissant un bar  me. Tous les crit  res n'ont pas la m  me valeur et on ne peut s'appuyer sur des grilles standard. Nous avons pr  par   pour chaque outil un bar  me pond  r   c'est « la grille de notation   tablie avant l'  valuation et indiquant quelle valeur accorder    chaque crit  re » (Tagliante : 2006).

En r  sum  , Chaque t  che du programme a   t   mod  lis  e    l'aide d'un sch  ma pr  sentant ses caract  ristiques et les comp  tences    prendre en consid  ration pour l'  valuer. Ensuite sont pr  par  s, l'activit   d'  valuation, les crit  res, les indices de performance et le bar  me. Afin d'illustrer notre travail, nous allons choisir, parmi les outils con  us, un exemple pour   valuer une t  che de notre formation.

5. Outils d'  valuation de la t  che : faire une fiche de pr  paration d'un cours de math  matiques au primaire

Concevoir une fiche de pr  paration d'une le  on est une t  che r  alis  e par l'  tudiant futur enseignant de math  matique au primaire durant sa formation initiale    l'universit  , elle fait aussi partie des activit  s professionnelles de l'enseignant. Cette t  che est en ad  quation avec le profil et le niveau de comp  tences des apprenants, il s'agit du niveau B1 du CECR.

Nous pr  sentons sa sch  matisation (sa table de sp  cification), l'activit   d'  valuation pr  par  e, les crit  res d'  valuation avec les indices de performance ou   l  ments observables et enfin le bar  me pond  r   en fonction des crit  res, du niveau et des indices d  termin  s.

5.1 Sch  ma de la t  che

C'est « un outil de travail qui concr  tise et fait visualiser, de fa  on non lin  aire, les diff  rents   l  ments qui entrent dans la r  alisation d'une t  che » (Tagliante : 2005). Ces   l  ments, pr  sent  s dans le CECR, sont :

- Le contexte de la communication,
- Les comp  tences g  n  rales c'est-  -dire les savoirs, savoir faire, savoir   tre et savoir apprendre non langagiers,
- Les comp  tences    communiquer en langue avec les trois composantes : pragmatique, linguistique et sociolinguistique d  finis ant  rieurement,
- Les strat  gies ou tactiques utilis  s pour faciliter la communication ou faire face aux difficult  s. On les r  partit en trois cat  gories : de pr  paration, de s  lection et de communication,
- Les activit  s langagi  res mises en jeu dans la r  alisation de la t  che c'est-  -dire la compr  hension orale ou   crite, l'expression orale ou   crite et l'interaction orale,
- Le domaine (public ou priv  ),
- Le th  me.

Pour d  terminer tous les   l  ments composant le sch  ma d'une t  che, il est n  cessaire de se baser sur une exp  rience r  elle de r  alisation de la t  che    sch  matiser et d'avoir analys   le produit communicatif obtenu. Les donn  es pr  sent  es dans le sch  ma (figure 1) sont constitu  es suite    un travail d'analyse,    la fois d'un corpus de fiches de pr  paration de le  ons, adopt  es dans notre institution mais aussi d'un corpus de cours de math  matiques enregistr  s dans les   coles au Liban.

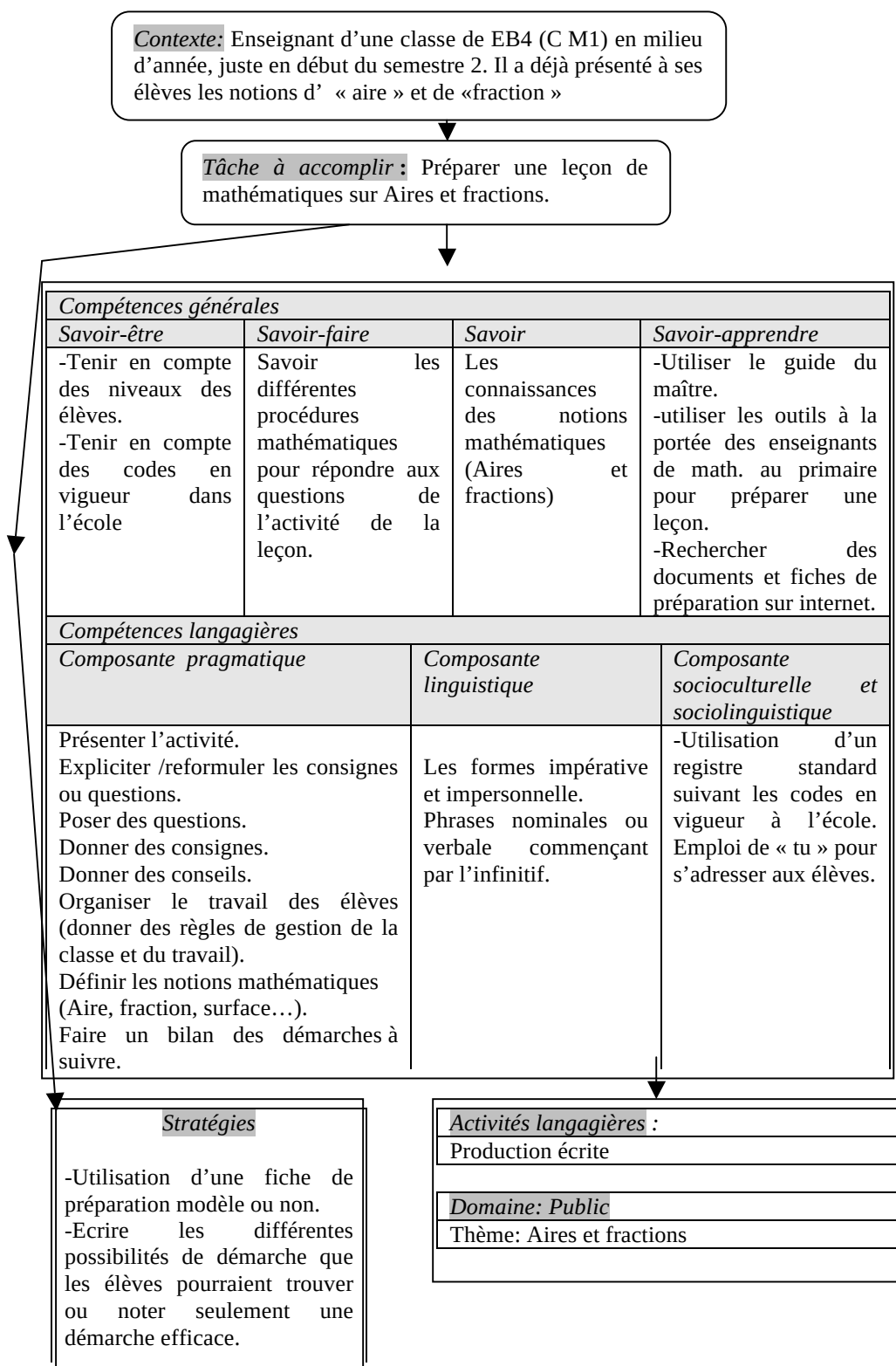


Figure 1 : Sch  ma de la t  che

5.2 L'activit   d'  valuation

Pour pr  parer une le  on l'enseignant    recours    un support p  dagogique comme le manuel. Nous avons choisi une le  on d'un manuel de math  matiques de EB4 (classe de 4^{  me} ann  e primaire) et l'activit   consiste    mettre l'apprenant   valu   en situation de communication r  elle, il est amen      r  aliser la fiche de pr  paration.

Il est important de signaler qu'une fiche de pr  paration de le  on peut   tre con  ue suivant des formes, composantes et modalit  s tr  s vari  es. Le mod  le et la structure qu'on a chois  s se base sur le corpus des fiches et les mod  les donn  s dans notre institution de formation. Des cours de math  matiques enregistr  s dans les   coles primaires au Liban nous ont guid   dans le choix des diff  rentes phases de la le  on et par cons  quent les diff  rentes parties de la fiche    pr  parer.

Nos consignes ont indiqu   d'une mani  re pr  cise ce que l'apprenant doit faire. Nous avons opt   pour une production tr  s guid  e afin d'  viter des variations dans la pr  sentation et les composantes de la fiche ce qui compliquera la t  che de l'  valuateur et rendra l'  valuation moins objective. Voici l'  preuve pr  par  e.

Module : Le fran  ais de l'enseignement
Epreuve de Production   crite B1
Dur  e : 1h

Vous trouverez    la page suivante une activit      pr  parer pour une le  on de math  matiques en classe de EB4 (CM1).

Faites une fiche de pr  paration de cette le  on, en y int  grant :

Partie 1 : fiche descriptive

Classe :

Titre de la le  on :

Dur  e totale:

Objectifs de l'activit   :

Mat  riel :

Modalit   du travail :

Partie 2 : Reformulation des consignes de l'activit  

Apr  s avoir   crits les donn  es de l'activit   au tableau certains   l  ves n'ont pas compris les consignes.

Pour les aider    comprendre proposer une reformulation de chacune de 4 consignes.

Partie 3 : Consignes de gestion du travail

Vous avez choisi une modalit   de travail en classe pour faire cette activit   par vos   l  ves, Donnez 4 consignes    vos   l  ves pour organiser le travail en classe.

Partie 4 : D  finir les notions de la le  on

Vous trouvez qu'il est n  cessaire d'expliquer    certains   l  ves les notions de « Aire », « surface », « fraction ».

Formulez une d  finition de chacune de ces trois notions    donner    vos   l  ves.

Partie 5 : Synth  se –bilan de la d  marche suivie

Vous formulez :

-une synth  se de la (des) d  marche(s)    suivre pour construire les surfaces dans 1 et 2

-un corrig   de 3 en expliquant la d  marche suivie

-une r  ponse    la question 4 en pr  sentant la d  marche    suivre pour obtenir cette r  ponse.

Figure 2 : Consignes de l'activit   d'  valuation

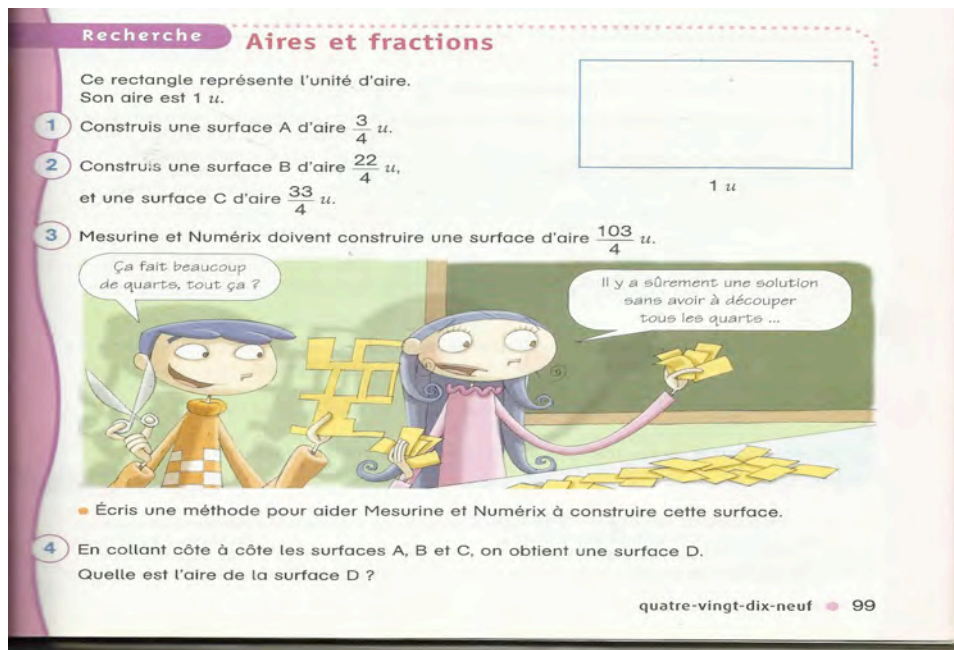


Figure 3 : Support de l'activit  

5.3 Crit  res d'  valuation et indices de performance (Tableau 3)

Afin de rendre l'  valuation objective, il faut d  terminer des crit  res auxquels peut se r  f  rer l'  valuateur. Ces crit  res sont plus pr  cis si on d  termine des indices observables dans la production des apprenants. La grille crit  ri  e (Tableau 3) retenue dans notre d  marche de conception comporte les 3 trois composantes d'une comp  tence de communication.

Dans la composante pragmatique, il s'agit de v  rifier si l'apprenant a :

- Respecter les consignes de l'activit   qui se traduit par le respect du type de discours demand   (informatif, injonctif, argumentatif), la pr  sence des fonctions langagi  res n  cessaires    la r  alisation de la t  che, du respect de la longueur demand  e,
- Formuler les actes de parole pertinents en fonction du discours    produire et surtout de la t  che r  alis  e,
- Pr  senter un discours coh  rent et encha  ner logiquement.

Dans la composante sociolinguistique et socioculturelle, on   value :

- Le code de pr  sentation qui est constitu   des   l  ments formels de pr  sentation dans les   crits, le non verbal, l'adaptation du comportement du locuteur    la situation de communication et le respect des tours de parole    l'oral,
- Le registre de langue s'il est adapt      la situation de communication.

Dans la composante linguistique, il s'agit de v  rifier :

- La morphologie c'est-  -dire la structure interne des mots : conjugaison, contraction des articles, formation des adjectifs...,
- La syntaxe qui concerne l'organisation des phrases, l'ad  quation des temps et modes verbaux utilis  s, l'accord des noms et des adjectifs, la structure des phrases, l'utilisation des articulateurs...,
- Le lexique dans sa vari  t   et sa pertinence par rapport au contexte,
- L'orthographe et la ma  trise de la ponctuation en production   crite,
- La prononciation dans la production orale.

Conclusion et perspective

En guise de validation, nous avons fait une premi  re exp  rimentation des outils pr  par  s avec trois groupes d'enseignants en formation continue. En comparaison avec les   valuations du DELF les apprenants en formation ont tous fait toutes les activit  s demand  es, il n'y avait pas de productions hors sujet, ils ont constat   la coh  rence entre les enseignements donn  s et l'  valuation et cela a   t   exprim   dans une « fiche d'  valuation de la formation » remplie au terme de leur session de formation. Du cot   des enseignants, ils sont amen  s    centrer leur cours sur les objectifs sp  cifiques du programme. Alors qu'auparavant, ils consacraient une partie de la formation    pr  parer aux   preuves du DELF et parfois    cibler d'autres objectifs non li  s au programme de la formation.

N'ayant pas encore constitu   une base d'outils qui couvre l'ensemble des t  ches du programme, la premi  re perspective de ce travail est de constituer une base d'outils compl  te,    alimenter r  guli  rement par des nouvelles activit  s. Nous envisageons   galement   valuer notre dispositif de formation afin d'observer ses cons  quences sur les enseignements, les apprentissages et particuli  rement sur le d  veloppement des comp  tences langagi  res professionnelles.

R  f  rences bibliographiques

- Bachman, L.F. & Palmer, A.S. (1987) The Construct Validation of Some Components of Communicative Proficiency *TESOL Quarterly*, 16, 4, 449 – 465.
- Barb  , G. (2005) Strat  gies de formation pour changer les pratiques   valuatives. In *Apprentissage d'une langue   trang  re/seconde*. (pp . 82 -101). Bruxelles : De Boeck.
- Canale, M. & Swain, M. (1981) Theoretical Bases of Communicative Approaches to Second Language and testing. *Applied Linguistics*, 1, 2-43.
- Centre international d'  tudes p  dagogiques CIEP (2012) *PRO FLE/   valuer, professionnalisation en fran  ais langue   trang  re*. <http://cned-bb9.blackboard.com/>
- Conseil de l'Europe (2000) *Cadre europ  en commun de r  f  rence* Paris : Didier.
- Cuq, J.-P. (2003) *Dictionnaire de didactique du fran  ais langue   trang  re et seconde*. Paris : Cl   international.
- Hymes, D. (1975) *Towards Communicative Competence*. Philadelphia : University of Pennsylvania Press.
- Lussier, D. (1992) *Evaluer les apprentissages dans une approche communicative*. Paris : Hachette.
- Mangiante, J.M. & Parpette, C. (2004) *Le Fran  ais sur Objectif Sp  cifique, de l'analyse des besoins    l'  laboration d'un cours*. Paris : Hachette FLE.
- Smaili, W. (2012) Des ressources p  dagogiques actionnelles pour les cours de fran  ais    l'universit  . in *Communiquer, Echanger, Collaborer en fran  ais dans l'espace m  diterran  en et balkanique. Actes du 7^o congr  s Panh  llenique et International des Professeurs de Fran  ais*. Ath  nes : Universit   d'Ath  nes.
- Tagliante, C. (2005) *L'  valuation et le cadre commun europ  en, Techniques et pratiques de classe*. Paris : Cl   international.
- Tagliante, C. (2006) *La classe de langue, Techniques et pratiques de classe*. Paris : Cl   international.

L'ÉVALUATION D'UN DISPOSITIF D'INSERTION ; ÉTUDE DU DÉVELOPPEMENT PROFESSIONNEL PERÇU DES ANIMATEURS PAR L'ANALYSE DES COMPÉTENCES

Lucie Aussel*

*Université de Toulouse

UMR « Éducation, Formation, Travail et Savoirs »

École Nationale de Formation Agronomique de Toulouse-Auzeville (ENFA)

2 route de Narbonne

31326 Castanet-Tolosan Cédex (France)

lucie.aussel@educagri.fr

Mots-clés : évaluation de dispositif – analyse du travail – développement professionnel perçu – entretiens d'autoconfrontation – évaluation de compétences

Résumé. Nous proposons d'appréhender le développement professionnel perçu à travers l'évaluation d'un dispositif d'insertion. Cette évaluation menée dans le cadre d'un appel à projet vise à appuyer le développement de ce dispositif. Nous expliciterons le dispositif d'évaluation que nous avons mis en place ainsi que les premiers résultats concernant le développement professionnel des animateurs de ce dispositif d'insertion. La mise en œuvre de cette expérimentation a entraîné une normalisation des pratiques modifiant le « travailler ensemble » de ces animateurs. Ainsi, des pratiques de type collaboratif ont émergé et les échanges entre les travailleurs ont augmenté. À travers l'analyse d'entretiens d'autoconfrontation, nous avons constaté que l'opérationnalisation de ce dispositif a entraîné pour ces animateurs l'accroissement de nouvelles compétences. Nous avons ciblé la compétence comme un indicateur de développement professionnel. Nous présenterons les trois catégories de compétences que l'analyse des entretiens d'autoconfrontation nous a permis de déceler.

Introduction

L'évaluation, au cœur de ce travail, a pour objet un dispositif expérimental destiné à favoriser l'orientation et l'insertion d'étudiants français inscrits dans l'enseignement supérieur agricole¹. Porté par une association spécialiste du recrutement et de l'insertion des cadres et techniciens du secteur agricole, il s'est déroulé sur les deux années de formation des étudiants. Tout d'abord, durant leur première année (2009/2010) les étudiants ont suivi un module devant leur permettre de découvrir les fonctions de l'entreprise, de décrypter les profils métiers. Ensuite, durant leur seconde année (2010/2011), ils ont pu participer à un module complémentaire davantage axé sur les démarches de recherche d'emploi. L'expérimentation a été mise en place sur l'ensemble du territoire national français, auprès de soixante-sept classes (770 étudiants) préparant quatre spécialités de l'enseignement supérieur agricole². Nous proposons dans cette contribution, d'appréhender dans un premier temps, l'évaluation du dispositif puis, dans un second temps, de développer les premiers résultats de notre recherche sur le développement perçu des animateurs de ce dispositif.

1 Cadre théorique

Nous envisageons le dispositif comme une organisation complexe (Morin, 1990) permettant d'organiser l'espace social par la médiation d'instruments (Rabardel, 1995) et de créer les

¹ Préparant un BTS : brevet de technicien supérieur agricole

² Les options : Analyse et Conduite des Systèmes d'Exploitation (ACSE), Gestion et Protection de la Nature (GPN), Service en Espace Rural (SER) et Sciences et Technologies des Aliments (STA) ont participé à l'expérimentation.

conditions nécessaires à la réalisation d'une intention. Le dispositif pour Peeters et Charlier (1999) permet de prendre en compte la dimension technique de certains phénomènes sociaux. Il est un moyen de traiter des zones d'ombres laissées auparavant de côté en dépassant les cadres initiaux d'analyse. Il existe selon Audran (2010) deux types de paradigmes permettent d'analyser le dispositif. Il présente tout d'abord, un paradigme déterministe dans lequel le dispositif existe par la volonté d'un concepteur et agit sur une situation. Selon l'auteur, on peut parler d'intentionnalité, de stimulus. Il présente ensuite, un paradigme phénoménologique, où la situation préexisterait à sa propre actualisation. Ainsi, selon ses deux cadres d'analyses, le dispositif pourrait être une forme d'influence dirigée « sur » dans l'intention de normaliser les individus, mais aussi un moyen, un support à l'autonomie individuelle dirigée « pour ».

L'évaluation est un système normé de jugements d'un référent vis-à-vis d'un référent en vue de prendre une décision quant à l'action évaluée. Le référent renvoie selon Hadji (1989) à « *des données qui sont de l'ordre du fait, et qui concerne l'objet réel à évaluer* » (p.25) et les référents à « *des données qui sont de l'ordre de l'idéal et qui concerne des attentes, des intentions ou des projets s'appliquant au même objet* » (*Ibid.*, p.25). Figari (2008) indique qu'en opposition à l'évaluation des apprentissages, plusieurs notions existent qui relèvent du « système », de « l'organisation », de « l'institution », du « curriculum », de la « politique éducative ». Selon lui, celles-ci sont larges et permettent difficilement de « *circonscrire des objets précis et isolables* » (p.78) rendant peu courantes les études portant sur le sujet en éducation. Même si l'emploi courant de « dispositif » est approximatif, il existe selon Figari (*Ibid.*) des définitions de référence et des utilisations qui s'imposent par leur récurrence. Il en fait ainsi la synthèse, le dispositif éducatif est :

- un objet concret observable (alternance, formation à distance, formation continue des enseignants, etc.) ;

- une abstraction (découpage de phénomènes donnant lieu à une représentation précise et une explication de leur régularité : dispositif de lutte contre l'échec scolaire, etc.).

Ainsi, dans le domaine de l'éducation se référer aux dispositifs reviendrait à convoquer un système complexe de situations reliées autour d'une intention de base articulée vers un public ciblé (des étudiants, des élèves, des décrocheurs, etc.) Son caractère générique lui permet d'être adapté à n'importe quelle situation. Ce qui le rend unique est qu'il est un artefact imaginé, construit et mis en place par l'homme pour résoudre ou améliorer une situation.

2 Évaluer un dispositif d'insertion

Nous proposons ici d'explicitier le dispositif d'évaluation que nous avons mis en place. Celui-ci est organisé selon deux orientations ; la première doit permettre de fournir des informations pour objectiver (a), et la seconde doit permettre de fournir des informations pour expliquer (b).

Ainsi, le premier objectif que nous poursuivons est un objectif de vérification. C'est une mesure de l'effet poursuivi. Nous l'obtenons par la comparaison de l'adéquation ou de l'inadéquation des résultats de l'expérimentation (référé) et des finalités visées par les expérimentateurs (référé). Cet axe concerne spécifiquement les étudiants à qui s'adresse cette expérimentation. Nous voulons connaître les effets de ce dispositif sur l'insertion réelle des étudiants d'une part (a) et sur les représentations d'autre part (b). Nous avons constitué un groupe témoin, composé d'étudiants ne participant pas au dispositif, afin de comparer leurs réponses à ceux du groupe expérimental (T= témoin, E= expérimental). Les étudiants du groupe E s'insèrent-ils mieux que ceux du groupe T ? L'idée d'une « meilleure » insertion s'étoffe sur le principe que les étudiants qui ont reçu de l'information sur le fonctionnement de l'entreprise ainsi qu'une préparation à la démarche de recherche d'emploi (E) auront plus de facilités à discerner les postes ou les poursuites d'études auxquelles ils peuvent prétendre que ceux qui n'ont pas eu accès au contenu du dispositif (T). Pour réaliser une « bonne » insertion, nous formulons l'hypothèse que les étudiants devraient recourir à l'orientation professionnelle, c'est-à-dire être capable de sélectionner les débouchés potentiels de leur formation et cibler leur demande. Ainsi, les étudiants qui parviendraient à s'orienter seraient plus rapidement que les autres en emploi ou en formation et dans une situation qu'ils jugeraient convenable. La maîtrise de l'orientation pourrait être un atout pour l'insertion. Nous formulons

l'hypothèse que la « meilleure » orientation des étudiants passe également par un changement des représentations que les étudiants se font du marché de l'emploi, du fonctionnement de l'entreprise, de leurs possibilités de poursuite d'études et/ou de carrière. La participation au dispositif permet-elle de faire évoluer les représentations des étudiants ? Une comparaison entre les deux groupes (T et E) a été mise en place afin de constater l'évolution de la représentation des étudiants. Ce niveau d'évaluation des représentations des étudiants permet d'expliquer le niveau d'efficacité de l'expérimentation. Le recueil d'information s'est effectué par questionnaires.

Deuxièmement, dans une approche complémentaire, l'évaluation permet de révéler un effet induit par l'innovation (b). Il s'agit de porter attention sur des aspects non envisagés par les expérimentateurs et de regarder ce qu'il se passe au-delà des effets recherchés sur les étudiants. Nous nous intéressons aux effets du dispositif sur les animateurs, c'est-à-dire au développement professionnel de ces travailleurs engagés dans le dispositif. Le développement professionnel correspond à la construction de « *compétences lors de formations individuelles ou collectives, mais aussi la construction de compétences nouvelles par la pratique et la réflexion sur la pratique ainsi que les transformations identitaires des individus ou des groupes* » (Marcel, 2009). Dans quelle mesure la participation à cette expérimentation permet-elle le développement professionnel de ces acteurs ?

Troisièmement, à travers l'étude de l'efficacité perçue de chacune des phases de l'expérimentation, nous nous intéressons au sentiment de réussite des acteurs sur sa mise en place et ses finalités (b). Dans ce volet il est question de mettre au jour l'efficacité³ du dispositif. C'est ce que nous déterminerons par le récit de l'expérience vécue des différents acteurs de l'expérimentation (étudiants/animateurs).

3 Appréhender le développement professionnel perçu des animateurs

Nous détaillons un des volets de l'évaluation centrés sur le développement professionnel perçu des animateurs de ce dispositif d'insertion. Nous envisageons le développement professionnel comme un ensemble « *de transformations individuelles et collectives de compétences et de composantes identitaires mobilisées ou susceptibles d'être mobilisées dans des situations professionnelles* » (Barbier, Chaix et Demailly, 1994). Ainsi il correspond à la construction de « *compétences lors de formations individuelles ou collectives, mais aussi à la construction de compétences nouvelles par la pratique et la réflexion sur la pratique ainsi que les transformations identitaires des individus ou des groupes.* » (Marcel, 2009, p.157). Nous nous intéressons au développement professionnel perçu que nous évaluons par la mise en valeur des compétences développées lors du déploiement de ce dispositif. Ainsi, nous utilisons les compétences comme indicateurs du développement professionnel perçu des animateurs. La compétence est le processus générateur du produit qu'est la performance, elle est finalisée, contextualisée, spécifique et contingente (Wittorski, 1998). La compétence vise l'efficacité, elle est orientée vers une finalité et se réalise dans une situation donnée. (Lenoir 2011, p.20) « *Est jugé compétent celui dont la performance, au regard de l'efficacité de son acte et de la qualité du résultat est appréciée comme positive* » (Wittorski, 2005 p.187). Il en va de la compétence un jugement de valeur impliquant un cadre de référence : une culture sociale, professionnelle. C'est un « *construit social dynamique* » (Lenoir, 2011, p.19). La compétence peut être individuelle ou collective qu'elle soit produite seule ou en interaction avec d'autres. La mise en place de ce nouveau dispositif entraîne une situation innovante de travail qui crée pour ces travailleurs des « zones d'incertitudes » (Crozier & Friedberg, 1977). Nous formulons l'hypothèse que les animateurs vont chercher à maîtriser ces « zones d'incertitudes », notamment par l'appropriation du nouveau projet avec la construction de compétences. Cela pourrait être l'acceptation du changement et son appropriation, le développement professionnel apparaîtrait alors comme un nouvel équilibre, comme le « *traitement de la complexité inédite* » (Crahay, 2006).

³ L'efficacité d'une cause est la capacité de celle-ci à produire un effet, elle est contextualisée ; une cause efficace est une cause qui obtient des résultats dans son activité.

4 Méthodologie de recherche

Étudier le développement professionnel perçu signifie se placer du point de vue des enquêtés. Pour être en mesure d'effectuer ce basculement nous avons décidé de travailler sur des propos recueillis. Plus encore que de mener des entretiens semi-directifs, nous avons choisi de faire passer des entretiens d'autoconfrontation simples (Clot et Faïta, 2001) sur la base des séquences filmées durant les observations. Cette méthode nous a paru la plus adéquate, car elle est la création d'une nouvelle situation : dite d'« *activité sur l'activité* » (Faïta, 2007). Nous avons demandé aux animateurs de visionner les films en même temps que nous, et de nous commenter leur activité. L'entretien d'autoconfrontation est dit simple, car aucune autre personne n'a assisté à l'entretien (collègues, étudiants, etc.) Cette pratique permet à l'animateur de devenir « *successivement acteur de la nouvelle situation initiée à partir de son vécu, puis analyste lui-même* » (Faïta, 2007, p.60/61). Cet exercice lui permet ainsi d'expliquer ses choix et ses stratégies, ses réussites et ses difficultés et de mettre au jour les compétences mobilisées. Le produit de l'autoconfrontation est le résultat d'une co-élaboration, Clot (2001) parle de co-analyse, car selon lui, le cadre de l'analyse est fait des rapports entre chercheur et individus entretenus. Ainsi, « *lorsqu'il est engagé dans un processus d'autoconfrontation [...] le sujet s'adresse en premier lieu à lui-même, sous couvert des commentaires qu'il destine à l'observateur, car il projette sur son activité passée un regard qui le transforme en interlocuteur actuel de celui qu'il a été au moment de l'action* » (Faïta, 2007, p.34).

Les animateurs que nous avons suivis sont tous salariés de la même association, porteuse de l'expérimentation sociale. Nous avons choisi d'observer les interventions de six animateurs sur les quatorze concernés par la mise en place du dispositif. Nous les avons sélectionnés en fonction de leur niveau d'implication dans le déroulement du dispositif. Nous avons pu constituer deux groupes. Deux femmes et un homme composent le premier groupe, ils sont fortement impliqués, ils se sont inscrits dans le groupe de travail qui a créé la trame pédagogique du module de formation. Ils sont considérés comme experts au moment où nous commençons l'enquête. Deux femmes et un homme composent également le second groupe. Ils ont une implication plus faible, ils sont considérés comme novices au démarrage de l'enquête.

Nous nous sommes intéressés au développement professionnel perçu de ces animateurs dans un contexte d'innovation au travail. Ils sont employés d'une association organisée en délégations couvrant l'ensemble du territoire métropolitain. Ils ont participé à la création et à la mise en place du dispositif qui est le point de départ du contexte d'innovation. Elle entraîne quatre changements ; une nouvelle organisation du travail, de nouvelles pratiques, un nouveau public (des étudiants plus jeunes), de nouveaux objectifs. Dans ce cadre particulier de rupture de la routine professionnelle, nous questionnons la capacité d'adaptation de ces travailleurs. Quels effets ces changements provoquent-ils sur leur développement professionnel perçu ?

5 Résultats

Nous présentons ici les premiers résultats de notre étude. Nous ne prétendons pas affirmer avoir évalué objectivement les compétences de ces animateurs, mais grâce au procédé des entretiens d'autoconfrontation nous leur avons permis d'exprimer les compétences qu'ils mobilisent dans cette activité précise. Nous sommes d'accord pour dire qu'ils ont pu mobiliser bien plus de compétences qu'ils n'en ont décrites, mais nous nous basons uniquement sur leur discours. Nous avons détecté dans les discours trois types de compétences.

- a) Les compétences de contenu correspondent à des compétences ayant un lien direct avec le contenu du dispositif d'insertion. Elles pourraient s'apparenter aux compétences fonctionnelles qui permettent de répondre au caractère technique des tâches « *en s'appuyant au savoir relatif au domaine de l'emploi de référence* » (Leduc & Valléry, 2006). Ce sont les gestes, les postures, les paroles, mais aussi le raisonnement des opérateurs en lien avec la résolution d'un problème spécifique.

Ex. : innover, maîtriser/ s'approprier les outils et illustrer son discours d'exemples concrets, utiliser l'espace, etc.

- b) Les compétences pédagogiques, qui relèvent de la mise en place de ce contenu. Elles pourraient s'apparenter aux compétences sociales qui sont, pour Leduc & Valléry (2006), des « *aptitudes à se comporter de manière socialement appropriée dans différents contextes* ». Elles interviennent selon eux dans la gestion des interactions avec l'environnement social, elles se fondent sur les connaissances de l'autre à un moment donné.

Ex. : repérer les profils des étudiants, rendre les étudiants acteurs de l'intervention, traduire son discours (reformuler pour être compris-e de tous), guider la progression du travail, créer une ambiance de travail, aider à la conceptualisation, etc.

- c) Les compétences de distanciation mobilisent la réflexivité, elles correspondent à l'analyse que font les animateurs de la situation qu'ils observent. Nous n'avons pas trouvé dans la littérature, pour le moment, une catégorie qui pourrait correspondre à ce type de compétence, nous les appelons provisoirement compétences réflexives.

Ex. : Repérer ce qu'on pourrait améliorer de sa pratique, avoir un regard critique (positif ou négatif) quant au contenu de la formation, repérer les points forts et points faibles de son intervention, faire des propositions d'amélioration du contenu de la formation, etc.

Nous avons été attentifs aux compétences énoncées en fonction des deux groupes que nous avons constitués : expert et novice. Voici les premiers constats que nous sommes en mesure de faire :

- les animateurs du groupe novice ont tendance à énoncer davantage les compétences de contenu. Nous pensons que cela peut s'expliquer parce qu'ils n'ont pas participé à la création du module de formation, ainsi il leur a été plus coûteux de s'approprier le dispositif, c'est cet effort qu'ils mettent en valeur ;
- les animateurs du groupe expert ont davantage mentionné des compétences pédagogiques. Nous pensons qu'ils sont plus à l'aise avec le contenu de la formation, et qu'ils axent davantage leur discours sur leurs relations avec les étudiants ;
- les compétences de distanciation ont été exposées de la même façon par les deux groupes. La distinction expert-novice n'est plus adéquate pour ce type de compétence basé sur l'analyse de pratique. Les animateurs se trouvent sur un « pied d'égalité » face à cette situation.

Par ailleurs, nous avons remarqué que tous (animateurs experts ou novices) ont moins cité de compétences lors du second entretien. Ceci s'explique par la diminution de l'aspect innovant de l'expérimentation ; une banalisation liée à l'expérience, la routine professionnelle s'installant autour de ce projet.

6 Conclusion

Nous avons abordé le développement professionnel des animateurs de ce dispositif à travers une approche par les compétences. Celle-ci à cela de particulier qu'elle mobilise les acteurs dans une démarche d'analyse de l'activité. Nous avons pensé la compétence comme un indicateur potentiel de développement professionnel dans un contexte de mise en œuvre de dispositif. La compétence énoncée, perçue, place l'individu au cœur de son récit n'éloignant de son discours ni « *la configuration de ses modalités d'action mises en œuvre en situation* », ni « *l'environnement professionnel* » (Lefevre, Garcia, & Namolovan, 2009, p. 290), les deux autres composants du développement professionnel. Au contraire, l'individu en tant que praticien réflexif (Schön, Heynemand, & Gagnon, 1994) fait le lien entre le contexte et ses actions. Pour Lefevre & al. (2009, p.290), « *ces trois éléments sont interdépendants et participent à décrire, à comprendre et à expliquer le développement professionnel* ».

Nous nous sommes, en effet, basés sur les discours des animateurs afin d'y observer le type de compétence qu'ils énonçaient dans un exercice d'analyse de la pratique. Le premier constat que nous avons pu établir est celui de la présence de compétences énoncées dans chacun des entretiens que nous avons menés. En effet, quel que soit le groupe auquel ils appartiennent, les animateurs commentent leur activité en décrivant les compétences qu'ils mobilisent. Néanmoins, nous avons pu observer des « comportements » différents en fonction des groupes d'animateurs que nous avons créés (expert/novice). En effet, les animateurs des deux groupes n'ont pas formulé les mêmes types de compétences. Nous avons lié ces particularités au niveau d'expérience acquise dans le projet.

Il nous paraît important de poursuivre cette recherche en explorant le sentiment d'efficacité professionnel (SEP) de ces animateurs. En effet, en complétant cette analyse avec une étude du SEP des animateurs nous pourrions approfondir les explications de ces « comportements ».

Références bibliographiques

- Audran, J. (2010). Dispositifs et situations. Note de synthèse présentée en vue d'obtenir l'Habilitation à la Direction de Recherches. Paris
- Clot, Y., & Faïta, D. (2000). Genre et styles en analyse du travail. Concepts et méthodes. Travailler (4), 7-42.
- Clot, Y., Faïta, D., Frenandez, G., & Scheller, L. (2001). Entretiens en autoconfrontation croisée une méthode en clinique de l'activité. Éducation permanente (146), 17-25.
- Crahay, M. (2006). Dangers, incertitudes et incomplétude de la logique de la compétence en éducation. Revue française de pédagogie, (154), 97-110.
- Crozier, M., & Friedberg, E. (1977). *L'acteur et le système les contraintes de l'action collective*. Sociologie politique. Paris: Éditions du Seuil.
- Faïta, D., & Maggi, B. (2007). Un débat en analyse du travail deux méthodes en synergie dans l'étude d'une situation d'enseignement. Toulouse : Octarès
- Figari, G. (2008). L'évaluation des dispositifs éducatifs. *Mesure et évaluation en éducation*, 31(3), 77-93.
- Hadji, C. (1989). L'évaluation, règles du jeu des intentions aux outils (ESF Editeur.). Paris.
- Lefevre, G., Garcia, A., & Namolovan, L. (2009). Les indicateurs de développement professionnel, 5(11), 277-314.
- Lenoir, Y. (2011). Conséquences des conceptions curriculaires actuelles sur les modes éducatifs. Les dossiers des sciences de l'éducation, (25), 13-28.
- Marcel, J.-F. (2009). Le sentiment d'Efficacité Professionnelle, un indicateur pour connaître le développement professionnel des « nouveaux » professeurs de l'enseignement agricole français. Questions vives, 5(11), 161-176.
- Morin, E. (1990). *Science avec conscience*. Points. Paris: Seuil.
- Peeters, H., & Charlier, P. (1999). Contributions à une théorie du dispositif. *Hermès*, (25), 15-23.
- Rabardel, P. (1995). Les hommes et les technologies approche cognitive des instruments contemporains. U. Paris: A. Colin.
- Schön, D. A., Heynemand, J., & Gagnon, D. (1994). *Le praticien réflexif à la recherche du savoir caché dans l'agir professionnel*. Collection Formation des maîtres. Montréal: les Éd. Logiques.
- Sorel, M. & Wittorski, R. (2005). La professionnalisation en actes et en questions. Paris : L'Harmattan.
- Wittorski, R. (1998). De la fabrication des compétences. Éducation permanente, n°135.

LES EVALUATIONS A FAIBLES ENJEUX : QUEL R  LE JOUE LA MOTIVATION ? UNE EXPERIENCE A PARTIR DE PISA.

Saskia Keskpaik*, Thierry Rocher**

* Minist  re de l'  ducation nationale – DEPP, saskia.keskpaik@education.gouv.fr

** Minist  re de l'  ducation nationale – DEPP, thierry.rocher@education.gouv.fr

Mots-cl  s : PISA, motivation, thermom  tre d'effort

R  sum  . Bien r  ussir un test n  cessite de la motivation de la part des r  pondants. On peut ainsi s'interroger sur le r  le de la motivation dans des   valuations    faibles enjeux, telles que l'  valuation internationale PISA. Cette communication vise      tudier cette question en pr  sentant une exp  rimentation r  alis  e lors du pr  -test PISA en printemps 2011. Des   l  ments d'analyse sont offerts afin de pouvoir juger de l'effet des actions entreprises dans le cadre de cette exp  rimentation sur la motivation et la performance des   l  ves participant au test. La mesure de motivation de PISA – le « thermom  tre d'effort » – est   galement pr  sent  e et discut  e en termes de variabilit   dans diff  rentes   valuations nationales.

1. Introduction

La France se caract  rise dans PISA (*Programme for International Student Assessment*) par un important taux de non-r  ponse aux questions ouvertes, qui appelle une r  ponse structur  e et construite. La question sous-jacente est de savoir pour quelles raisons les   l  ves fran  ais n'y r  pondent pas : le font-ils par m  connaissance intrins  que des connaissances et comp  tences requises ou par absence de motivation ?

L'  valuation internationale PISA ainsi que la majorit   des dispositifs d'  valuation de la DEPP¹ constituent, pour les   l  ves participants, des   valuations    faibles enjeux. N'  tant pas not  es et n'ayant pas d'impact sur le parcours de l'  l  ve, ces   valuations suscitent sans doute moins d'anxi  t   mais aussi moins d'implication de la part des   l  ves que des   preuves not  es.

Ceci a amen   certains pays participant    PISA    prendre des mesures afin de favoriser la motivation des   tablissements et des   l  ves. Tr  s vari  es selon les pays, ces mesures incluent des dispositifs d'animation et de communication (p.ex. Islande, Allemagne) ainsi que diff  rentes formes d'incitations (p.ex. Etats-Unis, Grande-Bretagne, Br  sil).

Nous nous sommes int  ress  s, dans ce travail, aux questions suivantes :

- Quelle est l'importance du degr   de motivation ?
- Comment favoriser la motivation ?
- Quel est le r  le de la motivation dans la validit   et la comparabilit   des   valuations ?

2. Exp  rience    partir de PISA

Les   preuves PISA 2012 ont fait l'objet d'un pr  -test en 2011 sur 56   tablissements fran  ais. Nous avons souhait   utiliser cette phase de pr  -test de l'interrogation pour   tudier l'effet de la motivation des   l  ves sur la participation et le niveau de performance obtenu. Ces

¹ Direction de l'  valuation, de la prospective et de la performance.

56   tablissements ont   t   r  partis en deux populations, l'une – dite « groupe t  moin » – pour laquelle aucune information suppl  mentaire    celles ordinairement diffus  es n'a   t   donn  e conform  ment au protocole d'  valuation, l'autre – appel   « groupe exp  rimental » – pour laquelle des actions sp  cifiques de sensibilisation    l'importance du dispositif PISA ont   t   engag  es.

S'agissant de cette deuxi  me population, une campagne d'information et de sensibilisation    l'importance de PISA a   t   men  e sous les formes cibl  es suivantes :

- r  union pr  alable des chefs d'  tablissement et fourniture de mat  riels de communication (affichage, pr  sentation Powerpoint, exemples de courriers aux f  d  rations de parents et aux syndicats, « dipl  mes » de passation du pr  -test de PISA 2012) ;
- information cibl  e vers les enseignants concern  s (information par mails, exemples d'items comment  s, etc.) ;
- information aupr  s des parents d'  l  ves sous la forme de r  unions d'information pr  alables ;
- sensibilisation aupr  s des   l  ves selon l'implication des chefs d'  tablissement (pr  sentation de PISA, d  dramatisation de l'  valuation, repas sp  ciaux, go  ters pour les   l  ves participant    PISA dans l'  tablissement).

Un troisi  me groupe – appel   « groupe actif » – a pu   tre distingu   au cours de l'exp  rimentation. Constituant un sous-ensemble du groupe exp  rimental, les   tablissements « actifs »   taient les plus impliqu  s dans cette exp  rimentation (figure 1). En particulier, les chefs d'  tablissement du groupe actif se sont d  plac  s    Paris,    la DEPP pour une r  union d'  changes et d'informations.

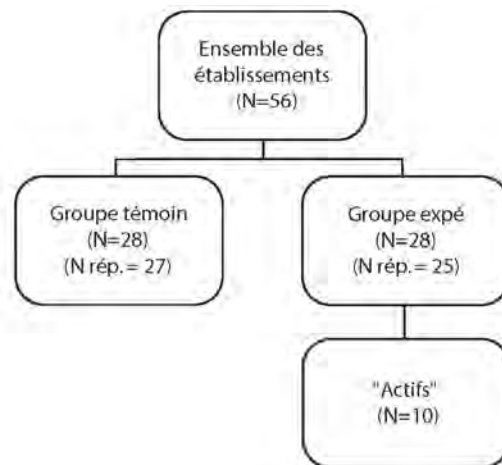


Figure 1 : R  partition des   tablissements dans les diff  rents groupes de l'exp  rimentation

Par ailleurs, un questionnaire a   t   adress      tous les   tablissements participant au pr  -test les interrogeant sur leur connaissance pr  alable de PISA ainsi que sur l'organisation de l'enqu  te dans l'  tablissement et sur les actions entreprises par l'  tablissement dans le but de motiver les   l  ves    la passation PISA. Les r  ponses    ce questionnaire montrent que les   tablissements connaissent relativement mal l'  valuation PISA : 2   tablissements sur 5 se disent peu ou pas inform  s sur PISA avant la passation. En outre, PISA est consid  r   comme un dispositif tr  s complexe : 2   tablissements sur 3 per  oivent la mise en place de PISA comme contraignante ou tr  s contraignante. En ce qui concerne les actions pour motiver les   l  ves, la mesure la plus souvent prise par les   tablissements est celle d'informer sur PISA : la plupart des   tablissements informe de la passation PISA les parents d'  l  ves et les enseignants, le faisant le plus souvent seulement par courrier.

En raison de la politique de redoublement et de la coupure coll  ge/lyc  e    15 ans, la population   valu  e dans PISA (les   l  ves de 15 ans) est tr  s diff  rente selon le type d'  tablissement : les   l  ves    l'heure sont au lyc  e (60 %) et les   l  ves en retard au coll  ge. Ce dernier point est

important dans la mesure o  , dans les coll  ges, les   l  ves s  lectionn  s pour repr  senter la France dans PISA sont exclusivement des redoublants. Ces   l  ves font plus souvent partie des   l  ves les moins performants de leur coll  ge ce qui rend plus difficile l'enjeu de la participation et de la motivation au sein des   tablissements vis-  -vis des enseignants tout particuli  rement.

3. R  sultats selon le groupe d'  tablissements

Nous attendions de ces mesures une augmentation de la participation des   l  ves, une diminution des non-r  ponses aux questions ouvertes, un degr   de motivation plus fort et, *in fine*, de meilleures performances observ  es dans l'  chantillon d'  tablissements ayant d  velopp   une politique active d'information et de sensibilisation.

Les r  sultats observ  s montrent une l  g  re augmentation de la participation des   l  ves du groupe exp  rimental par rapport    ceux du groupe t  moin dans les lyc  es sans que la participation augmente sensiblement dans les coll  ges (tableau 1). En revanche, l'exp  rimentation n'a pas eu d'impact sur les r  ponses aux questions ouvertes : il n'y a pas de variation significative du taux de non-r  ponse selon les diff  rents groupes d'  tablissements (tableau 2).

	Ensemble		Coll��ge		Lyc��e	
	Etab.	% particip.	Etab.	% particip.	Etab.	% particip.
T��moin	27	79,0	11	74,7	16	82,0
Exp��	25	81,2	11	77,3	14	93,6
Actifs	10	93,3	3	90,0	7	95,0
Ensemble	52	80,1	22	76,0	30	87,5

Tableau 1 : Taux de participation des   l  ves au test PISA

	Ensemble	Coll��ge	Lyc��e
T��moin	33,9	46,8	25,6
Exp��	31,7	46,8	20,7
Actif	29,1	47,7	21,3
Ensemble	32,8	46,8	23,1

Tableau 2 : Taux de non-r  ponse aux questions ouvertes    r  ponse construite (n=44)

En ce qui concerne l'impact de l'exp  rimentation sur la performance des   l  ves, on observe une l  g  re augmentation du score moyen dans des groupes exp  rimental et actif par rapport au groupe t  moin et cela pour toutes les   preuves PISA : l'  preuve sur ordinateur (*computer based assessment*, CBA), l'  preuve sur papier en culture financi  re (*financial literacy*, FL), l'  preuve sur papier en culture math  matique (*paper based assessment of mathematics literacy*, PBM) (tableau 3). En revanche, cet effet s'att  nue lorsqu'il est analys   « toutes choses   gales par ailleurs ». Les r  sultats d'une r  gression lin  aire montrent que le fait d'appartenir au groupe exp  rimental ou au groupe actif par rapport au groupe t  moin n'a pas d'impact significatif sur la performance dans le cas o   les diverses caract  ristiques des   l  ves telles que le sexe, le retard scolaire et le type d'  tablissement sont maintenues constantes. Ainsi, une meilleure motivation    r  pondre de la part des   l  ves suite    cette exp  rimentation ne se traduit pas clairement par une am  lioration dans la performance au test.

	CBA		FL		PBM	
	Moyenne	σ	Moyenne	σ	Moyenne	σ
T��moin	-0,3	1,0	0	1,1	-0,9	1,2
Exp��	0	1,3	0,1	1,3	-0,7	1,4
Actif	0,3	1,1	0,5	1,1	-0,4	1,2
Ensemble	-0,2	1,2	0	1,2	-0,8	1,3

Tableau 3 : Score moyen en CBA, FL et PBM selon le groupe d'  tablissements

4. Mesure de la motivation : le thermom  tre d'effort

4.1 Thermom  tre PISA : effort au test, effort    l'  preuve not  e, effort relatif

L'  valuation PISA mesure la motivation des   l  ves    l'aide d'un « thermom  tre d'effort » (figure 2). Les   l  ves r  pondent    deux questions concernant, d'une part, l'effort qu'ils ont fourni pour r  pondre au test PISA et, d'autre part, l'effort qu'ils auraient fourni si la note obtenue lors du test avait compt   pour leur bulletin scolaire. L'effort fourni au test PISA varie selon les pays. Les r  sultats de PISA 2006 montrent que les   l  ves fran  ais se caract  risent par un effort PISA relativement faible : la France occupait la 36^e place parmi 38 pays, ne d  passant que le Japon et l'Islande.

Quel effort avez-vous fourni pour r  pondre    ce test ?

Essayez de vous imaginer face    une situation de la vie r  elle (   l'  cole ou dans un autre contexte) qui est tr  s importante pour vous personnellement. Vous auriez envie de bien r  ussir, et pour cela, vous tentez de donner le meilleur de vous-m  me, en y consacrant le plus d'efforts possibles.

Dans cette situation, vous cocheriez la valeur la plus ��lev��e sur le « thermom��tre de l'effort », comme ci-dessous :	Par rapport �� la situation que vous venez d'imaginer, quel effort pensez-vous avoir fourni en r��pondant �� ce test ?	Si la note obtenue lors de ce test comptait pour votre bulletin scolaire, quel effort auriez-vous fourni ?
 ☒ 10 ☐ 9 ☐ 8 ☐ 7 ☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1	☐ 10 ☐ 9 ☐ 8 ☐ 7 ☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1	☐ 10 ☐ 9 ☐ 8 ☐ 7 ☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

Figure 2 : Thermom  tre d'effort de PISA

On note des variations inter-pays   galement pour l'effort fourni    l'  preuve not  e. Butler et Adams (2007) proposent l'effort relatif comme une mesure qui prend en compte ces variations et qui exprime, pour chaque pays, l'effort fourni au test PISA par rapport    l'effort fourni    l'  preuve not  e. Ces auteurs regroupent ensuite certaines des cat  gories de l'  chelle de l'effort relatif afin de d  gager des « profils » d'effort (tableau 4).

Diff��rence (Effort PISA - Effort ��preuve not��e)	Profils	Effort relatif	PISA 2006 France	
			%	score
-9		1	0,6	526
-8		2	1	467
-7		3	1,3	497
-6	Cyniques	4	1,7	485
-5		5	3,7	462
-4		6	7,2	491
-3		7	14,8	487
-2	R��alistes	8	21,3	494
-1	Assidus	9	17,1	479
0	Partisans	10	12,6	479
>0	Irr��alistes	0	2,3	433
-	Non-r��ponse	-	16,3	478

Tableau 4 : Niveaux d'effort selon les groupes

La figure 3 met en   vidence des diff  rences importantes entre les pays en ce qui concerne la r  partition des   l  ves selon le profil d'effort. On note   galement la variation consid  rable du score en fonction de ce profil. Par exemple, en France comme en Suisse, ce sont les   l  ves du profil « cyniques » qui sont les plus nombreux. Au Danemark, en revanche, ce sont les   l  ves « assidus » (qui d  clarent 1 point de diff  rence entre les deux   chelles). En outre, en Suisse, les   l  ves « cyniques » sont les plus performants alors que ce n'est pas le cas en France. Cela signifie que les m  mes r  ponses au thermom  tre traduisent potentiellement des comportements diff  rents selon les pays et le niveau des   l  ves.

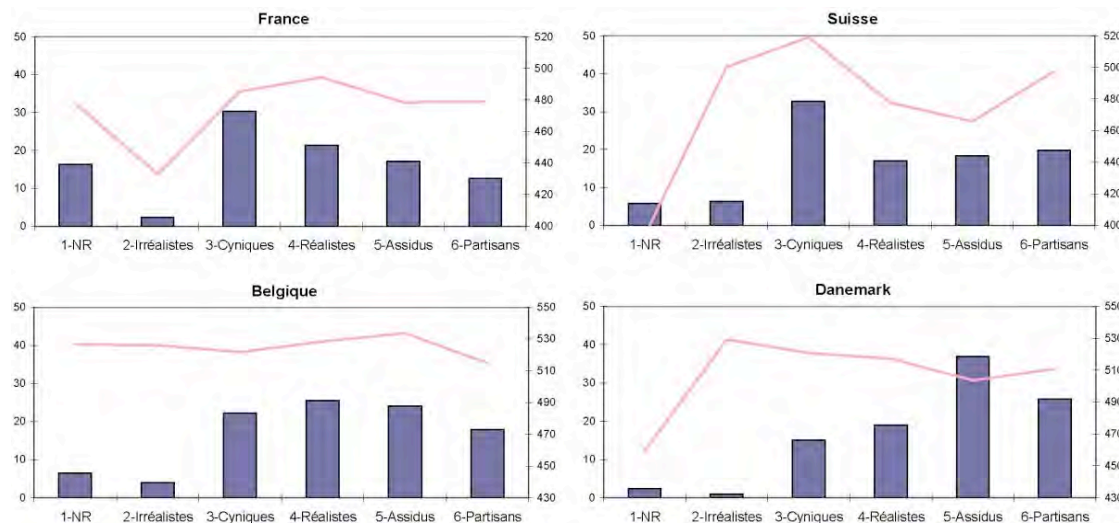


Figure 3 : Score selon profils d'effort relatif, PISA 2006

4.2 Thermom  tre dans l'exp  rimentation sur la motivation

Nous avons analys   les trois groupes d'  tablissements selon leurs profils d'effort. Les r  sultats de cette analyse montrent que les   l  ves du groupe exp  rimental ainsi que ceux du groupe actif sont proportionnellement plus nombreux      tre « assidus » et moins nombreux    avoir le profil « cyniques » que les   l  ves suivant leur scolarit   dans des   tablissements t  moins (figure 4). En

revanche, on note aussi un plus grand taux de non-r  ponse aux questions de motivation parmi ces   l  ves vis-  -vis des   l  ves du groupe t  moin. La figure 4 montre   galement que la relation entre le profil de motivation et la performance est plus affirm  e pour les groupes exp  rimental et actif que pour le groupe t  moin.

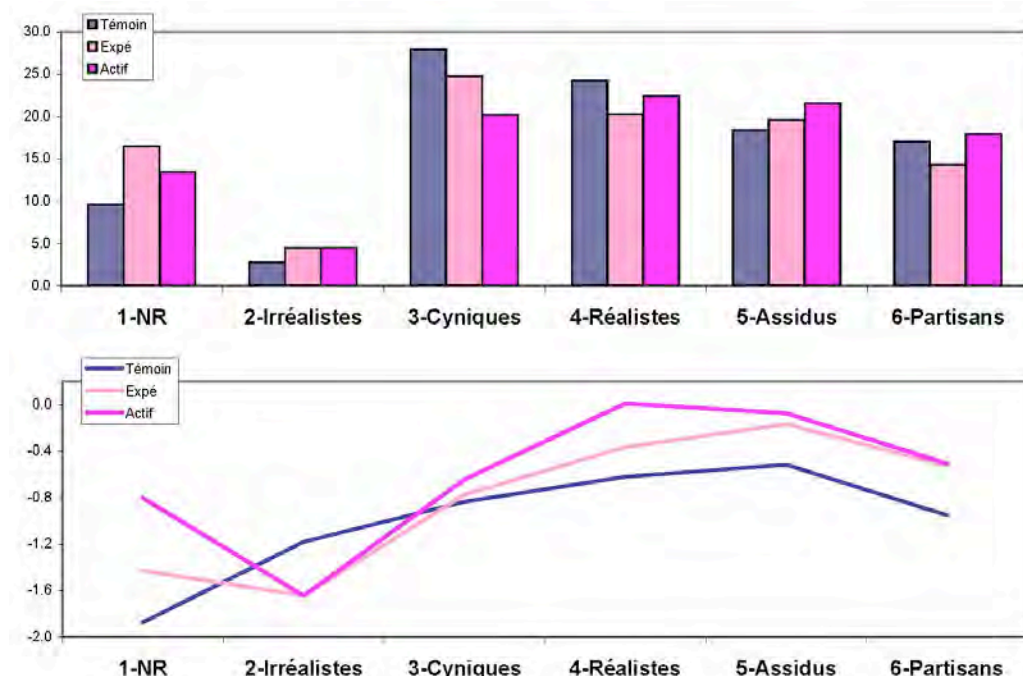


Figure 4 : Effort relatif et score selon le groupe d'  tablissements

4.3 R  sultats pour d'autres tests

Nous avons introduit le thermom  tre d'effort dans plusieurs   valuations nationales r  alis  es en 2011 :   valuation socle commun    l'  cole et au coll  ge, exp  rimentation « internat d'excellence » et exp  rimentation « enseignement int  gr   de science et technologie » (EIST). L'analyse des r  sultats de ces   valuations nous permet de juger de la variabilit   de cette mesure vis-  -vis des disciplines   valu  es, des niveaux scolaires et de certaines caract  ristiques des   l  ves.

L'effort fourni au test semble avoir un lien avec l'  ge des   l  ves. Selon les r  sultats de l'exp  rimentation « internat d'excellence », l'effort fourni aux   preuves de fran  ais et de math  matiques diminue avec l'  ge des   l  ves (figure 5).

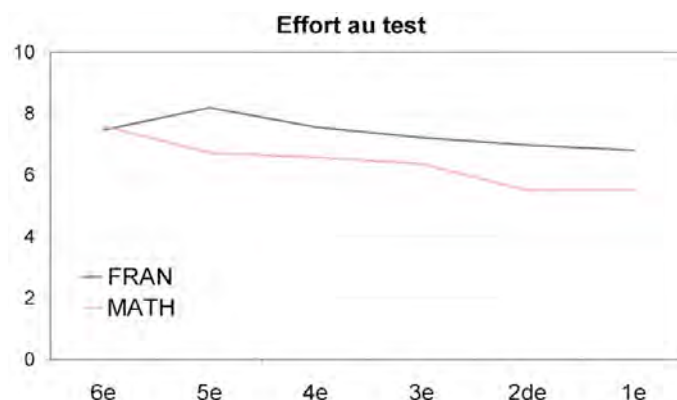


Figure 5 : Effort fourni aux   preuves de fran  ais et de math  matiques, internat d'excellence 2011

En outre, l'  ge ou le niveau scolaire des   l  ves joue   galement sur la relation entre l'effort fourni et le score. Les r  sultats de l'  valuation socle commun montrent qu'en fin de coll  ge, les scores obtenus en fran  ais (comp  tence 1) et en math  matiques (comp  tence 3) augmentent lorsque l'effort relatif augmente. En revanche, le lien entre la performance et l'effort relatif est nettement moins lin  aire en fin de primaire (figure 7).

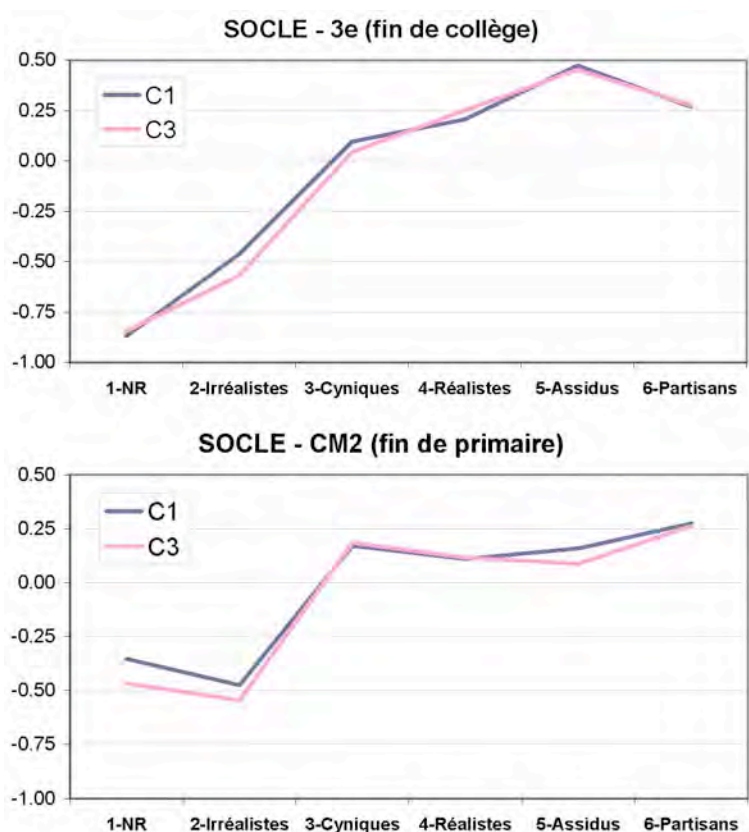


Figure 6 : Score en fonction de l'effort relatif, socle 2011

On note aussi des diff  rences entre les filles et les gar  ons et notamment une moindre motivation de la part des gar  ons (tableau 5).

	Diff��rences filles-gar��ons (standardis��es)	
	Test standardis��	Epreuve not��
PISA	0,13	0,09
Socle ��cole	0,22	0,09
Socle coll��ge	0,27	0,16
EIST	0,20	0,02

Tableau 5 : Diff  rences filles-gar  ons

5. Conclusion

L'exp  rimentation sur la motivation men  e    partir de PISA a donn   des effets relativement modestes. En termes d'enseignement, l'  chantillon trait   (56   tablissements) est peut-  tre trop restreint pour que l'am  lioration de la motivation et de la participation se traduise effectivement en am  lioration de la performance. Le fait que la campagne d'information et de sensibilisation ait plus influenc   le comportement des lyc  es que celui des coll  ges implique un effort particulier    porter sur les coll  ges. En 2012, la DEPP mettra en place l'  valuation PISA en l'accompagnant d'une campagne d'information (affiches, renforcement pr  alable de la communication autour de cette op  ration)    l'  chelle de tous les   tablissements de l'  chantillon.

L'exp  rimentation a pu mettre en   vidence l'int  r  t d'un indicateur de motivation pour le suivi des diff  rentes op  rations d'  valuation. Elle a aussi amen      revoir le terme « effort » qui pose un probl  me d'interpr  tation en m  langeant l'implication ou la motivation de l'  l  ve avec la difficult   de test.

6. R  f  rences et bibliographie

- Butler J., Adams R. J. (2007), The Impact of Differential Investment of Student Effort on the Outcomes of International Studies, *Journal of Applied Measurement*, 8(3), 279-304.
- Ndinga P., Frenette E. (2010), Elaboration et validation de l'Echelle de motivation    bien r  ussir un test (EMRT), *Mesure et   valuation en   ducation*, vol. 33, n   3, 99-123.
- Site internet de PISA : <http://www.pisa.oecd.org>

UN DISPOSITIF INTÉGRÉ POUR L'ÉVALUATION DES ACQUIS DES ÉLÈVES DE 7^E ANNÉE À MADAGASCAR

François-Marie Gerard*

* BIEF – fmg@bief.be

Mots-clés : dispositif, compétences, situation, évaluation

Résumé. Dans le cadre de la mise en œuvre du Plan « Éducation pour Tous », le Ministère de l'Éducation Nationale de Madagascar expérimente une réforme structurelle – le primaire est passé de 5 ans à 7 ans – et pédagogique – de nouveaux curriculums ont été élaborés, selon une approche par les situations. Pour évaluer les acquis des élèves de 7^e année issus de cette réforme, un dispositif d'évaluation a été élaboré et expérimenté. Ce dispositif doit permettre d'évaluer, tant de manière formative que certificative, l'ensemble des objectifs visés dans le cadre de la réforme pédagogique. Il est structuré autour de quatre modalités d'évaluation complémentaires les unes aux autres : l'évaluation classique des ressources, l'évaluation par situation complexe, la confrontation à une situation réelle à problèmes et des mini-projets. Ces quatre modalités d'évaluation se situent dans un continuum selon certaines dimensions essentielles de l'apprentissage et de l'évaluation.

1. Introduction

Dans le cadre de la mise en œuvre du Plan *Éducation pour Tous*, le Ministère de l'Éducation Nationale de Madagascar a entrepris une réforme du système éducatif. Dès 2003, des travaux – accompagnés par le BIEF – ont permis de réorganiser les curriculums pour passer d'une pédagogie par les contenus et les objectifs à une approche par les compétences (APC) au niveau de l'école primaire. Cette approche a été généralisée à toutes les écoles primaires en 2008, avec des résultats très positifs (Gerard & Roegiers, 2011). Depuis 2009, le Certificat d'Études Primaires Élémentaires (CEPE) est délivré sur la base d'un examen évaluant les savoirs et savoir-faire de manière traditionnelle, mais aussi les compétences, à travers des situations complexes.

Depuis l'année scolaire 2008-2009, une nouvelle réforme est expérimentée dans 20 CISCO, sur les 111 circonscriptions scolaires de la République. Dans les écoles concernées, le primaire est passé de 5 ans à 7 ans. De nouveaux curriculums ont été élaborés, selon une approche par les situations (APS). Étant donné la situation politique de Madagascar, la réforme n'a jusqu'à présent pas été étendue à d'autres CISCO, mais les curriculums des 7 années ont été élaborés et sont mis en œuvre dans les classes concernées.

Pour évaluer les acquis des élèves de 7^e année issus de cette réforme, un dispositif d'évaluation a été élaboré et expérimenté par le Service de la Conception, de l'Encadrement et de l'Évaluation (SCEE) de la Direction générale de l'Enseignement fondamental et de l'Alphabétisation (DGEFA), avec le soutien de l'UNICEF et l'appui technique du BIEF.

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar
Texte destiné aux Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

2. Présentation du dispositif d'évaluation

Lorsque les membres du SCEE se sont posé la question de l'évaluation des acquis des élèves de 7^e année, ils ont rapidement pris conscience de la nécessité d'un dispositif varié, car – pour tenir compte des objectifs de la réforme – il fallait évaluer une pluralité d'objets :

- Les changements de comportement face aux situations que les élèves rencontrent, vers plus de responsabilité, d'autonomie, de compétence et d'intégration des valeurs socio-culturelles malagasy ;
- La mise en œuvre, face à une situation réelle, d'une démarche d'analyse, de recueil d'informations, de comparaison et de prise de décision pour améliorer des situations réelles de la vie courante ;
- La mobilisation des ressources nécessaires et pertinentes pour mettre en œuvre cette démarche et améliorer la situation ;
- La maîtrise des ressources qui pourront être mobilisées en situation.

En d'autres termes, le dispositif à élaborer ne devait pas seulement mesurer la quantité de savoirs acquis par l'élève. Il devait aussi mettre en valeur les compétences, et notamment la capacité à analyser une situation concrète, à construire un raisonnement et à adopter un comportement pour résoudre le problème que cette situation pose.

2.1 La définition du profil de sortie

Pour affiner les objets à évaluer, le SCEE – en collaboration avec des représentants d'autres directions et en particulier de la Direction des Curriculums et des Intrants (DCI) – a dégagé un profil de sortie de l'élève de 7^e année, en se fondant sur différents documents¹.

Pour pouvoir poursuivre ses études ou s'insérer dans la vie active de la société, à la fin de la 7 ^e année, l'élève :	
A	Maîtrise au niveau attendu la langue malagasy, de par son statut de langue maternelle et nationale
B	Maîtrise au niveau attendu les deux autres langues officielles : le français et l'anglais
C	Réinvestit les acquis langagiers dans des situations de communication orale et/ou écrite de la vie courante
D	Utilise les connaissances et les aptitudes requises dans les domaines des mathématiques, des sciences, de la technologie, des sciences humaines, des arts et de l'éducation physique et sportive, et les réinvestit dans des situations de la vie courante
E	Manifeste de la créativité, de l'initiative et l'esprit d'équipe face aux tâches à réaliser
F	Utilise une démarche de recueil d'informations, d'analyse, de comparaison et de prise de décision pour améliorer des situations réelles de la vie courante, dans le respect des valeurs civiques, environnementales et culturelles malagasy
G	Agit en fonction de ses droits et de ses devoirs de jeune Malagasy pour améliorer la qualité de sa vie et celle de ses proches
H	Se prend en charge en santé et promeut un comportement préventif face aux maladies transmissibles, au VIH/SIDA, à la toxicomanie pour améliorer la qualité de sa vie et celle de ses proches.

Tableau 1 - Profil de sortie de 7^e année

¹ Les finalités de l'éducation suggérées par la « concertation technique de juin 09 » ; le projet de profil de sortie de 7^e année, défini en février 2008, dans le document « Cadre d'orientation du curriculum » ; le projet de profil de sortie de fin de primaire, défini en juin-juillet 2005, dans le cadre de l'APC.

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar
Texte destiné aux Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

Ce profil a servi de base pour élaborer le dispositif d'évaluation, en permettant de dégager des modalités complémentaires d'évaluation.

2.2 Le dispositif d'évaluation

Les acteurs malgaches ont estimé que tous les énoncés du profil devaient faire l'objet d'évaluations formative et certificative. Pour chaque énoncé, les modalités possibles d'évaluation ont été identifiées, en fonction de la spécificité de l'énoncé. Parmi ces modalités possibles, une modalité préférentielle a été identifiée pour chaque énoncé. Cette modalité préférentielle n'exclut pas les autres, mais elle est celle qui apparaît comme réunissant les meilleurs degrés de pertinence, de validité et de faisabilité pour évaluer l'énoncé concerné selon chaque fonction de l'évaluation. Le dispositif d'évaluation des acquis des élèves de 7^e année est dès lors un ensemble cohérent de modalités et d'outils destinés à évaluer les progrès des élèves dans le cadre de la mise en œuvre de la démarche d'apprentissage.

Le dispositif est structuré autour de quatre modalités d'évaluation complémentaires les unes aux autres :

1. *L'évaluation classique des ressources* porte sur la maîtrise des savoirs et des savoir-faire, dans toutes les disciplines évaluées (Malagasy – Français – Anglais – Mathématiques – Sciences et technologie – Sciences sociales – Arts). Elle concerne avant tout l'évaluation formative des énoncés A, B et D du profil de sortie, mais peut aussi entrer dans l'évaluation certificative. Dans ce cas, elle doit porter sur les ressources principales, considérées comme essentielles. Le plus souvent, elle fait l'objet de questionnaires avec des réponses fermées (V/F, QCM, appariement) ou des réponses ouvertes courtes ou moyennement longues. Elle peut aussi se dérouler oralement.
2. *L'évaluation par situation complexe* porte sur le réinvestissement et la mobilisation des savoirs et savoir-faire pour résoudre une situation délimitée. Elle consiste – dans les 7 disciplines évaluées – à proposer à l'élève une situation complexe, c'est-à-dire nécessitant de combiner plusieurs ressources pour pouvoir la résoudre. La situation proposée est construite par l'évaluateur qui s'efforcera de la rendre la plus proche possible de la vie courante. Elle est particulièrement pertinente pour évaluer la maîtrise des « catégories d'action » ou des « compétences ». Face à la situation, l'élève doit faire une production elle-même complexe qui nécessite le recours à des critères et des indicateurs pour son appréciation. Elle est utile pour l'évaluation formative des énoncés C et D du profil de sortie, mais permet aussi l'évaluation certificative des « catégories d'action » essentielles.
3. *La confrontation à une situation réelle à problèmes* porte sur la mise en œuvre de la démarche d'analyse telle qu'elle est travaillée dans le cadre de l'APS. La situation devrait être une situation réelle issue de l'environnement des élèves et qui présente un certain nombre de problèmes. Face à cette situation, les élèves vont mettre en œuvre la démarche préconisée en l'analysant, en recherchant de l'information, en la comparant avec des situations passées, en la reliant à de nouvelles technologies ou à de nouveaux modes de pensée et en proposant des solutions. Pour réaliser cette démarche, ils devront mobiliser des ressources pertinentes. Ce qui est évalué est essentiellement la mise en œuvre de la démarche, mais aussi ses produits. Le traitement de la situation est interdisciplinaire, mais peut déboucher sur des informations disciplinaires. La confrontation à une situation réelle à problèmes sera surtout utilisée pour l'évaluation formative des énoncés F, G et H du profil de sortie, soit pour identifier les éléments nouveaux qui méritent d'être traités ultérieurement (orientation), soit pour identifier les erreurs commises par les élèves et les corriger (régulation). L'auto- et la co-évaluation peuvent être mobilisées.
4. Les *mini-projets* sont une occasion pour les élèves de vivre à travers le déroulement d'un projet qu'ils ont choisi une démarche d'exploration, de compréhension, de résolution de problèmes en vue d'un résultat concret et tangible. Ces mini-projets sont réalisés en partie en petits groupes, permettant une évaluation collective, mais doivent comporter également un travail individuel. Ils permettent d'évaluer la mise en œuvre de la démarche, la mobilisation

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar

Texte destiné aux Actes du 24^e colloque de l'Admécé-Europe

L'évaluation des compétences en milieu scolaire et en milieu professionnel

et la maîtrise de ressources ainsi que les attitudes des élèves². C'est la modalité privilégiée pour l'évaluation certificative finale de l'ensemble des objectifs définis dans le profil de sortie de l'élève de 7^e année, mais en particulier les énoncés E et F, en combinaison avec les autres modalités. Pour préparer cette évaluation finale, il est indispensable que les élèves vivent deux ou trois mini-projets durant l'année scolaire dans une perspective d'évaluation formative, avec une part importante d'auto- et de co-évaluation.

2.3 Tableau récapitulatif des modalités d'évaluation

Le tableau suivant reprend les quatre modalités du dispositif d'évaluation en mettant en évidence un certain nombre de caractéristiques.

	Évaluation classique des ressources	Évaluation par situation complexe	Confrontation à une situation réelle à problèmes	Mini-projets
Objets d'évaluation	Savoirs et savoir-faire (S et SF)	Mobilisation des S et SF pour résoudre une situation construite, délimitée et proche de la vie « Compétences » ou Catégories d'action	Mise en œuvre de la démarche d'analyse (APS)	Mise en œuvre d'une démarche d'exploration, de compréhension, de résolution de problèmes en vue d'un résultat concret et tangible Attitudes (savoir-être)
Objet d'observation	Produit surtout	Produit surtout	Processus surtout	Processus et produit
Types d'évaluation	Formative avant tout Certificative (sur les S et SF essentiels)	Formative avant tout Certificative (sur les catégories d'action essentielles)	Formative (orientation par identification des S et SF à travailler, régulation par identification et remédiation des erreurs)	Formative Certificative
Moments d'évaluation	Après (ou pendant) les apprentissages des S et SF	Après (ou pendant) les apprentissages des S et SF et de leur intégration	Avant les apprentissages de S et SF Après l'apprentissage de la démarche APS	À la suite d'une confrontation à une situation réelle à problèmes
Type de recueil d'informations	Dispersé : différentes questions sans lien entre elles	Intégré : présentation d'une situation complexe construite à résoudre, avec tous les supports	Confrontation à une situation réelle issue de l'environnement des élèves et qui présente un certain nombre de problèmes	Réalisation du mini-projet
Type de production de l'élève	QCM – V/F – Appariement – Réponses courtes ou moyennes	Production complexe pour résoudre la situation	Mise en œuvre des différentes étapes de la démarche APS	Production en fonction du mini-projet
Types de critères d'évaluation	Exactitude de la réponse	Pertinence de la production Utilisation correcte des outils de la	Pertinence des données collectées Pertinence de la problématique	Pertinence du mini-projet Respect de la démarche de projet

² Cette évaluation des attitudes n'est pas sans poser de difficultés. Voir notamment Gerard (2011).

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar
Texte destiné aux Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

		discipline Cohérence de la production	Cohérence des savoirs mobilisés avec le problème Fiabilité de la solution	Pertinence des ressources Créativité Participation au groupe Écoute et respect des autres
Portée de l'évaluation	Individuelle	Individuelle	Collective	Collective et individuelle

Tableau 2 - Tableau récapitulatif des modalités d'évaluation

2.4 Analyse comparative des modalités

Les quatre modalités d'évaluation sont complémentaires les unes aux autres. Elles se situent dans un continuum selon certaines dimensions essentielles de l'apprentissage et de l'évaluation :

Évaluation classique des ressources	Évaluation par situation complexe	Confrontation à une situation réelle à problèmes	Mini-projets
Individuel	—————→		Collectif
Ponctuel	—————→		Global
Simple	—————→		Complexe
Contrôlé	—————→		Autonome
Décontextualisé	—————→		Contextualisé
Disciplinaire	—————→		Interdisciplinaire

Figure 1 - Analyse comparative des modalités

Les quatre modalités d'évaluation peuvent aussi être rangées dans le même ordre, mais de manière contrastée, en fonction des trois qualités essentielles de toute évaluation : la pertinence, la validité et la fiabilité. En prolongement de ce que De Ketele & Gerard (2005) ont pu mettre en évidence, on peut montrer que – dans le cadre de la démarche pédagogique expérimentée par la réforme du système éducatif malgache – les modalités les plus pertinentes sont la confrontation à une situation réelle à problèmes et les mini-projets, mais aussi que ces modalités posent des difficultés en termes de validité et de fiabilité. Pour réduire ces difficultés de validité et de fiabilité, qui sont propres à de nombreux outils ouverts d'évaluation, il convient de déterminer ce que Gerard (2006) a appelé des « clés de fermeture » - notamment grâce à des grilles d'observation construites sur des critères et des indicateurs précis – afin de bien cadrer ces modalités, de définir leur portée et le sens qu'elles apportent à l'évaluation.

3. Utilisation du dispositif d'évaluation

En juin 2010, une première expérimentation d'un dispositif d'évaluation des acquis des élèves de 7^e année a été réalisée, mais elle portait uniquement sur la modalité « Mini-projets » et avait une portée exclusivement collective. Cette expérimentation a cependant permis de dégager des observations concrètes qui ont aidé l'élaboration plus complète du dispositif.

En juin 2011, le dispositif complet a été testé dans 6 CISCO (sur les 20 CISCO expérimentales). L'objectif de ce test était avant tout de vérifier la validité et la faisabilité du dispositif. À la suite de ce test, un grand travail a été accompli pour améliorer les outils utilisés, notamment en termes de validité et de fiabilité.

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar

Texte destiné aux Actes du 24^e colloque de l'Adméc-Europe

L'évaluation des compétences en milieu scolaire et en milieu professionnel

En juin 2012, le dispositif devrait être utilisé pour réaliser l'évaluation certificative des élèves des 20 CISCO. Dans le même temps, le dispositif devrait permettre de réaliser une étude comparative entre les acquis des élèves de « 7^e année » (ceux de la réforme) et ceux des « classes de 5^e » (ceux des classes suivant les curriculums traditionnels).

Cette étude comparative devrait permettre de mieux analyser et comprendre certains constats qui ont été observés sur la base des résultats obtenus en juin 2011. Notamment, de manière relativement paradoxale par rapport aux travaux réalisés dans le cadre de l'évaluation des compétences (Gerard & BIEF, 2008 ; Gerard & Roegiers, 2011 ; Rey, Carette, Defrance & Kahn, 2003 ; Roegiers, 2004 ; Scallon, 2004), il est apparu que les résultats les moins performants ont été obtenus pour la modalité « Évaluation classique des ressources » (M1), alors que les résultats pour la modalité « Évaluation par situation complexe » (M2) étaient meilleurs. Assez logiquement par contre, on a pu observer les meilleurs résultats pour l'évaluation par « Mini-projets » (M4), comme le montre le graphique suivant.

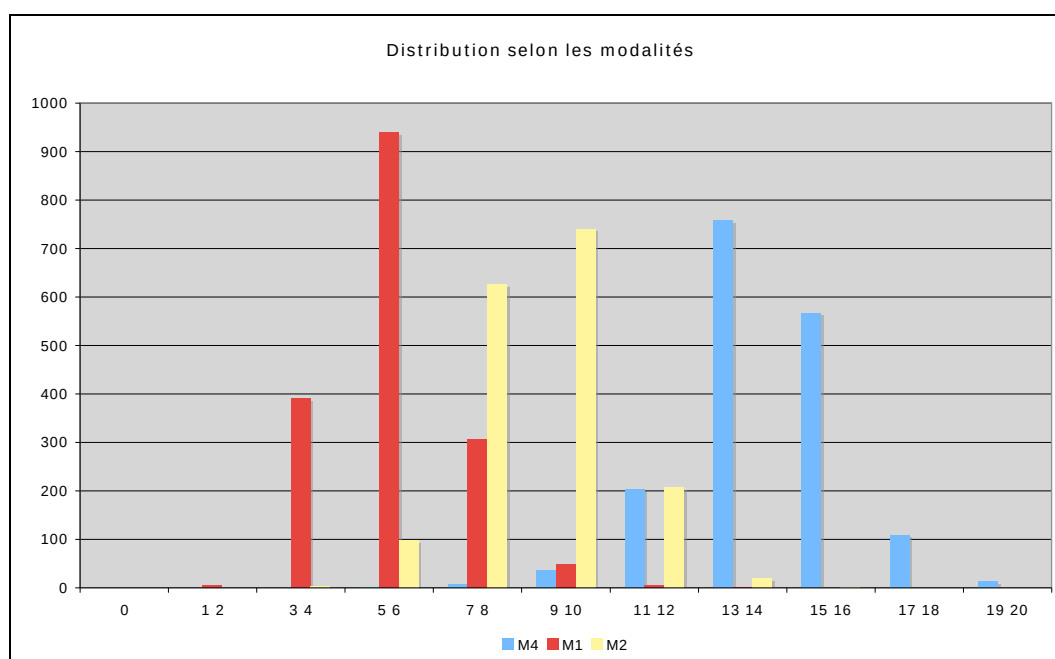


Figure 2 - Distribution des résultats selon les modalités, lors de l'expérimentation du dispositif

L'opération de juin 2011 visait l'expérimentation du dispositif. Sur la base de celle-ci et des informations recueillies, il n'est pas possible d'analyser ces résultats paradoxaux de manière suffisamment fine et valide. L'étude comparative entre la population concernée par la réforme pédagogique et un échantillon témoin composé d'écoles non impliquées dans la réforme devrait permettre de recueillir l'information nécessaire pour mieux analyser ces éléments.

4. Conclusion

La principale qualité du dispositif d'évaluation des acquis des élèves de 7^e année de Madagascar réside vraisemblablement dans la diversité des modalités et des outils d'évaluation mis en place. Face à une réforme pédagogique qui vise le développement de différentes dimensions – depuis l'acquisition de savoirs et de savoir-faire typiquement scolaires jusqu'à des changements de comportement en termes d'autonomie, de responsabilité et de citoyenneté en passant par le développement d'une méthode d'analyse de situations réelles à problèmes afin de leur apporter des

Un dispositif intégré pour l'évaluation des acquis des élèves de 7^e année à Madagascar
Texte destiné aux Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

améliorations – les acteurs malgaches ont choisi de n'éluder aucune de ces dimensions, mais d'élaborer des instruments permettant de les prendre chacune en compte.

Inévitablement, le dispositif est complexe. Il nécessite des investissements importants non seulement pour la conception des outils, mais aussi pour la formation des enseignants qui sont appelés à les mettre en œuvre. Dans un pays où les difficultés de tout ordre ne manquent pas, le dynamisme et l'engagement des acteurs ne suffisent sans doute pas pour déboucher sur des outils parfaitement pertinents, valides et fiables. Néanmoins, l'utilisation du dispositif sur le terrain a pu montrer non seulement sa capacité à évaluer les acquis des élèves, mais aussi à participer à la formation des enseignants et par là à celle de leurs élèves.

5. Références et bibliographie

- De Ketele, J.-M. & Gerard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences, *Mesure et Évaluation en Éducation*, Volume 28, n°3, 1-26.
- Gerard, F.-M. (2006). Les outils ouverts d'évaluation, ou la nécessité de clés de fermeture, in G. Figari & L. Mottier Lopez (Ed.). *Recherche sur l'évaluation en éducation - Problématiques, méthodologies et épistémologie* (pp. 147-154). Paris : L'Harmattan.
- Gerard, F.-M. (2011), L'apprentissage et l'évaluation des attitudes et des savoir-être, in X. Roegiers. *Curricula et apprentissages au primaire et au secondaire – La Pédagogie de l'Intégration comme cadre de réflexion et d'action* (pp. 146-150). Bruxelles : De Boeck.
- Gerard, F.-M. - BIEF. (2008, 2^e édition 2009). *Évaluer des compétences – Guide pratique*. Bruxelles : De Boeck.
- Gerard, F.-M. & Roegiers, X. (2011), Currículo e Avaliação: ligações que nunca serão suficientemente fortes, in M.P. Alves, & J.-M. De Ketele. (Dir.). *Do Currículo à avaliação, da avaliação ao currículo* (pp. 143-158). Porto : Porto Editora.
- Rey, B., Carette, V., Defrance, A. & Kahn, S. (2003). *Les compétences à l'école — Apprentissage et évaluation*. Bruxelles : De Boeck.
- Roegiers, X. (2004, 2^e édition 2010). *L'école et l'évaluation*. Bruxelles : De Boeck.
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par les compétences*. Bruxelles : De Boeck Université.

DETERMINATION DU SEUIL DE MAITRISE POUR L'ATTRIBUTION DES COMPETENCES DU SOCLE COMMUN COMPARAISON DE DIFFERENTES METHODES

Pascal Bessonneau*, Thierry Rocher**

* Ministère de l'éducation nationale – DEPP, pascal.bessonneau@education.gouv.fr

****** *Ministère de l'éducation nationale – DEPP, thierry.rocher@education.gouv.fr*

Mots-clés : création de tests, modèle de réponse à l'item, standard settings, points de césure.

Résumé. Le socle commun de connaissances et de compétences est un référentiel de compétences très variées que tout élève doit maîtriser en fin de scolarité obligatoire en France. Entré en vigueur en 2005, le socle commun occupe une place grandissante dans le système éducatif français : la maîtrise des sept compétences qui le composent est, depuis 2011, nécessaire pour obtenir le diplôme national du brevet (9^{ème} année de scolarité). La direction de l'évaluation, de la prospective et de la performance (DEPP) s'est vue confier la mission de mesurer la maîtrise de ces compétences au moyen de tests standardisés, sur des échantillons représentatifs de plusieurs milliers d'élèves de CM2 (5^{ème} année) et de 3^e (9^{ème} année), dans le but de produire des indicateurs nationaux de suivi temporel des performances des élèves. Cette présentation vise à décrire et à comparer les différentes procédures utilisées pour la détermination des seuils (« standard-setting ») portant sur la maîtrise de la langue française et sur les principaux éléments de mathématiques et la culture scientifique et technologique en 2011.

1. Contexte général

1.1 Le socle commun de connaissances et de compétences

Le socle commun de connaissances et de compétences (SOCLE) été créé en 2005. C'est un standard de compétences et de connaissances minimales requises à la fin de la scolarité obligatoire :

« La scolarité obligatoire doit au moins garantir à chaque élève les moyens nécessaires à l'acquisition d'un socle commun constitué d'un ensemble de connaissances et de compétences qu'il est indispensable de maîtriser pour accomplir avec succès sa scolarité, poursuivre sa formation, construire son avenir personnel et réussir sa vie en société »

Les attestations de maîtrise sont délivrées par les équipes enseignantes au sein d'un livret personnel de compétences.

Depuis 2011, une nouvelle étape a été franchie puisque la maîtrise du SOCLE est nécessaire pour obtenir le diplôme national du brevet (D.N.B.).

Le SOCLE se décline selon 7 compétences :

1. La maîtrise de la langue française (C1)
2. La pratique d'une langue vivante étrangère (C2)

3. Les principaux   l  ments de math  matiques et la culture scientifique et technologique (C3)
4. La ma  trise des techniques usuelles de l'information et de la communication (C4)
5. La culture humaniste (C5)
6. Les comp  tences sociales et civiques (C6)
7. L'autonomie et l'initiative (C7)

L'approche p  dagogique est transdisciplinaire. Dans ce sens, il est possible de la rapprocher de l'approche des « Comp  tences Clefs » pour l'Union Europ  enne.

1.2 La Loi organique de la loi de finances

La Loi organique de la loi de finances (LOLF) impose des indicateurs de mesure de l'efficacit   des politiques publiques. Cette mesure de l'efficacit   se fait    travers la d  finition dans la loi d'indicateurs de performance. Par exemple, parmi les crit  res, un crit  re porte sur la performance    PISA (par exemple OCDE, 2011) des   l  ves fran  ais.

Ainsi, avec le SOCLE, doit se mettre en place un processus permettant la proportion d'  l  ves ma  trisant le SOCLE :

- pour chaque comp  tence ;
- les   l  ves    la fin de l'  cole primaire (CM2) et en fin de coll  ge (3^e) ;
- d  cliner selon le secteur de scolarisation¹ ;

Ces indicateurs doivent   tre pertinents pour chacune des strates form  es par les crit  res pr  sent  s ci-dessus. En outre ils doivent garder leur pertinence au niveau national. Enfin la mesure des politiques   ducatives suppose que ces indicateurs doivent   tre objectifs et doivent permettre des comparaisons temporelles.

En France, la DEPP a mis en place depuis 2003, un dispositif d'  valuations standardis  es dont l'objectif est d'assurer le suivi global des performances des   l  ves de CM2 et de 3^e : le « Cycle d'Evaluations Disciplinaire R  alis  es sur Echantillon » (CEDRE). Le lecteur pourra se reporter par exemple    la Note d'information pour des r  sultats r  cents (Bourny et al., 2010). Cependant, ces   valuations sont organis  es selon un cycle de six ans alors que les indicateurs de la LOLF sont annuels. En outre, les   valuations CEDRE adoptent une perspective disciplinaire et non le point de vue du socle commun.

2. L'exp  rimentation des   preuves en fin de primaire (CM2)

La DEPP, comme pour certains indicateurs de la LOLF, a   t   charg  e d'  laborer des   valuations permettant la mesure de la proportion d'  l  ves ma  trisant le SOCLE pour ces diff  rentes populations et comp  tences. Ces   preuves ont   t   r  alis  es en collaboration avec l'Inspection G  n  rale de l'Education Nationale (IGEN).

Afin d'  tre simple    administrer, les   preuves se pr  sentent sous la forme de QCM, au moins pour la premi  re ann  e. D'autres types d'  preuves exp  rimentales viendront s'ajouter en 2011 pour r  pondre aux nouvelles attentes p  dagogiques li  es    la sp  cificit   du SOCLE. Les cahiers utilis  s se pr  sentaient sous la forme de cahiers tournants. Pour la comp  tence C1, 122 items ont   t   test  s et 148 items pour la comp  tence C3.

¹ Quatre grandes cat  gories de type d'  tablissements sont distingu  es : les   tablissements priv  s, les   tablissements publics hors   ducation prioritaire et deux types d'  tablissements relevant de l'  ducation prioritaire (   l'  poque du recueil des donn  es : RRS, r  seau de r  ussite scolaire et RAR –r  seau ambition r  ussite).

Entre le 16 et le 27 mai 2011, 277   coles ont   t   sollicit  es pour participer    cette   valuation. Le taux de r  ponse est   lev   puisque les donn  es finales comptent 262   coles participantes soit 7489   l  ves.

3. D  termination d'un seuil

Les indicateurs de la LOLF portent sur la proportion d'  l  ves qui ma  trisent les comp  tences du SOCLE. Par cons  quent un seuil de ma  trise doit   tre d  fini pour chaque comp  tence,    partir des r  sultats obtenus au test portant sur la comp  tence en question. Les propri  t  s attendus pour ce seuil doivent   tre les m  mes que pour le test : fid  le, valide, stable dans le temps, ... En outre dans notre cas, ce seuil doit pouvoir   tre d  fini ind  pendamment des items choisis parmi les items exp  riment  s, c'est-  -dire que ce seuil doit   tre valide ind  pendamment du cahier administr  .

La d  finition de score seuil a fait l'objet d'une litt  rature importante notamment dans la litt  rature anglo-saxonne. Le processus visant    d  finir un score seuil est qualifi   en anglais de « standard setting ». Nous allons illustrer diff  rentes m  thodes de « standard setting » employ  es dans le cadre des indicateurs de la LOLF, en se basant le plus souvent sur les donn  es issues de l'exp  rimentation de l'  preuve du SOCLE C1 en primaire.

2.1 Introduction

La d  termination normative est commune dans le domaine de la sant   (George et al.). Il s'agit,    partir d'une population de r  f  rence, de d  finir un seuil    l'aide d'indicateurs statistiques tels que les quantiles ou l'  cart-type. Dans le domaine de l'  du  m  trie, le seuil peut   tre d  termin      partir des donn  es en fonction par exemple des 15 % d'  l  ves les plus en difficult  .

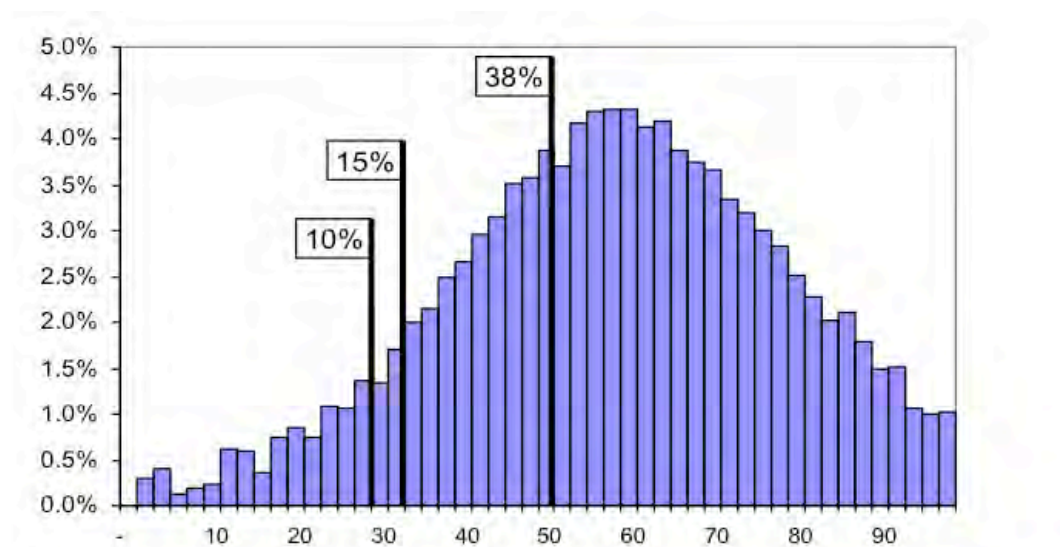


Figure 1. D  termination normative d'un seuil avec diff  rents pourcentages de r  ussite au test

L'inconv  nient de cette m  thode est   vident. La d  cision revient au statisticien (ou    une d  cision politique) ou    un autre acteur qui fixe ce seuil *a priori*. Elle est donc ind  pendante du contenu du test et est subjective. Toutefois c'est une m  thode tr  s simple pouvant pr  senter un int  r  t dans le cadre de comparaisons temporelles comme dans CEDRE ou PISA.

Dans la suite de ce document, il s'agira d'envisager une approche qui croise les exigences et les attentes des p  dagogues avec les r  sultats empiriques observ  s aux items du test. Afin d'assurer la

validit   de ce score seuil au niveau psychom  trique, nous d  crivons le processus d'  troite collaboration entre statisticiens et p  dagogues qui permet d'aboutir    un seuil valide des deux points de vue.

2.2 Donn  es recueillies lors de l'  preuve du SOCLE

Les donn  es recueillies lors des   preuves exp  rimentales du SOCLE en 2011 sont multiples. En effet le but   tait de disposer de plusieurs faisceaux d'informations afin de se placer dans des conditions optimales pour d  cider d'un score seuil. Ainsi nous disposions de trois sources d'information :

3.1.1 Les   l  ves

Les informations portent sur la r  ussite aux items de la population d'  l  ves test  s en 2011 qui concernent sur un   chantillon « repr  sentatif » d'environ 8000   l  ves de CM2. Nous disposons donc d'une information sur les items d'une part et des scores d'autre part. Le score des   l  ves et les param  tres d'items des items sont issus d'une mod  lisation par un mod  le de r  ponse    l'item    2 param  tres. Il est important de noter que seuls les items valides d'un point de vue psychom  trique sont utilis  s dans cette   tude.

3.1.2 Les professeurs des   l  ves   chantillonn  s

Pour chaque   l  ve, le professeur indique si l'  l  ve ma  trise ou non la comp  tence C1 du SOCLE, la comp  tence C3 ou les deux. Un cahier sp  cial contenant une partie des items exp  riment  s et repr  sentatifs des items des cahiers leur est fourni. Une d  cision leur est demand  e pour chaque item (cf. Tableau 1). Enfin deux questions sont pos  es dans ce cahier concernant le minimum et le maximum d'  l  ves qui, selon eux, ma  trisent le SOCLE (cf. p6).

<i>Selon vous, les ��l��ves ma��trisant la comp��tence ...</i>	<i>...doivent absolument r��ussir cet item (A)</i>	<i>...devrait pouvoir r��ussir cet item (B)</i>	<i>...ne doivent pas forc��ment r��ussir cet item (C)</i>
Item 1	X		
Item 2		X	
...	...	...	...
Item N			X
Total	Σ = score minimal		N items - Σ = score maximal

Tableau 1. Importance des items pour la ma  trise du SOCLE (cahier professeur).

3.1.3 Des experts en p  dagogie

Un groupe d'experts d'environ une dizaine de personnes constitu  s par des professeurs, des concepteurs et des inspecteurs g  n  raux sont sollicit  s en premier lieu pour effectuer le travail de jugement sur chaque item test   (cf. Tableau 1) et pour indiquer le pourcentage minimum et maximum d'  l  ves ma  trisant le SOCLE (cf. p6). Enfin ils ont   t   convoqu  s pour finaliser le choix du score seuil.

4. M  thodes envisag  es

Les m  thodes envisag  es sont au nombre de trois :

- M  thode des *contrastes* : comparaison des scores des   l  ves et des attestations donn  es par leurs enseignants
- M  thode des *zones de jugement (Hofstee)* : comparaison r  sultats statistiques et jugements par les experts des items
- M  thode des *marque-pages (bookmarks)* : positionnement d'un seuil dans une liste d'items class  s par ordre de r  ussite d  croissante

Pour la comp  tence C3, une premi  re approche disciplinaire a   t   utilis  e en s  parant les donn  es sur les items de math  matiques et de sciences afin de faciliter le travail des enseignants et des juges. N  anmoins les attentes portaient sur une seule proportion de r  ussite pour cette comp  tence. En outre l'analyse du test C3 (sciences et math  matiques) montrait une unidimensionnalit   suffisante.

Pour une revue des m  thodes de d  termination d'un ou de plusieurs scores de c  sure, le lecteur est invit      se reporter par exemple    : Bunch M. B. et al., 2007. Cizek G. J. et al., 2004. Rupp A. A. et al., 2009.

4.1 M  thodes des contrastes

Cette m  thode consiste    confronter les r  sultats obtenus au test avec le jugement des enseignements sur les   l  ves. La figure 2 illustre cette m  thode essentiellement graphique.

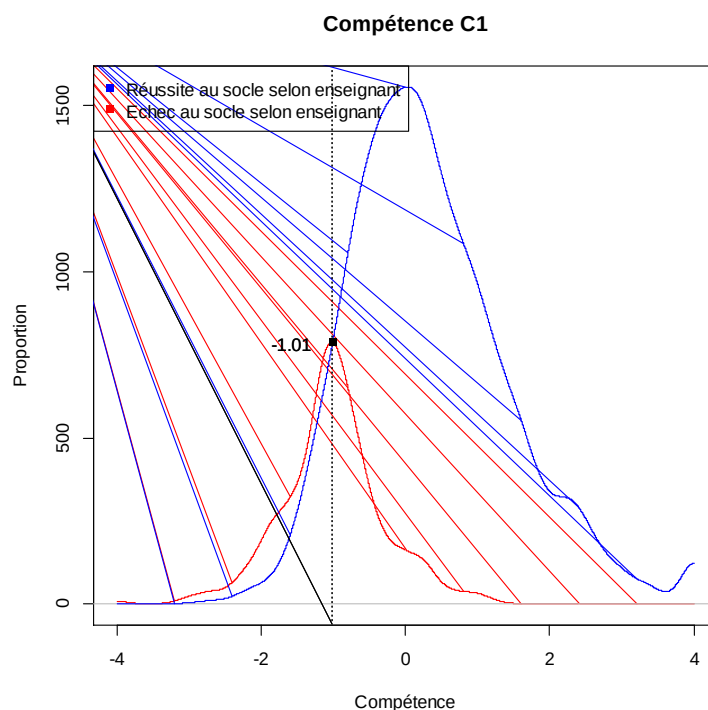


Figure 2. M  thode des contrastes pour la comp  tence C1 du SOCLE.

Les deux distributions des scores se croisent en un point qui d'une part minimis   les faux n  gatifs et les faux positifs. Ici ce point de croisement est positionn      une habilet   de -1,01².

² L'unit   est l'habilet   calcul  e par le mod  le de r  ponse    l'item. Elle est approximativement centr  e en z  ro et d'  cart-type 1.

Toutefois on peut noter que les distributions ne sont pas optimales. Ainsi la proportion d'  l  ves pour lesquels l'enseignant n'a pas attribu   le SOCLE et qui « r  ussissent », c'est-  -dire qu'ils ont une habilet   sup  rieure au point de c  sure est importante.

Dans le cas de la figure 3, il appara  t nettement que le test est un indicateur relativement pauvre de la d  cision du professeur d'attribuer ou non le SOCLE.

4.2 M  thodes des zones de jugement (m  thode d'Hofstee)

Le principe de cette m  thode est de confronter toujours graphiquement les informations donn  es par les professeurs    la courbe empirique cumul  e de r  ussite au test.

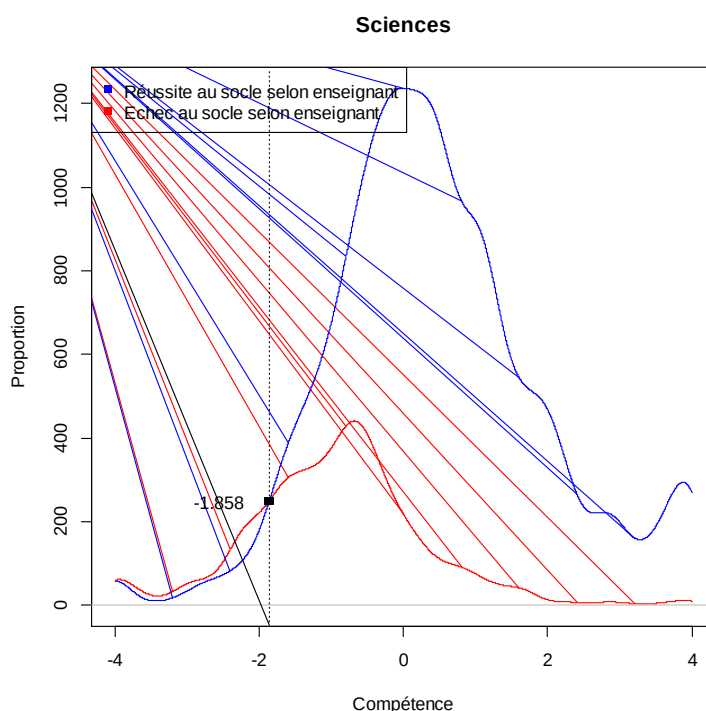


Figure 3. M  thode des contrastes pour la comp  tence en sciences

Le principe de cette m  thode est r  sum   sur la figure 4. Les deux questions formul  es pour d  finir le pourcentage minimum et maximum d'  l  ves ma  trisant le SOCLE sont les suivantes :

- « Sur la base de votre exp  rience professionnelle, quels sont, selon vous, le pourcentage minimum et le pourcentage maximum d'  l  ves ma  trisant la comp  tence ? »
- « Quel est le score – en terme de nombre de bonnes r  ponses – au-del   duquel les   l  ves peuvent   tre consid  r  s comme ma  trisant la comp  tence ? Plus pr  cis  ment, il s'agit de d  terminer un score minimal « acceptable » correspondant    une notation large et un score maximal « acceptable » correspondant    une notation s  v  re.

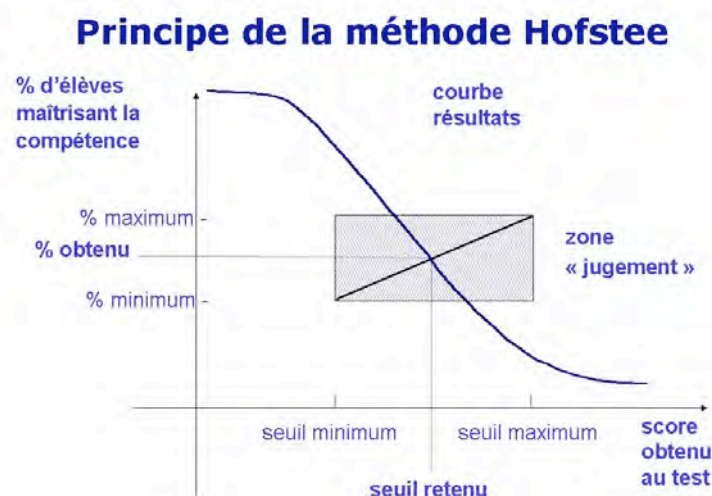


Figure 4. Principe de la m  thode d'Hofstee

Alors que la m  thode d'Hofstee utilise dans la litt  rature plut  t un jugement bas   sur le test (le score minimum et maximum au test), les scores minimaux et maximaux ont   t   d  finis    partir d'un jugement sur les items composant le test. En effet il nous semblait plus ad  quat,   tant donn   la structure en cahiers tournants et d'autre part plus fiable, de demander aux professeurs leur jugement sur les items plut  t que sur un cahier    mod  le   .

Le score minimum est calcul   en sommant le nombre d'items class   (A) et le score maximum en calculant le nombre d'items class  s (A) et (B). Au total 263 enregistrements sont utilisables.

A l'aide de ces r  sultats, nous pouvons tracer la courbe empirique cumul  e les droites : les droites coupant la courbe en bleu et les courbes ne la croisant en rouge sur la figure 5.

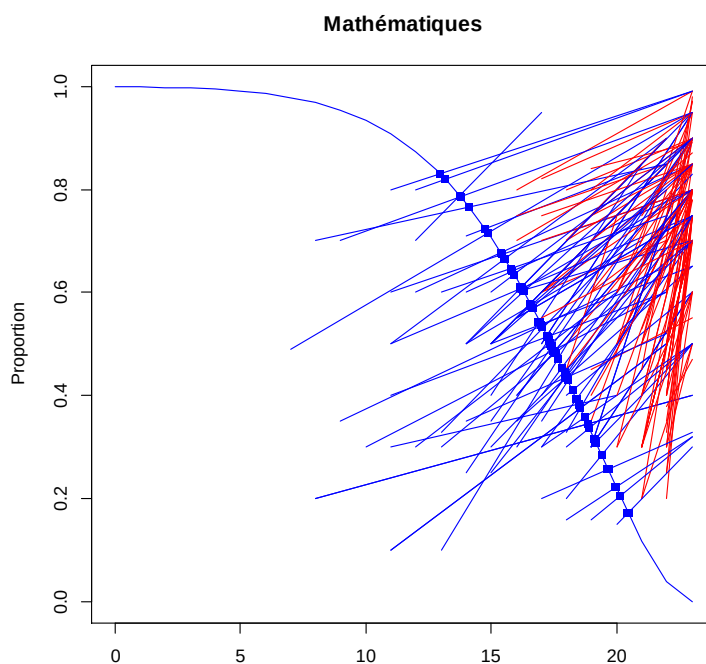


Figure 4. M  thode d'Hofstee pour les math  matiques (sous comp  tence de C3)

3.4 Marque-pages (bookmarks)

Depuis la g  n  ralisation de l'utilisation des mod  les de r  ponse    l'item, la m  thode des *marque-pages* s'est souvent impos  e dans les pratiques. Elle est issue d'une vision des « standards settings » o   le jugement des experts se porte sur les items. Elle se situe donc dans la lign  e des m  thodes dites de « Angoff » tr  s usit  es avant l'utilisation des mod  les de r  ponse    l'item.

Dans cette m  thode, un groupe « flottant » correspond    des   l  ves « fragiles » est pris comme r  f  rence. L'objectif est que les juges d  terminent le niveau de difficult   des items qui caract  risent le mieux les   l  ves de ce groupe « flottant ».

Pour le socle, une premi  re   tape a   t   de demander aux experts de d  finir une zone de 5 items au maximum. Pour les items de cette zone, les   l  ves « fragiles » doivent avoir une r  ussite de 50 % aux items.

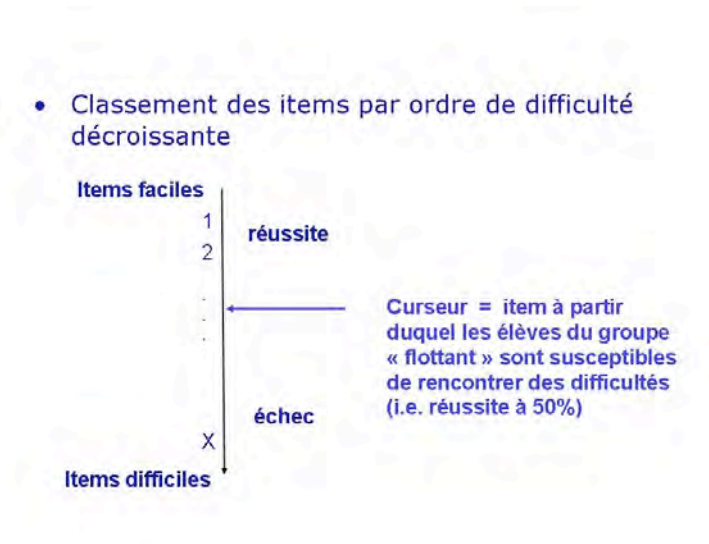


Figure 5. M  thode des marque-pages

Puis dans une seconde phase, afin d'obtenir un consensus, des informations sur le choix de la zone de d  cision pour chaque juge est donn  e anonymement ainsi que le pourcentage d'  l  ve ma  trisant le SOCLE en utilisant chacun des points de c  sure propos  . Pour la comp  tence C3, un soin particulier est port   sur l'alignement entre les maths et les sciences. Au final, la zone retenue pour la comp  tence C1 am  ne    un pourcentage d'  l  ves ma  trisant le SOCLE de 80 % et pour la comp  tence C3    un pourcentage de 72 %.

5. Conclusion

Les r  sultats obtenus au final pour chaque m  thode sont les suivants sont figur  s dans le tableau 6.

M��thode	Intervenants	R��ussite au SOCLE
Contrastes	Professeurs	83 %
Hofstee	Professeurs	66 %
Hofstee	Experts	62 %
Marque-pages	Experts	80 %

Tableau 5. Comparaison des r  sultats pour la comp  tence C1

La d  finition des points de c  sure et le calcul des proportions de r  ussite selon les   tablissements montre une certaine disparit   par rapport aux jugements des enseignants illustrant une certaine disparit   dans le processus de d  cision des enseignants.

	C1		C3	
	Test	Attestations	Test	Attestations
Public hors EP	81,4	80	75	79,4
RRS	64,3	66,6	51,2	66,3
RAR	59,9	64,7	46	61,9
Priv��	84,3	86,9	72,8	86,3
Ensemble	79,5	79,3	71,5	78,6

Tableau 4. Points de c  sure obtenus par type d'  tablissements

Pour mai 2012,    partir des informations collect  es, les items du test final seront choisis afin de maximiser la pr  cision sur la proportion de r  ussite obtenue. En outre des items additionnels, exp  rimentaux seront ajout  s (r  ponse ouverte longue, ...).

Ce travail ouvre la voie vers le m  me type de d  marche pour les autres comp  tences du SOCLE. Toutefois ce travail a pu mettre en lumi  re la fragilit   des informations fournis par les enseignants, en particulier en 3^e o   de tr  s nombreuses classes n'  taient compos  es que d'  l  ves ma  trisant le SOCLE, selon l'attestation des enseignants. Cela renforce le doute qui p  se sur la qualit   des attestations de validation du SOCLE actuellement   tablies par les enseignants (Daussin et al., 2010).

D'un point de vue m  thodologique, la m  thode d'Hofstee a pos   des difficult  s car un nombre non n  gligeable des zones de jugement n'incluait pas la courbe empirique. Par cons  quent la m  thode de Beuk sera peut   tre privil  gi  e dans l'avenir car elle pose moins de probl  me de ce type.

En conclusion, malgr   certains   cueils   voqu  s, les r  sultats fournis par les enseignants   taient int  ressants pour aider les juges dans la phase finale de d  termination du point de c  sure. Notamment le premier tour lors d'une session de « marque-page » montre une certaine disparit   naturelle entre les juges o   toutes les informations suppl  mentaires sont appr  ciables pour aboutir au cours des tours ult  rieurs    un consensus.

Des   tudes ult  rieures vont   tre men  es afin de tenter d'expliquer les diff  rences observ  es par types d'  tablissement, par professeurs, ...

6. R  f  rences et bibliographie

Bourny G., Bessonneau P., Daussin J.-M et Keskpaik S. (2010), *L'  volution des comp  tences g  n  rales des   l  ves en fin de coll  ge de 2003    2009*, Minist  re de l'  ducation nationale, Direction de l'  valuation, de la prospective et de la performance.

Bunch M. B., Cizek G. J. (2007), *Standard Settings*, SAGE.

Cizek G. J., Bunch M.B. et Koons H. (2004), *Setting Performance Standards : Contemporary Methods*. NCME.

Daussin J.-M., Rocher T. et Trosseille B. (2010), *L'attestation de la ma  trise du socle commun est-elle soluble dans le jugement des enseignants ?*, Educations et formations, n  79.

George S., Sayeed Haque M. et Oyeboode F. (2006), *Standard setting: Comparison of two methods*, BMC Medical Education, Vol 6. n  46.

OCDE (2011), *R  sultats du PISA 2009 : Savoirs et savoir-faire des   l  ves – Performance des   l  ves en compr  hension de l'  crit, en math  matiques et en sciences*, Editions OCDE.

Rupp A. A., Vock M. and Harsch C. (2009), *The Development, Calibration, and Validation of Standards-based Tests for English as a First Foreign Language at the IQB*, IQB. Humboldt University Berlin.

EVALUER DES COMPETENCES DANS UNE OPTIQUE SITUATIONNELLE : UN GAGE DE VALIDITE ? L'EXPERIENCE DE DEVELOPPEMENT D'OUTILS D'EVALUATION DES COMPETENCES EN HISTOIRE AU SECONDAIRE EN BELGIQUE FRANCOPHONE

Jean-Louis JADOULLE*

Université de Liège, Service de Didactique Spéciale de l'Histoire, ljadoulle@ulg.ac.be

Mots-clés : compétence, évaluation, familles de situations, validité, fiabilité

Résumé. *L'expérience d'élaboration des outils d'évaluation des compétences en histoire telle qu'elle a été conduite dans le cadre d'une recherche « technologique » qui a été réalisée en 2000-2001, puis prolongée au sein de la Commission inter-réseaux des outils d'évaluation en histoire, tend à montrer que la définition de familles de situations permet d'assurer la validité des situations d'évaluation des compétences des élèves. La définition de critères et d'indicateurs constitue quant à elle un moyen d'assurer, autant que faire se peut, la fiabilité de l'évaluation.*

La présente contribution s'organise autour d'une triple préoccupation. Il s'agit d'abord de présenter la teneur des outils d'évaluation des compétences en histoire qui ont été élaborés, sous notre direction, au sein de l'Unité de Didactique et de Communication en Histoire de l'Université catholique de Louvain ainsi que le processus de recherche qui y a mené. Une attention particulière sera également portée à l'approche situationnelle sur laquelle ce travail d'élaboration d'outils d'évaluation a pris appui. Nous tenterons de montrer en quoi le choix de cette approche constitue un gage de validité au plan de l'évaluation des productions des élèves. Enfin, nous expliciterons les éléments qui ont été mis en place pour assurer la fiabilité de l'évaluation.

1. Quelques éléments de contexte

Initiée en 1997 suite au vote, par le Parlement de la Communauté française de Belgique du « Décret-Missions »¹, la réforme des programmes d'enseignement de l'école secondaire² est entrée en application en 2001-2002 en classes de 1^{ère}, 3^{ème} et 5^{ème} années, en 2002-2003 en 2^{ème}, 4^{ème} et 6^{ème} années. Parmi les attendus de cette réforme, il faut souligner le projet de développer les « compétences » des apprenants. Le « Décret-Missions » définit la compétence comme l'« aptitude à mettre en œuvre un ensemble organisé de savoirs, de savoir-faire et d'attitudes permettant d'accomplir un certain nombre de tâches »³. La compétence est donc vue comme une habileté intégrative qui impose que l'élève mobilise un ensemble de ressources préalablement apprises. Cette définition rejoint celle de nombreux auteurs dont, en Communauté française de Belgique, celles de Roegiers et De Ketele (2000), Beckers (2002) et encore Roegiers (2010a et b).

Dès avant l'entrée en application de la réforme, les enseignants d'histoire ont manifesté leur perplexité par rapport à la manière d'évaluer les compétences des élèves. Cette perplexité allait de pair avec un grand scepticisme quant à la probabilité que la réforme leur apporte un surcroît d'efficacité sur le plan de l'évaluation des performances des élèves. Ainsi, en avril 2001, soit cinq mois avant l'entrée en vigueur des nouveaux programmes, nous avons interrogé l'ensemble des enseignants d'histoire, de français, de sciences, de mathématiques et d'« étude du milieu »⁴ travaillant dans les établissements affiliés à la Fédération de

¹ Décret définissant les Missions prioritaires de l'Enseignement Fondamental et de l'Enseignement Secondaire et organisant les Structures propres à les atteindre, 24 juillet 1997.

² Sur les antécédents historiques et le contexte qui a mené à cette réforme, voir JADOULLE, J.-L. (2011a).

³ Décret définissant les Missions prioritaires de l'Enseignement Fondamental et de l'Enseignement Secondaire et organisant les Structures propres à les atteindre, 24 juillet 1997, p. 90.

⁴ Dans l'enseignement « libre confessionnel », au premier degré (1^{ère} et 2^{ème} années), l'histoire est intégrée dans un cours dit d'« étude du milieu » dont l'objet est l'étude du mode de vie des hommes, aujourd'hui, en Occident. Cette étude doit porter sur six activités essentielles (habiter, produire, consommer, se cultiver, circuler, vivre en société) et être menée dans une optique interdisciplinaire (histoire, géographie, sciences sociales et économiques).

l'Enseignement secondaire catholique⁵. Ces enseignants avaient pu prendre connaissance des programmes qu'ils avaient la responsabilité de mettre en œuvre à partir du mois de septembre. Le questionnaire qui leur a été administré visait à cerner leur position par rapport aux nouveaux curricula et leur motivation à les mettre en œuvre (Jadoulle, J.-L. et Letor, C. [2002]). Un des items qui leur était soumis portait sur le gain qu'ils escomptaient au plan de l'évaluation des performances des élèves.

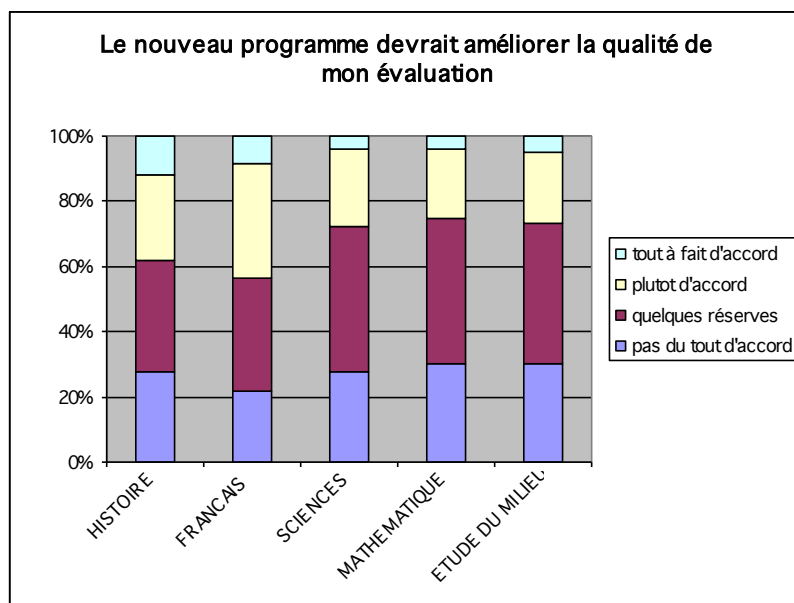


Figure 1 : Gain escompté par les enseignants au plan de l'évaluation des performances des élèves
(D'après Jadoulle, J.-L. et Letor, C. [2002]).

D'une manière générale, la majorité des enseignants se montraient pour le moins réservés à l'idée que leurs pratiques d'évaluation puissent évoluer positivement, à la faveur des nouveaux curricula. Parmi les enseignants d'histoire, 59,05 % des répondants étaient soit convaincus que leur évaluation ne devrait pas s'améliorer soit très réservés à ce sujet. Une minorité de répondants (36,17 %) escomptait un bénéfice.

2. Le dispositif de recherche

C'est dans ce contexte que, en 2000-2001, nous avons dirigé, au sein de l'Unité de Didactique et de Communication en Histoire de l'Université catholique de Louvain, une recherche de type « technologique » financée par le Ministère de l'Education et de la Recherche scientifique de la Communauté française. Cette recherche a été réalisée par M. Bouhon et C. Dambroise (Bouhon, M. et Dambroise, C. [2002]). Il s'agissait, pour deux des quatre compétences (tableau 2) dont les enseignants d'histoire avaient la charge, d'élaborer, à destination des professeurs, des « outils d'évaluation ». Ces deux compétences portaient d'une part sur l'énoncé de questions de recherche en histoire et, d'autre part, sur la fabrication d'un outil de communication d'un savoir historique préalablement appris (tableau 2). Le dispositif de recherche (figure 2) était organisé en cinq étapes.

⁵ La Belgique étant un Etat fédéral, l'éducation est gérée par les Communautés qui sont au nombre de trois : la Communauté française, la Communauté flamande et la Communauté germanophone. Ces communautés financent quatre « réseaux d'enseignement ». Le réseau « officiel » est financé et organisé directement par le Ministère de l'Education de la Communauté. En Belgique francophone, en 2009-2010, il scolarisait environ 23,7 % des élèves de l'enseignement secondaire. Le réseau organisé par les provinces et les communes (CEPEONS) est subventionné au plan du fonctionnement et de la rémunération des enseignants par le Ministère de l'Education : il scolarisait, en 2009-2010, environ 16 % des élèves de l'enseignement secondaire. Le réseau « libre confessionnel », organisé par des « pouvoirs organisateurs » réunis au sein de la Fédération de l'Enseignement secondaire catholique (Fesec) et subventionné par le Ministère de l'Education scolarisait, en 2009-2010, environ 60,3 % des élèves de l'enseignement secondaire. Enfin, le réseau « libre non confessionnel », organisé par des « pouvoirs organisateurs » privés mais subventionné au plan du fonctionnement et de la rémunération des enseignants par le Ministère de l'Education.

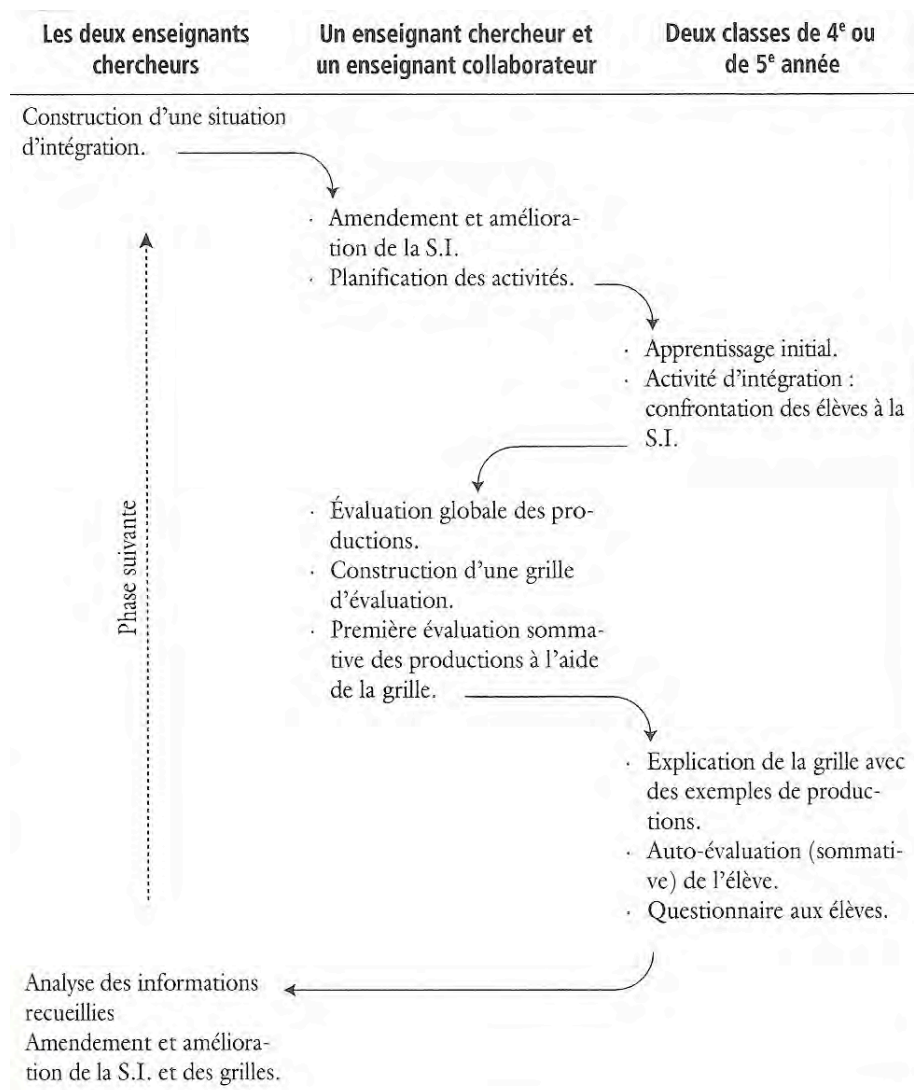


Figure 2 : Description du dispositif de recherche
(D'après Bouhon, M. et Dambroise, C. [2002], p. 47).

2.1 La conception des situations d'intégration et la planification des activités d'apprentissage des ressources

Le travail d'élaboration des situations d'intégration a été réalisé conjointement par les deux chercheurs, également enseignants dans le secondaire, et, pour chacune des deux compétences choisies, par un autre professeur. Par « situation d'intégration », on entend l'activité à travers laquelle les élèves doivent faire la preuve de leur capacité à intégrer ou à mobiliser conjointement les ressources qu'ils ont apprises pour résoudre le type de problème qui définit chacune des deux compétences choisies (tableau 3). Ces situations d'intégration étaient donc, tout à la fois, des situations d'apprentissage et d'évaluation des compétences. La tâche des chercheurs et des enseignants consistait aussi à préparer la séquence d'apprentissage des ressources qui allait se déployer avant la phase d'intégration.

2.2 La mise en œuvre de la séquence et l'évaluation de la maîtrise de la compétence

Cette phase de préparation de la séquence d'apprentissage était suivie d'une phase de mise en œuvre. Celle-ci était organisée en deux temps : d'abord l'apprentissage des savoirs et des savoir-faire dont les élèves devaient avoir la maîtrise en phase d'intégration, ensuite l'exercice de la compétence proprement dite. Pour chaque compétence, la séquence a été mise en œuvre dans deux classes de deux niveaux d'étude différents (4^{ème} et 5^{ème} années) par deux professeurs différents.

2.3 L'analyse du matériel récolté, la fabrication d'une grille d'évaluation et l'évaluation des productions des élèves

La mise en œuvre des situations d'intégration dans les différentes classes a permis de rassembler, pour chaque compétence évaluée, entre 40 et 60 productions d'élèves. Celles-ci ont été examinées par les chercheurs et les enseignants. Dans un premier temps, les professeurs ont été invités à classer de manière intuitive les productions des élèves selon qu'elles leur semblaient d'une qualité très bonne ou bonne, satisfaisante, faible ou insuffisante. Ils ont ensuite été amenés à verbaliser les raisons qui les conduisaient à les apprécier de cette manière. Ce travail a permis de mettre au jour un ensemble d'éléments, lesquels ont été progressivement rassemblés et organisés en critères et en indicateurs (cfr *infra* 3.2). Le travail de verbalisation et d'échange s'est prolongé jusqu'à saturation c'est-à-dire jusqu'au moment où il ne permettait plus de mettre au jour de nouveaux critères ni de nouveaux indicateurs. Ceux-ci ont ensuite été organisés sous forme d'une grille d'évaluation adaptée à chaque situation d'intégration. Sur base de cette grille, l'enseignant-chercheur et l'enseignant qui avaient mis en œuvre la même situation d'intégration ont alors passé au crible les productions des élèves et les ont évaluées.

Tableau 1 : Exemple de grille d'évaluation relative à une situation d'évaluation de la compétence 1 sur l'organisation sociale au Moyen Âge (D'après Bouhon, M. et Dambroise, C. [2002], p. 63).

Critères	Éléments à observer	(1)			(2)		
		Q1	Q2	Q3	Q1	Q2	Q3
Pertinence des questions	<i>Cochez les cases qui ne satisfont pas aux items ci-dessous. Si une des cases est cochée dans une colonne la question vaut 0 point.</i>						
	1 La question est en lien avec les trois dimensions du thème de recherche : · espace (Occident) · époque (Moyen Âge) · concept (stratification sociale)						
	2 La question s'appuie sur une interprétation correcte des documents.						
	3 La question est une vraie question de recherche. La réponse à la question ne se trouve ni dans les documents ni dans le cours.						
	4 La question est différente des autres.	/2	/2	/2	/2	/2	/2
		/6			/6		
Précision des questions	Sont explicitement mentionnés dans la formulation de la question : · le ou les document(s) sur le(s)quel(s) la question s'appuie ; · les dimensions du thème de la question : la chronologie ou le lieu ou les classes sociales ...						
	<i>Retirez 0,5 point par case cochée</i>				/2		/2
Correction du langage	La question est formulée dans un français correct. Elle a du sens pour tout lecteur.						
	<i>Retirez 1 point par case cochée</i>				/2		/2
Exploitation des documents	<i>Cochez d'abord ci-dessous les documents qui ont été de toute évidence exploités pour formuler une question (2 points par ☐ cochée).</i>			/4			/4
Confrontation des documents	<i>Reliez ensuite les documents qui ont été confrontés de façon pertinente pour formuler une question (1 point par confrontation pertinente).</i>			/2			/2
	(1) Doc 1 : ☐ Doc 2 : ☐ Doc 3 : ☐						
	(2) Doc 1 : ☐ Doc 2 : ☐ Doc 3 : ☐						
Exploitation des connaissances	Les connaissances suivantes sont exploitées correctement : <i>cochez les cases en fonction du type de connaissance et de la question.</i>						
	· Division de la société en groupes selon des critères. · Inégalité de conditions entre ces différents groupes. · Autres connaissances mobilisées correctement :						
	<i>Comptez 2 points par connaissance exploitée une première fois et 0,5 point par occasion supplémentaire si celle-ci apporte quelque chose en plus par rapport à la première fois.</i>			/4			/4

(1) Évaluation de l'élève (2) Évaluation du professeur

2.4 L'évaluation formative des élèves

Dans chaque classe, les enseignants ont conduit ensuite une séance d'évaluation formative. Celle-ci était organisée en quatre temps. Les élèves étaient d'abord amenés à découvrir la grille qui avait servi à évaluer leur production. Ensuite, le professeur leur soumettait quelques productions, venant de l'autre classe qui avait été confrontée à la même situation d'intégration. Ces productions avaient été choisies pour les qualités et les défauts qu'elles présentent, de telle sorte que les élèves puissent, par l'analyse, mettre en évidence les principaux critères et indicateurs présentés dans la grille d'évaluation et donc s'approprier celle-ci. Chacun était ensuite invité à auto-évaluer sa production, à l'aide de la grille. La séance se terminait par l'administration d'un questionnaire qui interrogeait les élèves sur la lisibilité des grilles, leur pertinence et leur caractère formatif.

2.5 L'analyse des réponses des élèves, de leur auto-évaluation et des grilles d'évaluation

La démarche se terminait par l'analyse, par l'enseignant-chercheur et son collègue professeur d'histoire, des réponses des élèves au questionnaire, de leurs réactions durant la phase d'auto-évaluation ainsi que de l'écart entre la note qu'ils s'étaient chacun attribuée et celle octroyée par l'enseignant. La grille d'évaluation a fait également l'objet d'un examen attentif. Les critères et les indicateurs retenus sont-ils pertinents ? La pondération semble-t-elle judicieuse ? La grille est-elle aisément maniable par l'enseignant ? Et l'élève ? ...

L'ensemble de cette démarche a ensuite été réitérée à deux reprises, les élèves étant donc soumis à deux nouvelles séquences d'apprentissage, chacune comprenant une situation d'intégration portant sur la même compétence. À chaque reprise, la grille d'évaluation était ajustée de telle sorte que, à l'issue du dispositif de recherche, les chercheurs ont été à même de proposer, pour chacune des deux compétences qui avaient été choisies, une grille d'évaluation-type qui propose aux enseignants un choix de critères et des types d'indicateurs (cfr *infra* 3.2).

3. La construction des situations d'intégration et des grilles d'évaluation : une question de validité et de fiabilité

Parmi les nombreux problèmes méthodologiques qu'a posés cette recherche, il en est deux que rencontrent aussi les enseignants, dès lors qu'ils entreprennent d'évaluer les compétences des apprenants.

- Le premier problème concerne l'élaboration des situations d'intégration elle-même : comment faire en sorte que ces situations donnent lieu à une production dont l'évaluation permettra d'inférer la maîtrise par l'élève d'une compétence, celle dont l'apprentissage est visé ? Poser cette question revient à s'interroger sur la manière d'assurer au mieux la *validité* de l'évaluation.
- Le second problème concerne la fabrication de la grille d'évaluation et la manière de s'y prendre pour réduire, autant que faire se peut, l'écart entre les appréciations que pourraient porter des correcteurs différents, au premier chef l'enseignant et l'apprenant lui-même. Poser cette question revient à s'interroger sur la manière d'assurer au mieux la *fiabilité* de l'évaluation.

3.1 La conception des situations d'intégration

3.1.1 Une question de validité... mais aussi de cohérence

Pour De Ketele (1989), évaluer consiste à :

- recueillir un ensemble d'informations ;
- examiner le degré d'adéquation entre cet ensemble d'informations et un ensemble de critères adéquats ;
- et ce en vue de prendre une décision : celle-ci peut être de type diagnostique, formatif ou certificatif.

Dès le recueil d'informations, une question de validité se pose : comment faire en sorte que les informations, de par leur nature et les conditions de leur recueil, renseignent bien sur la maîtrise de la compétence à propos de laquelle le professeur ou le chercheur tentent de se documenter ? Ce problème de validité prend une dimension plus large quand il s'agit d'outiller un grand nombre d'enseignants, par exemple l'ensemble des professeurs d'histoire en Communauté française de Belgique : quelles orientations leur fournir pour assurer la validité du recueil auquel ils procèderont dans leurs différentes classes, au fil des semaines et des années ? Quel outillage leur fournir pour que les situations d'intégration qu'ils fabriqueront évaluent bien les compétences énoncées dans les programmes et qu'ils prétendent évaluer ?

Appliquée à l'échelle d'un système éducatif, cette question de validité recèle donc aussi un enjeu en terme de *cohérence* : sur un plan « macro » ou « méso », comment faire en sorte qu'un enseignant, qu'elle que soit son école ou sa classe, mette en place des situations d'évaluation suffisamment isomorphes par rapport à celles qu'un autre enseignant mettra en place dans une autre école, une autre classe ? Et sur un plan « micro », comment faire en sorte que, d'une séquence d'apprentissage à l'autre, les situations d'évaluation construites par un même enseignant à propos d'une même compétence soient suffisamment isomorphes les unes par rapport aux autres ?

La première question intéressera d'abord les responsables des systèmes d'enseignement (niveau « macro ») mais aussi les chefs d'établissements (niveau « méso »), soucieux l'un et l'autre que les objectifs de formation poursuivis par les enseignants d'une discipline donnée, dans l'ensemble du système éducatif ou dans leur école, soient proches sinon identiques. La seconde question intéressera davantage le pédagogue, le didacticien et l'enseignant (niveau « micro ») soucieux que l'« approche par compétences » ne fasse pas le jeu des élèves dont les pré-acquis mais aussi les déterminants socioculturels les rendent plus aptes à gérer la complexité. Il convient en effet de rappeler que faire apprendre des compétences constitue une entreprise ambitieuse en ce qu'elle est nourrie par le projet de rendre les élèves capables de transférer leurs connaissances et donc d'accéder à la maîtrise de tâches complexes d'un haut niveau taxonomique. À moins de se résoudre à faire le jeu des déterminants socioculturels, pareille ambition suppose le respect d'un certain nombre de conditions au nombre desquelles il faut citer la garantie ferme qui est faite à l'apprenant que les conditions d'évaluation d'une même compétence seront identiques, d'une situation d'intégration à l'autre (Jadoulle, J.-L. [2011b]).

Il convient de souligner combien ces questions relatives à la cohérence des situations d'intégration, au niveau du système éducatif et de l'établissement d'enseignement d'une part, au niveau de la classe d'autre part, sont loin d'être anecdotiques. Elles renvoient en effet à deux objectifs fondamentaux de la réforme des programmes qui a été mise en œuvre en Communauté française de Belgique (Jadoulle, J.-L. [2011a]). Le premier objectif concerne la volonté de clarifier les tâches attendues des élèves en fin de formation et s'origine dans le constat d'un déficit de précision tel que les enseignants étaient devenus, en somme, les détenteurs de ce pouvoir exorbitant de définir ce qui est attendu des élèves et qu'il était, dès lors, quasi impossible d'évaluer l'efficacité du système éducatif. Le second objectif concerne le projet de réduire les inégalités scolaires, tellement accusées en Communauté française de Belgique.

Pour permettre aux enseignants d'assurer la validité et la cohérence des situations d'intégration élaborées en vue de faire apprendre et d'évaluer une même compétence, la recherche qui a été conduite a pris appui sur les *familles de situations* qui avaient été énoncées, pour chaque compétence, dans le programme de la Fédération de l'Enseignement secondaire catholique.

3.1.2 Le choix d'une optique situationnelle : un gage de validité

Suite au « Décret-Missions » de 1997, la Ministre de l'Éducation avait désigné, dans chaque discipline scolaire, une commission composée d'experts et d'enseignants de tous les « réseaux » d'enseignement. En histoire, la commission a proposé à la Ministre une liste de quatre « compétences terminales » que le Parlement de la Communauté française a votée, le 24 avril 1999.

Compétence terminale n°1 Au départ d'une situation du passé ou du présent, élaborer une problématique de recherche et sélectionner dans divers lieux d'information et de documentation des renseignements utiles.	Compétence terminale n°3 Sur base d'un nombre limité de données, organiser une synthèse mettant en évidence, selon les cas, des permanences, des processus évolutifs, des changements ou des synchronismes et formuler des hypothèses explicatives.
Compétence terminale n°2 En fonction d'une question déterminée, remettre dans son contexte historique, analyser et critiquer un ensemble limité de sources.	Compétence terminale n° 4 Concevoir, préparer et mener à bien une stratégie de communication d'un savoir historique en ayant recours à différents modes d'expression, écrit, oral, visuel ou audiovisuel.

Tableau 2 : Les « compétences terminales » en histoire votées par le Parlement de la Communauté française, le 24 avril 1999 (D'après Ministère de la Communauté française, *Compétences terminales et savoirs requis en histoire*, Bruxelles, 1999, p. 11.)

La constitution belge garantissant à tous les « réseaux » la « liberté de méthode », l'appropriation, par les deux principaux « réseaux » d'enseignement⁶, des énoncés de « compétences terminales », s'est faite en des termes sensiblement différents.

⁶ Les réseaux « officiel » et « libre confessionnel ».

Dans l'enseignement « libre confessionnel » (Fesec), au sein du « groupe à tâche » chargé, dès 1999-2000, de rédiger le nouveau programme d'histoire, l'absence de précisions quant aux tâches attendues des élèves et aux situations dans lesquelles elles devront être mises en œuvre, a paru source de problèmes. Ceux-ci ont été évoqués plus haut et concernent à la fois le pilotage du système éducatif (niveau « macro »), la cohérence des objectifs de formation poursuivis au sein d'un même établissement (niveau « méso ») et l'isomorphisme des situations d'intégration proposées par un même enseignant en vue de l'apprentissage et de l'évaluation d'une même compétence (niveau « micro »).

Ce constat a conduit le groupe à opter pour une approche dite « opérationnelle » ou « situationnelle »⁷. Inspirée des travaux de De Ketele et Roegiers (2000) puis de Beckers (2002) et Jonnaert (2003), cette approche suppose la définition, pour chaque compétence, d'un type ou d'une famille de situations, nouvelles et complexes, dont les contours sont délimités par un certain nombre de paramètres. Ils sont censés assurer que les différentes situations d'apprentissage et d'évaluation d'une même compétence, situations proposées par un même enseignant ou par des enseignants différents, appartiennent bien à la même famille.

Le « groupe à tâche » chargé de l'écriture du nouveau programme d'histoire de la Fesec a donc assorti chaque énoncé de compétence d'une famille de situations et à ainsi précisé et interprété les énoncés votés par le Parlement. Le tableau 3 en précise les paramètres principaux.

Compétence 1 — Énoncer un ensemble organisé de questions de recherche pertinentes et originales ➤ En fonction d'un objet de recherche ➤ Sur base de l'analyse d'un ensemble de documents : — d'un genre déjà rencontré — suffisamment fiables par rapport à l'objet de recherche — comportant des paradoxes, des contradictions, des divergences entre eux et/ou avec des informations apprises ➤ En mobilisant un concept qui soit : — un réel outil de compréhension de l'objet de recherche, — un réel outil d'investigation de la documentation	Compétence 3 — Répondre à une question de recherche sous forme d'un texte de synthèse ➤ En fonction d'une question de recherche ➤ Sur base d'un ensemble de documents : — d'un genre déjà rencontré — suffisamment fiables par rapport à l'objet de recherche — comprenant des informations pouvant être mises en relation entre elles et/ou avec des connaissances apprises ➤ En mobilisant un concept qui soit : — un réel outil de compréhension de l'objet de recherche, — un réel outil d'investigation de la documentation
Compétence 2 — Identifier les documents pertinents et, pour une source au moins, énoncer les raisons de s'y fier ou de s'en méfier ➤ En fonction d'un objet ou d'une question de recherche ➤ Sur base d'un ensemble de documents — d'un genre déjà rencontré — pertinents et non pertinents par rapport à l'objet de recherche — posant des problèmes de critique ➤ Sur base d'informations relatives au contexte de production des documents	Compétence 4 — Communiquer des savoirs préalablement appris sous forme d'une ligne du temps, d'un schéma, d'un tableau, d'un plan, d'un « poster » ➤ Des savoirs préalablement appris ➤ Des matériaux nouveaux à traiter (documents d'un genre familier) pour fabriquer le vecteur de communication auquel les élèves ont été initiés

Tableau 3 : Les énoncés des compétences en histoire dans les établissements affiliés à la Fesec ou le « réseau libre confessionnel » (D'après *Programme d'histoire*, Bruxelles, Fesec, 2000)

- La première compétence a donc pour objectif d'apprendre à l'élève à énoncer des questions de recherche sur base de l'analyse d'un ensemble documentaire nouveau et en mobilisant un concept étudié auparavant. La sélection par l'enseignant des documents doit donc être l'objet de la plus grande attention : rassemblés autour d'un objet de recherche nouveau, les informations qu'ils contiennent doivent pouvoir être confrontées entre elles ou avec des connaissances apprises et être éclairées par la mobilisation d'un des concepts préalablement étudiés.
- La deuxième compétence vise à rendre l'élève capable d'identifier, parmi un ensemble de documents et en fonction d'un objet ou d'une question de recherche, la ou les sources pertinente(s) et d'en évaluer la crédibilité. Pour ce faire, il dispose, à propos de chaque document, d'informations fournies par l'enseignant et/ou d'instruments de travail et qui peuvent le renseigner sur les différents témoins, leurs options, le contexte dans lequel ils ont témoigné, leurs intentions et leur(s) destinataire(s)...

⁷ Roegiers, X. et De Ketele, J.-M. (2000), p. 96.

- La troisième compétence a pour finalité d'apprendre à l'élève à rédiger une synthèse en vue de répondre à une question de recherche nouvelle, en analysant un ensemble de documents nouveaux et en mobilisant un concept préalablement appris.
- La quatrième compétence porte sur l'aptitude de l'élève à transposer des connaissances préalablement apprises, et ce sous forme d'un « poster », d'un tableau, d'une ligne du temps ou d'un schéma.

Ce choix distingue, assez nettement, l'histoire des autres disciplines du cursus secondaire en Belgique francophone. En effet, dans la plupart d'entre elles, les compétences qui ont été définies correspondent tantôt à des capacités générales (savoir écrire, savoir observer, savoir analyser...), tantôt à des savoir faire techniques (résoudre une équation, accorder un participe passé...), tantôt encore à des attitudes (faire preuve d'autonomie, planifier son travail...). Ces énoncés se situent donc dans une optique que l'on peut qualifier de « généraliste » ou « longitudinale »⁸. Dans cette optique, la compétence est vue comme une capacité générale dont le développement passe par la rencontre de situations très diverses, au contraire de l'approche « opérationnelle » ou « situationnelle » qui postule que l'apprentissage d'une compétence suppose que cette diversité soit contenue dans le paramétrage de la famille de situations attenante à cette compétence. Dans le « réseau officiel », le nouveau programme d'histoire publié en 2000, tout en déclinant les quatre « compétences terminales » en sept compétences, s'inscrit dans cette optique « généraliste » ou « longitudinale ».

Toutefois, depuis 2000, un processus de rapprochement est en cours à la faveur des travaux de la Commission inter-réseaux des outils d'évaluation en histoire. Chargée de guider les enseignants et de fournir à l'inspection inter-réseaux des modèles pour apprécier le niveau des études, cette commission élabore et diffuse des outils d'évaluation des compétences en histoire valables pour tous les « réseaux » d'enseignement⁹. Ils ont été élaborés sur base de l'énoncé, pour chacune des quatre compétences terminales, d'une famille de situations, de telle sorte que les différents « réseaux » d'enseignement s'acheminent, de facto, vers une interprétation commune des compétences terminales en histoire.

Dans le cadre de la recherche sur l'évaluation des compétences dont il est question dans la présente contribution, ces familles de situations, alors énoncées au sein du programme de l'enseignement « libre confessionnel », se sont révélés de précieux outils pour garantir la validité des outils d'évaluation élaborés par les chercheurs et les enseignants.

3.2 La conception des grilles d'évaluation

Une fois le recueil d'informations réalisé, d'une manière qui le rende aussi valide que possible, il convient d'assurer la fiabilité de l'examen de ces données et de la prise de décision qui s'en suit. Pour outiller en ce sens les professeurs du secondaire, les deux chercheurs et les enseignants participants à la recherche ont dégagé, pour chaque compétence, une série de critères c'est-à-dire de qualités attendues de la production de l'élève (Roegiers, X. [2010 a et b]). Ceux-ci ont été hiérarchisés en critères principaux et critères secondaires, la maîtrise des premiers étant jugée nécessaire et suffisante pour décider de la maîtrise de la compétence. Pour chaque critère, des indicateurs ou des types d'indicateurs ont été identifiés de telle sorte que le professeur dispose d'éléments observables sur base desquels il puisse déterminer si le critère est satisfait ou non (Roegiers, X. [2010 a et b]).

Cet ensemble de critères et d'indicateurs a été présenté sous forme de grilles d'évaluation-typ, lesquelles ont été ensuite complétées par des propositions en termes de seuils de maîtrise pour chaque critère et pour l'appréciation de la performance globale.

⁸ Roegiers, X. et De Ketele, J.-M. (2000), p. 96. Elle est aussi qualifiée de « virtuelle » par Jonnaert, Ph. (2003), p. 39-40.

⁹ Ils sont téléchargeables à la page <http://www.enseignement.be/index.php?page=24420>.

Critères	Éléments à observer
<i>CRITÈRES MINIMAUX</i>	
Pertinence des questions	· Chaque question est-elle en lien avec les trois dimensions du thème de recherche ? · espace · époque · concept · Chaque question s'appuie-t-elle sur une interprétation correcte des documents ? · Chaque question est-elle une vraie question de recherche ? La réponse à la question ne se trouve-t-elle pas dans les documents ou dans le cours ? · Chaque question est-elle différente des autres ?
Exploitation des connaissances	· L'élève a-t-il mobilisé les principales caractéristiques du concept susceptibles d'être exploitées pour formuler ses questions ? · L'élève a-t-il exploité des savoirs pertinents du cours pour formuler ses questions ? · L'élève a-t-il exploité d'autres savoirs pertinents ?
Prise en compte des documents	· Les questions se fondent-elles sur les documents ?
<i>CRITÈRES DE PERFECTIONNEMENT</i>	
Précision des questions	· L'élève a-t-il explicitement mentionné dans la formulation de chaque question : · le ou les document(s) sur le(s)quel(s) la question s'appuie ? · les composantes de la situation historique sur laquelle la question s'appuie ? · la chronologie, les faits, les lieux, les institutions, les personnes
Organisation des questions	· Les diverses parties qui composent la question sont-elles cohérentes ? · Certaines sous-questions ne devraient-elles pas être rapprochées d'autres questions ?
Correction du langage	· La question est-elle formulée dans un français correct ? A-t-elle du sens pour tout lecteur ?
Confrontation des documents	· L'élève confronte-t-il les documents entre eux pour rechercher des informations qui posent question ?

Tableau 4 : Critères et type d'indicateurs proposés pour l'évaluation de la compétence 1 (D'après Bouhon, M. et Dambroise, C. [2002], p. 24).

L'ensemble de cet outillage a été diffusé sous la forme d'une publication (Bouhon, M. et Dambroise, C. [2002]) et dans le cadre de séminaires de formation qui ont été et sont organisés périodiquement. Ces formations ont également incorporés peu à peu des outils conçus, dans le même esprit, et relatifs aux deux compétences sur lesquelles la recherche dont il est question dans la présente contribution ne s'était pas arrêtée. Ces autres outils ont été diffusés dans d'autres publications (Jadoulle, J.-L. et Bouhon, M. [2001, nouv. éd. : 2003] ; Jadoulle, J.-L., Bouhon, M. et Nys, A. [2004]). L'ensemble de ces éléments a également nourri les travaux de la Commission inter-réseaux des outils d'évaluation en histoire dont les situations d'évaluation des compétences ont, à leur tour, alimenté les formations et complété l'équipement des enseignants, tous « réseaux » confondus.

Notre propre expérience et celle des nombreux enseignants mais aussi des étudiants en formation initiale que nous côtoyons nous indiquent que ces grilles d'évaluation constituent un moyen efficace qui permet d' « éviter le pire en matière de subjectivité... » (Beckers, J. [2002]). Il convient en effet de rappeler que si le recours à une certaine technicité permet effectivement de réduire les écarts entre correcteurs, il ne permet évidemment pas de les neutraliser tous, à coup sûr. L'évaluation restant un processus inférentiel, le jugement de l'enseignant, fut ce-

t-il outillé et encadré, demeure une composante de l'acte d'évaluer (Roegiers, X. [2010b]). Pour réduire les écarts, il convient, nous semble-t-il, d'assurer une réelle appropriation de la grille d'évaluation par les enseignants et, dans le cas où l'on pratique l'auto- et la co-évaluation, par les élèves.

4. Les outils d'évaluation des compétences en histoire en Communauté française de Belgique : quel impact ?

En l'absence de données empiriques, l'impact de ces outils sur les pratiques d'évaluation des enseignants demeure inconnue. Largement diffusés, ils ont été repris par de nombreux enseignants. Mais sous quelle forme et avec quelles adaptations ? Et pour quels résultats ? Ces questions demeurent sans réponse.

La seule donnée empirique dont nous disposons provient d'une enquête que nous avons conduite, à la faveur de la recherche doctorale de Bouhon (2009). Ainsi, entre avril et juillet 2008, le questionnaire qui avait été soumis à l'ensemble des enseignants d'histoire, en avril 2001, a été réadministré auprès de 130 professeurs d'histoire de l'enseignement « libre confessionnel ». Même s'ils n'ont pas été choisis sur base d'une procédure d'échantillonnage, ces enseignants peuvent être considérés comme représentatifs des professeurs d'histoire oeuvrant dans les établissements de ce réseau.

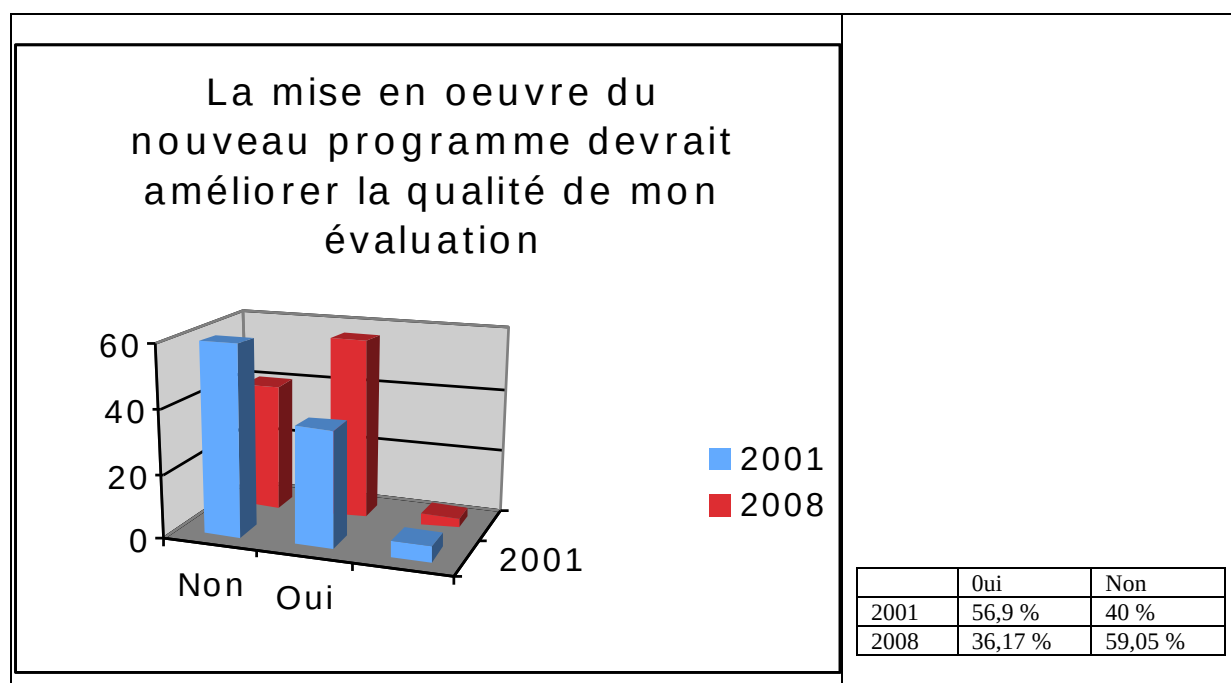


Figure 3 : Evolution entre 2001 et 2008 du gain escompté par les enseignants au plan de l'évaluation des performances des élèves

(D'après Jadoulle, J.-L. et Letor, C. [2002] et Bouhon, M. [2009]).

En avril 2001, à la question de savoir si le nouveau programme allait modifier la qualité de leur évaluation, 59,05 % des répondants étaient soit convaincus que leur évaluation ne devrait pas s'améliorer soit très réservés à ce sujet. Seule une minorité de répondants (36,17 %) escomptait un bénéfice. En 2008, par contre, ils sont 56 % à adhérer à l'affirmation selon laquelle la mise en œuvre du nouveau programme a amélioré la qualité de leur évaluation.

Ce résultat doit évidemment être pris avec les précautions d'usage. Il témoigne, au mieux, d'un sentiment de plus grande efficacité des enseignants au plan de l'évaluation des compétences de leurs élèves. L'imputer au nouveau programme et aux outils d'évaluation à la réalisation desquels nous avons contribué serait hasardeux.

5. Références et bibliographie

Beckers, J. (2002). *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*. Bruxelles : Labor.

Bouhon, M. (2009). *Les représentations sociales des enseignants d'histoire relatives à leur discipline et son enseignement*. Louvain-la-Neuve : UCL.

Bouhon, M. et Dambroise, C. (2002). *Evaluer des compétences en classe d'histoire*. Louvain-la-Neuve : Unité de Didactique et de Communication en Histoire de l'UCL.

De Ketele, J.-M. (1989). L'évaluation de la productivité des institutions d'éducation. In *Université et société : le rendement de l'enseignement universitaire*. Bruxelles : Fondation Universitaire.

Jadoulle, J.-L. (2000). Développer des compétences « historiennes » en classe d'histoire. In *Actes du I^{er} Congrès des chercheurs en éducation. Le point sur la recherche en éducation en Communauté française* (p. 257-259). Bruxelles : Ministère de l'Enseignement supérieur et de la recherche scientifique.

Jadoulle, J.-L. (2003). Evaluer des compétences en histoire : oui mais comment ? In *Les outils d'évaluation des compétences disciplinaires dans l'enseignement secondaire. Actes de l'Université d'été, Université de Liège, 23 août 2002* (p. 26-29). Liège : ULg.

Jadoulle, J.-L. (2009). Faire apprendre et évaluer des compétences en classe d'histoire. *Cahiers Pédagogiques*, 471, 21-23.

Jadoulle, J.-L. (2011a). Faire apprendre des compétences en classe d'histoire : tenants et aboutissants d'une réforme au long cours en Communauté française de Belgique. In M.-A. Ethier, D. Lefrançois et J.-F. Cardin (Ed.), *Enseigner et apprendre l'histoire : manuels, enseignants et élèves* (p. 331-357). Québec : Presses universitaires de Laval.

Jadoulle, J.-L. (2011b). Approche par compétences et équité des apprentissages : un mariage impossible ? Etude de cas à propos de l'enseignement de l'histoire. *Puzzle. Revue du Centre interfacultaire de formation des enseignants de l'ULg*, 30, 33-37.

Jadoulle, J.-L. et Bouhon, M. (2000 ; nouv. éd. : 2003). *Développer des compétences en classe d'histoire*. Louvain-la-Neuve : Unité de Didactique et de Communication en Histoire de l'UCL.

Jadoulle, J.-L., Bouhon, M. et Dambroise, C. (2000). Comment évaluer des compétences en classe d'histoire ? In L. Paquay (Ed.), *L'évaluation des compétences chez l'apprenant. Pratiques, méthodes et fondements* (p. 117-124). Louvain-la-Neuve : UCL-GRIFED.

Jadoulle, J.-L., Bouhon, M. et Dambroise, C. (2002). Evaluer des compétences en classe d'histoire. *Le point sur la recherche en éducation*, 23, 3-29.

Jadoulle, J.-L. et Bouhon, M. et Nys, A. (2004). *Conceptualiser le passé pour comprendre le présent. Conceptualisation et pédagogie de l'intégration en classe d'histoire*. Louvain-la-Neuve : Unité de Didactique et de Communication en Histoire de l'UCL.

Jadoulle, J. -L. et Letor, C. (2002). Que pensent les enseignants des nouveaux programmes ? *Exposant neuf*, 10, 30-31.

Roegiers, X. (2010a). *La pédagogie de l'intégration. Des systèmes d'éducation et de formation au cœur des sociétés*. Paris-Bruxelles : De Boeck.

Roegiers, X. (2010b). *L'école et l'évaluation. Des situations complexes pour évaluer les acquis des élèves*. Paris-Bruxelles : De Boeck.

Roegiers, X. et De Ketele, J.-M. (2000). *Une pédagogie de l'intégration. Compétences et intégration des acquis dans l'enseignement*. Paris-Bruxelles : De Boeck.

EVALUER L'HISTOIRE DES ARTS A L'ECOLE PRIMAIRE, DISCOURS ET DISPOSITIFS

Gaëlle Boubennec*, Laurent Séjourné**, Cécile Vendramini***

* IUFM de Bretagne - Université de Bretagne Occidentale, gaelle.boubennec@bretagne.iufm.fr

** IUFM de Bretagne- UBO, laurent.sejourne@bretagne.iufm.fr

*** IUFM de Bretagne- UBO, cecile.vendramini@bretagne.iufm.fr

Mots-clés : histoire des arts, éducation artistique, école primaire, livret de compétences, transdisciplinarité

Résumé. L'histoire des arts est un nouvel enseignement transdisciplinaire qui est apparu dans les programmes de l'école primaire française en 2008. Dès son apparition, son évaluation a été prévue dans le livret personnel de compétences de l'élève, introduit la même année dans le cursus du premier degré. La teneur des items proposés pour cette évaluation pose quelques questions d'interprétation. L'équipe de recherche a élaboré un scénario sous forme audio-numérique dont la pertinence en matière d'évaluation transdisciplinaire a été validée par des formateurs d'IUFM en arts (musique, arts visuels, cinéma), en histoire et en lettres. L'étude rend compte de la réception de ce dispositif par des acteurs du terrain et de leur discours quant à sa mise en œuvre possible en classe (articulation entre les référentiels, les compétences des maîtres et l'évaluation par compétences en éducation artistique).

Introduction

Cette recherche en cours fait suite à une étude récente sur l'introduction de l'histoire des arts dans le premier degré de l'école¹. Apparu au lycée dans les années 1990, l'enseignement de l'histoire des arts a été intégré aux programmes de l'école primaire en 2008 et du collège en 2009². Une première réflexion a été effectuée sur les interrogations³ et polémiques soulevées à propos de la réception de cette injonction ministérielle, en formation initiale des maîtres (Vendramini, Gourlaouen, 2011). Cette communication fait état des travaux de recherche menés actuellement par ce même groupe à propos des dispositifs de mise en œuvre et de l'évaluation de ce nouveau domaine d'enseignement dans le premier degré de l'école. Dès son apparition, l'évaluation de l'histoire des arts a été prévue, comme le soulignent les divers discours d'accompagnement mis en place : « Tout projet nouveau nécessite qu'on envisage le processus d'évaluation (...) Comment aider un nouvel enseignement à être considéré comme tel, dans un système qui reste très marqué

¹ L'équipe de recherche (2010-2013) comprend trois enseignants-chercheurs (Musique, Cinéma, Histoire et géographie), six formateurs disciplinaires (deux en musique, deux en arts visuels, une formatrice lettres et un formateur Tice), et une professeure des écoles.)

² L'organisation de l'enseignement de l'histoire des arts à l'école primaire a été publiée dans le B.O. n°32 du 28 août 2008.

³ La première interrogation portait sur le pluriel employé (histoire des arts et non de l'art). Le concept d'« histoire des arts », a été défini par le recteur Philippe Joutard en 1989, mais Marie Lavin souligne également l'influence de Gérard Monnier dans cette décision de mettre les arts au pluriel : « Gérard Monnier repère des voies nouvelles : succès récent de l'art des jardins, espaces publics investis par de nouvelles pratiques artistiques dans des relations temporaires: danse, mode, musique, cinéma, arts de la rue, nouvel art mural conçu comme intervention de proximité en rapport avec une problématique locale (...) La prolifération de tous ces arts justifiant le pluriel. » (Lavin, 1998, 77-78).

par la notation, la correction, le classement, si on ne lui permet pas d'accéder au rang de discipline officielle par l'octroi d'une épreuve, quelle qu'elle soit ? »⁴ Ainsi, parallèlement à l'introduction de ce nouvel enseignement, une épreuve en histoire des arts au brevet des collèges faisait son apparition dans le second degré et l'évaluation de ce domaine prenait toute sa place à l'école primaire dans les items du livret de compétences du socle commun, à partir du cycle 2. Il est ainsi précisé que si l'évaluation « au sens traditionnel » se pratique uniquement au collège, lors du diplôme national du brevet (DNB), « on saura apprécier le degré d'atteinte d'un certain nombre d'acquis à chaque niveau, car ce qu'on vise relève davantage du savoir-être que du savoir-faire. Et on gardera la trace des parcours culturels effectués »⁵. Pour avancer dans la réflexion sur les compétences à acquérir en histoire des arts à l'école primaire, l'équipe de recherche a élaboré un scénario sous forme audio-numérique dont la pertinence en matière d'évaluation transdisciplinaire a été validée par des formateurs d'IUFM en arts (musique, arts visuels, cinéma), en histoire et en lettres. Cet article rend compte de la réception de ce dispositif par des acteurs du terrain et de leur discours quant à sa mise en œuvre possible en classe (articulation entre les référentiels, les compétences des maîtres et l'évaluation par compétences en éducation artistique).

1. Du discours des enseignants sur les dispositifs choisis en histoire des arts

Lors des premières réflexions sur l'introduction de l'histoire des arts à l'école primaire, la question des compétences des enseignants dans ce nouveau domaine « sans spécialiste à la clé »⁶, avait été posée. Si depuis 2004 une certification en histoire des arts est nécessaire pour l'encadrement de l'histoire des arts au lycée⁷, elle n'est pas demandée au niveau du collège, ni imposée aux formateurs d'institut de formation des maîtres (les formateurs concernés sont en général des enseignants en histoire, lettres ou éducation artistique). Dans un article récent sur l'intégration de l'histoire des arts dans la scolarité, le sociologue Jean-Michel Pire n'avait d'ailleurs pas hésité à mettre en doute les compétences des préposés à cet enseignement (Pire, 2010, 341). Nos nouvelles investigations se sont portées sur le discours en la matière que tiennent les enseignants du terrain de l'école primaire. Dans un premier temps, des questionnaires ont été passés à des enseignants de cycle 3 ayant suivi un stage de formation continue en histoire des arts. Puis deux enseignantes de CM1 ont été sollicitées pour réagir à un dispositif possible d'enseignement de l'histoire des arts.

1.1 Quelques dispositifs d'enseignement de l'histoire des arts

Pour avoir un indicateur sur l'auto-évaluation des compétences en histoire des arts du côté des acteurs du terrain, un questionnaire a été passé à une soixantaine d'enseignants du premier degré, dans le cadre d'une formation continue dans ce domaine⁸. Les questions portaient sur l'organisation choisie pour mettre en œuvre ce nouvel enseignement et sur l'appréciation des problèmes soulevés. L'analyse des réponses tient compte du fait que ces enseignants avaient choisi cette formation continue, et les propos tenus ne reflètent que l'état d'esprit de ce groupe de maîtres du premier degré. Leurs réponses montrent que l'introduction de l'histoire des arts à l'école

⁴ E.S.E.N (2010). L'ESEN, Ecole supérieure de l'éducation nationale, de l'enseignement supérieur et de la recherche est un service à compétence nationale, fondé par arrêté du 24 août 2011 et rattaché à la direction générale des ressources humaines des ministères de l'éducation nationale et de l'enseignement supérieur et de la recherche.

⁵ Ibid note 4.

⁶ Il n'y a pas en France de CAPES d'histoire des arts.

⁷ B.O. n° 39 du 28 octobre 2004.

⁸ Dès la rentrée 2008, il est demandé aux corps d'inspection de « veiller à la mise en place effective de ce nouvel enseignement, dans toutes les écoles primaires. » (Préparation de la rentrée 2008 – BO n° 15 du 10 avril 2008.) Cette formation continue avait lieu dans le département du Finistère.

primaire a été dans l'ensemble plutôt bienvenue, ce nouveau domaine d'enseignement étant considéré comme « très riche et intéressant ». A travers lui, les enseignants visent plutôt à faire naître « la curiosité, l'envie de connaître, le sensible plutôt que l'acquisition de connaissances ». Ils semblent se sentir « *moyennement compétents* » en la matière, tout en se disant très démunis au niveau des ressources et du temps à trouver pour une pratique cohérente. Rappelons que le cadrage de cet enseignement repose sur trois incontournables : 1) les cinq périodes historiques suivantes : Préhistoire et Antiquité gallo-romaine, Moyen-Âge, Temps modernes, XIX^{ème} siècle, XX^{ème} siècle et notre époque ; 2) les six grands domaines artistiques que sont « Arts de l'espace », « Arts du langage », « Arts du quotidien », « Arts du son », « Arts du spectacle vivant » et les « Arts visuels » ; 3) la liste de références d'œuvres artistiques destinées à aider les enseignants dans leur choix des supports qui seront étudiées en classe. La seule indication donnée aux enseignants est un volume horaire annuel de vingt heures qui concerne l'ensemble des domaines disciplinaires. Chaque enseignant est libre d'organiser l'histoire des arts comme il le souhaite. Pour mieux cerner cette organisation, un entretien a été mené avec la conseillère pédagogique en éducation musicale qui encadrerait la formation, et qui a pu décrire quatre organisations adoptées par des enseignants du premier degré dans leur classe :

1. « Les 20 heures annuelles sont intégrées dans l'ensemble des disciplines en fonction de l'entrée choisie, de l'œuvre abordée, d'un événement, d'un projet. Il n'y a pas de plage horaire spécifique « histoire des arts » dans l'emploi du temps.
2. Un moment de « découverte artistique » est instauré et organisé chaque semaine. Il figure à l'emploi du temps : 30 à 35 minutes autour d'une œuvre artistique.
3. Un temps d'histoire des arts est programmé régulièrement (ex : une demi-journée par mois). Il figure à l'emploi du temps. Une liste d'œuvres et de sujets d'étude est établie pour l'année.
4. Des projets liant pratiques artistiques et histoire des arts sont proposés aux élèves.
5. Une demi-journée par semaine est consacrée à ces projets artistiques et culturels. Le temps consacré à ces projets dépasse le volume horaire strictement prévu pour les pratiques artistiques et l'histoire des arts car ces projets permettent un lien avec d'autres disciplines (maîtrise de la langue et littérature, histoire, TICE...) »

Elle précise que la première organisation est la plus courante. Les séances d'histoire des arts sont alors essentiellement liées au programme d'histoire et concernent plus particulièrement les arts de l'espace, arts du quotidien et arts visuels. Cette conseillère en arts note que la seconde organisation pose le problème de la mise en relation des œuvres avec d'autres domaines artistiques. Cela aboutit souvent à un catalogue d'œuvres décontextualisées, n'ayant aucune relation les unes avec les autres. Les deux dernières organisations, pourtant très peu pratiquées, permettent un travail d'appropriation et de réflexion plus approfondi autour des œuvres et donnent davantage lieu à une mise en réseau d'œuvres de différents domaines artistiques. Si les organisations des dispositifs en histoire des arts des maîtres du premier degré peuvent être listés, la question de leur évaluation est très peu évoquée. En général, les enseignants semblent assez réticents à évaluer ce domaine, comme l'explique l'une d'entre eux : « Pour le livret de compétences, nous sommes nous-même en pleine réflexion dessus car l'année dernière nous n'avions pas à le remplir en entier mais seulement sa première page qui est la feuille récapitulative des domaines de compétences à valider. Nous ne savons pas encore clairement si nous devons le remplir en entier cette année, nous attendons des éclaircissements de la part de l'inspectrice.»

2. Du discours officiel sur l'évaluation en histoire des arts

2.1. Les compétences affichées du livret personnel de compétences

Cette enseignante fait allusion à un nouveau livret de compétences qui a fait son apparition parallèlement à l'introduction de l'histoire des arts à l'école primaire. En 2010, ce livret personnel

des compétences du socle commun de 25 pages⁹ a été conçu comme le fil rouge de la scolarité de l'élève, du CE2 à la fin du collège. Il s'agit de la mise en forme d'une évaluation des compétences du socle commun, institué, rappelons-le, en 2005, pour répondre à une harmonisation européenne des compétences scolaires. Si l'on se réfère aux items spécifiques à l'histoire des arts qui apparaissent dans ce livret au cycle 3, ils se résument à deux compétences d'ordre historique et sociologique/ « situer une œuvre dans son époque » et « comprendre une œuvre à travers son époque » ; et à deux compétences à dominante on ne peut plus transversales : « rencontrer des œuvres » et « développer la curiosité pour de nouvelles œuvres ». Les deux premières compétences sous-entendent un éventail implicite de connaissances qu'un enseignant de l'école primaire ne pourra pas toujours déployer. Quant aux deux autres compétences, elles se réfugient dans le lot habituel de compétences transversales que l'on retrouve en général en éducation artistique.

Si l'on résume les attentes officielles, l'enseignement de l'histoire des arts aurait pour objectif « d'offrir à tous les élèves, de tous âges, des situations de rencontres, sensibles et réfléchies, avec des œuvres relevant de différents domaines artistiques, de différentes époques et civilisations ; de les amener à se construire une culture personnelle à valeur universelle fondée sur des œuvres de référence ; de leur permettre d'accéder progressivement au rang « d'amateurs éclairés », maniant de façon pertinente un premier vocabulaire sensible, maîtrisant des repères dans le temps et dans l'espace et appréciant le plaisir que procure la rencontre avec l'art ; de les aider à franchir spontanément les portes d'un musée, d'une galerie, d'une salle de concert, d'un cinéma d'art et d'essai, d'un théâtre, d'un opéra, et tout autre lieu de conversation, de création et de diffusion du patrimoine artistique ; de donner des éléments d'information sur les métiers liés aux domaines des arts et de la culture. A l'école primaire, l'enseignement de l'histoire des arts se veut progressif, cohérent et toujours connecté aux disciplines. Les objectifs, à ce niveau, seraient de susciter la curiosité, de développer le désir d'apprendre, stimuler la créativité, notamment en lien avec une pratique sensible. Il s'agit aussi de développer l'aptitude à voir et regarder, entendre et écouter, observer, décrire, comprendre »¹⁰.

2.2 Les acquis attendus

Plus précisément, à la fin du cycle 3, à travers l'étude d'œuvres relevant de la liste de référence, appartenant aux six domaines artistiques et à chacune des périodes historiques, l'élève aura acquis des connaissances, des capacités et des attitudes. On attendra ainsi de lui qu'il connaisse :

« - des formes d'expression, des matériaux, des techniques et des outils, un premier vocabulaire spécifique ;

- des œuvres d'art appartenant aux différents domaines artistiques ;
- des grands repères historiques ».

Par ailleurs, il devra être capable :

« - de mobiliser ses connaissances pour parler de façon sensible d'œuvres d'art ;

- d'utiliser des critères simples pour aborder ces œuvres, avec l'aide de l'enseignant ;
- d'identifier une œuvre étudiée par son titre, le nom de l'auteur, l'époque à laquelle cette œuvre a été créée ».

Enfin, on attend de l'élève « qu'il fasse preuve de curiosité et de créativité artistiques, qu'il développe une aptitude au dialogue et à l'échange, qu'il découvre et apprécie la diversité des domaines artistiques, des cultures, la pluralité des goûts dans le respect de l'autre et la tolérance¹¹ ».

⁹ Bulletin officiel n°27 du 8 juillet 2010.

¹⁰ Organisation de l'enseignement de l'histoire des arts – BO n° 32 du 28 août 2008.

¹¹ Ibid note 9.

3. Construction d'un outil pour une évaluation transdisciplinaire en histoire des arts

3.1 Présentation de l'outil

Le *Pouvoir du Pingouin Blanc* est un jeu sérieux (élaboré par quatre membres de l'équipe de recherche¹²) qui se présente comme une situation-problème. Pour sa résolution, cette situation nécessite la mobilisation d'un nombre important de procédures. Ce jeu sérieux propose des tâches à caractère interdisciplinaire dans les domaines de la littérature, de la géographie, des arts visuels, de l'éducation musicale et de l'histoire des arts. C'est la mise en perspective des savoirs, des savoirs-faire et des attitudes dans chacune de ces disciplines qui permet *in fine* de résoudre l'énigme¹³. Ce jeu sérieux est destiné aux élèves du Cycle 3, du collège et de l'ASH et nécessite pour sa réalisation une mise en œuvre de trois à six semaines selon le niveau de classe et le temps que l'enseignant décide d'y consacrer.

3.1.1 Situation

Le jeu sérieux débute par la lecture aux élèves d'une lettre prétendument reçue par le professeur. Il s'agit d'un vieux garçon, Arthur Bardonchet qui vient demander de l'aide, car on lui a volé la carte ultime de sa collection *Cacao Pop*, celle du *Pingouin Blanc* : « C'est pourquoi je fais appel à vous. Un regard nouveau sur cette affaire sera peut-être plus perspicace que le mien. A présent, pour vous aider dans la résolution de cette énigme, je me propose de vous retracer ma vie du mieux possible. J'espère que cela vous aidera un peu dans vos recherches. Je ne sais pas trop par quel bout commencer... »

3.1.2 Documents d'accompagnement

Les élèves disposent pour résoudre cette énigme d'un corpus composé :

1. De la lettre d'Arthur Bardonchet qui raconte son histoire (support papier).
2. De l'accès au forum des collectionneurs qu'il fréquente sur internet et de sa page Facebook (support papier et numérique).
3. D'un plan de la ville où se déroule l'action du jeu (support papier / support numérique).
4. De la série de messages qu'a laissée le voleur sur le répondeur téléphonique d'Arthur Bardonchet (support audio).
5. D'une série de prises de vue de différents quartiers de la ville où sévit un énigmatique colleur / décolleur d'affiches *Cacao Pop* (support papier / support numérique).
6. Du jeu de cartes à collectionner *Cacao Pop* (format papier).

3.2 Procédures permettant la résolution de l'énigme

La stratégie la plus courte et la plus efficace pour résoudre la situation-problème est la suivante :

1. Lecture de la nouvelle, relevé d'indices et identification des suspects
2. Lecture du plan de la ville et mise en relation du plan et de la nouvelle
3. Compréhension acousmatique de l'enregistrement audio
4. Mise en relation de l'enregistrement audio avec le plan de la ville et la nouvelle

¹² Laurent Séjourné, Isabelle Zimmermann, Gaëlle Boubennec et Pascal Bertrand.

¹³ La notion d'énigme est à comprendre ici dans son double sens : énigme policière comme le laisse entrevoir la thématique de la situation-problème, mais également comme l'écrit Philippe Meirieu « (...) créer l'énigme : faire du savoir une énigme susceptible de mobiliser le désir de l'apprenant...chercher à faire naître l'intérêt de l'élève en laissant entrevoir le plaisir de savoir sans livrer le secret de la connaissance. » (Meirieu, 1987, p.92).

5. Lecture d'images : les affiches lacérées *Cacao Pop*
6. Lecture du forum des collectionneurs et mise en relation avec la lecture d'images et la nouvelle
7. Synthèse des preuves récupérées et écriture d'un e-mail à Arthur Bardonchet

3.3 Avantages et précautions liés à l'outil

3.3.1 Situation-problème, jeu et motivation

Il existe une littérature abondante qui traite des liens entre jeu / motivation et apprentissages. (Brougère, 2005, Sauvé et al. 2007)¹⁴ L'ensemble des activités proposées dans notre situation-problème est tourné vers la résolution de l'énigme. Bien que pluridisciplinaire, et mettant en jeu un grand nombre de procédures, l'objectif est perceptible de tous. A cet égard la situation-problème constitue une véritable plus-value dans la mesure où elle donne réellement du sens aux apprentissages. C'est cette compréhension du sens qui est la garantie d'une grande motivation de la part des élèves. Elle favorise également un engagement immédiat des élèves dans les activités. En outre, la situation-problème proposée relève pour une grande part de l'activité ludique telle que la définit Brougère.

1. Le second degré ou l'absence de littéralité de l'activité ludique ce qui renvoie au différentes modalités du faire-semblant
2. La décision, non seulement celle d'entrer dans le jeu mais plus encore le fait que le jeu n'est qu'une succession de décisions ou que jouer c'est décider.
3. La règle, qu'elle s'impose de l'extérieur au joueur, qu'elle soit négociée ou qu'elle se construise au fur et à mesure du jeu.
4. La frivolité ou la minimisation des conséquences de l'activité.
5. L'incertitude (Brougère, 2005).

La limite de cette définition au regard de notre situation-problème réside d'une part dans la décision de jouer ou non et d'autre part dans le doute que suscite la situation.

3.3.2 Le vrai et le vraisemblable

Si la majorité des élèves devrait comprendre et accepter la situation du jeu, quitte à faire semblant de croire que Arthur Bardonchet a réellement besoin d'eux pour retrouver sa carte à collectionner, quelques-uns pourront douter d'être effectivement en train de jouer. La documentation qui accompagne le jeu est vraisemblable. Que ce soit la nouvelle qui accompagne l'énoncé de la situation, la page Facebook, l'affiche photographiée en ville, le plan de la ville et même le message téléphonique, l'ensemble de ces éléments donne une certaine crédibilité à la situation. Si cette documentation, qui semble authentique, n'est sans doute pas étrangère à l'engagement des élèves dans l'activité, elle contribue à semer le doute dans l'esprit de certains... La gestion du doute nécessitera une réflexion de la part de l'enseignant de manière à ne pas « décevoir » outre mesure les élèves qui pourraient hésiter à s'engager de nouveau dans ce type d'activité.

3.3.3 Sortir du jeu. Le rôle du maître

Les élèves, pour mener à bien leur enquête, doivent mettre en œuvre un nombre significatif de procédures dont certaines sont nouvelles pour les plus jeunes. C'est pour cette raison que l'enseignant doit, par moment, extirper les élèves du jeu afin de leur proposer des procédures

¹⁴ La motivation crée « des conditions favorables à l'apprentissage et, en ayant un impact positif sur les apprentissages cognitif, affectif et psychomoteur [...], le jeu motive l'apprenant, structure et consolide les connaissances, favorise la résolution de problèmes et influence le changement des comportements et des attitudes des jeunes » (Sauvé et al., 2007, p.89).

élémentaires décontextualisées. Ce temps d'analyse doit également permettre aux élèves de rapprocher les situations nouvelles à des « familles de tâches » (Beckers, 2002) auxquelles ils sont familiers. L'enseignant doit également veiller à ce que les élèves ne soient pas exclusivement centrés sur la résolution pragmatique de l'énigme. Cette approche serait sans doute inopérante mais également dommageable pour la construction des compétences. Cette « visée instruite » (Rey, 2006) qui doit être privilégiée ici doit favoriser une démarche réflexive plutôt que spontanée. Le nombre de suspects étant limité, le hasard ou l'intuition peuvent suffire à démasquer le coupable, mais en aucun cas cette logique de l'efficacité ne permettrait de le compromettre de manière irréfutable. Si l'élève doit être invité à sortir de la logique de l'efficacité, le professeur doit également proposer des temps de stabilisation du savoir. Ces temps ne peuvent avoir lieu qu'en dehors du jeu de manière à permettre une systématisation des tâches. En outre, il peut être nécessaire de définir et de théoriser certaines notions qui « ne s'inventent pas ». C'est cette stabilisation des savoirs, qui permettra de transférer les procédures rencontrées en dehors du contexte particulier du jeu. Le rôle de l'enseignant dans la gestion des jeux dans le cadre des apprentissages est donc différent du rôle du maître « générateur de savoir » (Barthélémy-Ruiz, 2006). Sa situation se rapproche de celle qui est mise en œuvre dans les approches socio-constructivistes. Son rôle est alors de déstabiliser l'élève, de lui faire prendre conscience de la limite de ses conceptions et de le faire renoncer à ses convictions pour appréhender une posture nouvelle¹⁵.

3.3.4 Problèmes liés à la didactique des disciplines artistiques

Les habitudes de classe en matière de didactique des arts sont variées d'un enseignant à l'autre. Néanmoins, l'approche par compétences et par la résolution de situation-problème dans les domaines artistiques sont rares. En éducation musicale, les pratiques vocales sont majoritairement proposées sous forme de chants ou de jeux vocaux et la culture de l'oreille sous forme d'auditions dirigées. Ces cultures de l'oreille et de la voix sont parfois rattachées à l'histoire des arts. Dans l'outil proposé, culture vocale, culture de l'oreille et histoire des arts sont non seulement liées, mais rattachées aux autres disciplines. Pour palier au manque d'habitude lié à la pratique de situation-problème dans le domaine des arts, trois livrets pédagogiques sont proposés en éducation musicale, en arts visuel et en histoire des arts.

3.4 Évaluation des compétences proposée par l'outil

L'outil propose plusieurs entrées pour l'évaluation. La première, dichotomique, peut se résumer à la question suivante : « au terme de la séquence, l'élève ou le groupe d'élèves est-il parvenu à identifier le coupable et à justifier de sa culpabilité ? ». Cette évaluation ne propose aucun diagnostic et ne rend compte ni des procédures engagées ni des interactions au sein du groupe. Elle permet en revanche dans le cas où la réponse est affirmative de supposer les procédures engagées et de valider leur mise en œuvre. Afin de proposer un outil de diagnostic, les évaluations en éducation musicale ont été divisées en trois phases (Rey, 2006)¹⁶. Trois évaluations sont proposées dans le livret du professeur, chacune devant correspondre à trois moments de la séquence (début – milieu – fin). Ces évaluations ne comportent aucun caractère obligatoire et ne permettent pas à

¹⁵ « On ne distribue plus les connaissances, on accompagne plutôt les élèves dans leurs découvertes. D'enseignant, on devient animateur, meneur de jeu [...] il faut donc coiffer une autre casquette [...]. Et, même si l'on voit comment les élèves devraient s'y prendre, rester très discret pendant le temps de jeu, consacré à une observation dont les éléments seront utilisés au moment du débriefing. » (Barthélémy-Ruiz, 2006, p.20)

¹⁶ « 1. Savoir exécuter une opération (...) en réponse à un signal (...) « compétence élémentaire » ou « compétence de premier degré ». 2. posséder toute une gamme de ces compétences élémentaires et savoir, dans une situation inédite, choisir celle qui convient (...) nous parlerons de « compétences élémentaires avec cadrage » ou de compétence de deuxième degré » ; 3. Savoir choisir et combiner correctement plusieurs compétences élémentaires pour traiter une situation complexe. Nous parlerons alors de « compétence complexe » ou de « compétence de troisième degré » (Rey, 2006, p. 26).

proprement parler de résoudre l'énigme. Leur mise en œuvre est donc à la discrétion de l'enseignant.

4. La réception du dispositif

4.1 De la définition de l'histoire des arts

Deux enseignantes de CM1 ont été volontaires pour se prêter à l'évaluation du dispositif. La première enseignante (que nous appellerons E1) a 38 ans et 12 ans d'ancienneté dans la même école rurale. Elle dit pratiquer depuis toujours l'histoire de l'art, mais « en partant seulement de l'histoire » (discipline qu'elle assure, dit-elle, avec aisance) et en s'appuyant seulement sur les « arts plastiques ». Concernant l'éducation artistique, elle a plus de penchant pour la musique, mais se dit maintenant attentive, avec l'introduction de « l'histoire des arts » à toucher à tous les arts. La seconde (E2) a une quarante d'années. Elle enseigne depuis 2 ans dans une école péri-urbaine, après avoir passé plusieurs années en poste à l'étranger. Elle dit « ne pas savoir exactement ce que l'on entend par histoire des arts » et attend « qu'on lui explique ». Elle pratique des échanges de service en assurant l'histoire en CM2 et affiche une plus grande expertise en arts visuels et en théâtre. Les deux enseignantes n'ont pas caché que leur motivation première pour participer à cette recherche était un besoin d'outils pour mener à bien ce nouvel enseignement.

Les propos tenus par la conseillère pédagogique finistérienne montrent que cette recherche d'outils est la principale question qui se pose en dehors du problème du temps et de l'évaluation. « Quelques éditeurs ont commencé à s'approprier ce nouveau domaine, mais les outils proposés restent souvent très disciplinaires. Souvent, le choix a été fait d'intégrer ce domaine au manuel d'histoire. Un certain nombre de manuels ou guides pédagogiques sont parus dans le domaine de l'histoire de l'art (arts de l'espace, arts du quotidien, arts visuels), beaucoup moins dans le domaine des « arts du son » ou « des arts du spectacle vivant ». D'autres outils, tenant le pari d'intégrer l'ensemble des domaines artistiques, se sont attachés à suivre la chronologie définie par les périodes historiques au programme. S'ils couvrent l'ensemble des périodes ainsi que l'ensemble des domaines artistiques, ces outils n'invitent pas véritablement à créer de liens entre les domaines et les périodes. Pour le moment, les enseignants convaincus de l'importance de cet enseignement et soucieux de l'introduire dans leur classe travaillent avec des outils divers : CDthèques d'écoles ou personnelles, DVD, Sites internet, reproductions d'œuvres picturales essentiellement. Très peu d'entre eux ont opté pour un manuel. Tous avouent tâtonner, s'éparpiller et ne pas parvenir à structurer cet enseignement de l'histoire des arts »¹⁷.

4.2 Leur vision de l'outil

Pour E1, ce genre de support s'apparente à un « projet arts », comme un autre puisqu'elle en fait un chaque année. Pour E2, c'est un projet « multi-activités », qui lui permet « de rompre un peu avec le travail traditionnel, pour qu'il y ait une rupture dans l'enchaînement des séances ». Toutes les deux n'ont pas l'habitude de mener des projets collectifs avec toute l'école : « Dans quelques directions qu'on aille, la classe et moi c'est notre groupe, ensemble et on va dans une direction » (E2). Quand on aborde la question du jeu, les réponses sont partagées : « ça semble être un jeu à certains moments, mais même temps, ça ressemble à des activités scolaires, comme on a l'habitude de les faire en classe, lire un texte, des mots que l'on ne comprend pas, on va chercher dans le dictionnaire du vocabulaire, on fait du travail sur les synonymes, ça c'est du travail classique et puis après, quand on cherche le voleur, là, ça sera plutôt du jeu. » (E2).

¹⁷ Entretien CPEM, juin 2011.

4. 3 Les évaluations envisagées

D'emblée, il s'agira essentiellement pour toutes les deux d'évaluations par compétences. Tous les domaines abordés leur semblent évaluable. En ce qui concerne l'histoire des arts, pour E1, c'est très clair, puisque lors de l'entretien, elle liste spontanément les compétences en question, en précisant qu'elle a toujours travaillé ainsi dans sa classe. « Concernant le livret des compétences du socle commun, j'utilise les compétences qui y figurent (pas toutes) et j'ai complété mon propre livret avec des compétences plus détaillées car certaines compétences proposées me semblaient trop vagues ou pas évaluable (...). Certaines ne sont même pas des compétences d'ailleurs ! Enfin bref, je n'y adhère pas complètement alors j'ai fait ma propre petite cuisine... Nous on a décidé que chacun avait son livret, Gaëlle (*enseignante de la même école, en CM2*) et moi, on a notre propre livret. En fait, ce qui s'est passé, c'est que l'année dernière on a acheté celui préconisé par l'inspection, qu'on ne connaissait pas du tout, et quand on l'a reçu, Gaëlle et moi, on a été catastrophées, parce que c'était trop général, y'avait des compétences qui n'étaient pas des compétences, c'était des titres, comme s'il fallait évaluer un titre de leçon, enfin...il est pas si bien fait que ça ! On trouve ça très succinct, c'est difficile d'évaluer sur des choses aussi générales. »

Les compétences qu'elle a retenues et listées en histoire des arts dans le cadre d'un travail avec le dispositif proposé sont les suivantes : a) avoir des connaissances et des repères relevant de la culture artistique ; b) rencontrer des œuvres, les identifier et nommer leur auteur ; c) situer dans le temps, l'espace et les civilisations des œuvres artistiques ; d) définir et caractériser simplement une œuvre ; e) être sensible aux enjeux esthétiques et humains d'une œuvre artistique ; f) être capable de porter un regard critique sur une œuvre.

Pour E2, l'évaluation envisagée en histoire des arts est beaucoup moins précise, même si elle a prévu de critères d'évaluation pour tous les domaines abordés dans le support. Lorsqu'on aborde la question, elle revient sur son souci de définir ce qu'est l'histoire des arts : « alors, d'abord, je ne sais pas ce que c'est que l'histoire des arts, je sais ce que ça veut dire pour moi, mais vous allez me dire ce que c'est que l'histoire des arts pour vous parce que moi j'aimerais savoir si on parle de la même chose, vous et moi, quand on dit histoire des arts ». Lorsqu'on lui demande ce qu'elle pense des items figurant dans le livret des compétences du socle commun, elle semble assez perplexe : « Oui il y a, je ne sais même plus ce qu'il y a (...) il peut y avoir le discours et puis il peut y avoir la réalité. (*Elle va chercher le livret*). Ici, par exemple, on vous dit rencontrer des œuvres, oui nous rencontrerons des œuvres. Développer la curiosité pour de nouvelles œuvres, oui, moi j'ai des élèves qui diront : « j'ai vu ça chez moi dans un livre. » On voit que quand on reste sur un œuvre, sur un artiste, ça reste, il y a des choses qui restent, c'est un bagage qui reste (...) Ce que je vais cocher chez les enfants, c'est *Rencontrer des œuvres* et *Développer la curiosité*, ça oui ! (...) *Situer dans son époque*, et *Comprendre une œuvre dans son époque*, je le ferai à un autre moment, oui plutôt en histoire. » Dans le dispositif, il y a possibilité de travailler sur des événements artistiques précis (années 70-80), mais, au contraire de E1, elle ne compte pas s'emparer pas de cette possibilité, qui pour elle, dépasse l'histoire des arts.

Conclusion

Il s'agit ici d'une étude en cours qui ne prend en compte qu'une analyse à priori des enseignants. En amont de l'évaluation en histoire des arts, la question de l'évaluation en éducation artistique se pose avec toujours autant d'acuité. Les recherches sur l'évaluation des compétences en éducation artistique à l'école primaire restent rares. Les ouvrages généraux sur l'évaluation des compétences scolaires éludent souvent d'ailleurs les savoirs à acquérir dans ce domaine du « sensible », cantonnant ce terrain au domaine du seul « savoir être » (Guimard, 2010, Gimmonnet, 2007). Il semble que les recherches concernant les théories de la transposition didactique en éducation artistique continuent de se heurter au manque de lisibilité de l'enjeu didactique, du contenu à apprendre et d'une évaluation appropriée (Vendramini, séjourné, 2009). Cette recherche tente d'apporter un éclairage nouveau sur ces questions en proposant des cheminements possibles et abordables pour les enseignants polyvalents du premier degré. La mise en œuvre du dispositif proposé est actuellement en cours dans trois écoles primaires et devrait continuer d'alimenter le

débat entre chercheurs et acteurs de terrain à propos des stratégies didactiques soulevées par cet élargissement de l'éducation à l'art.

Bibliographie

Barthélémy-Ruiz Chantal (2006). Le mariage de l'eau et du feu ? Jeu et éducation à travers l'histoire. *Les cahiers pédagogiques*, 448, 20.

Beckers J. (2002) *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*. Bruxelles: Labor

Brougère G. (2005). *Jouer/Apprendre*. Paris : Economica-Anthropos.

Brougère G. (1995). *Jeu et éducation*. Paris : L'Harmattan

E.S.E.N (2010). Dossier d'accompagnement des enseignants pour la mise en œuvre de l'histoire des arts <http://www.esen.education.fr/fr/ressources-par-theme/enseignements/l-enseignement-de-l-histoire-des-arts/>.

Gimmonet B. (2007). *Les notes à l'école*. Paris : L'harmattan

Guimard P. (2010). *L'évaluation des compétences scolaires*. Rennes: PUR.

Lavin M. (1998). *Histoire des arts, l'émergence d'un enseignement*. Paris : Hachette Education.

Meirieu P. (1987). *Apprendre, oui... mais comment*. Paris : ESF

Rey B. (1996) *Les compétences transversales en question*. Paris : ESF

Pivec M., Pivec P.(2009). *What do we know from research about the use of games in education ?*, http://games.eun.org/upload/gis-full_report_en.pdf (Chapter 7). Consulté en janvier 2012.

Rey, B., Carette, V., Defrance, A. & Kahn, S. (2006). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles : De Boeck

Pire J-M (2010). Vers l'intégration de l'histoire des arts dans la scolarité : la fin d'une exception française. In D.Poulot, J.M. Pire & A. Bonnet (Ed.) « *L'Education artistique en France, du modèle académique aux pratiques actuelles* ». Rennes : PUR.

Sauvé L, Renaud L. & Gauvin M. (2007). Une analyse des écrits sur les impacts du jeu sur l'apprentissage. *Revue des sciences de l'éducation*, 33, 89-107.

Vendramini C., Gourlaouen G. (2011). *Histoire des arts à l'école primaire : questions sur une didactique nouvelle*, Que valent les apprentissages en histoire, géographie et éducation à la citoyenneté?" INRP, Lyon, actes à paraître.

Vendramini, Séjourné (2009), « *Création artistique via une interface de captations sonores et gestuelles en milieu scolaire : étude de faisabilité* ». In Georges-Louis Baron, Eric Bruillard, Luc-Olivier Pochon, *Informatique et progiciels en éducation et en formation* », Collection Technologies nouvelles et éducation, INRP, 132-149.

LES COMPETENCES PERSONNELLES ET SOCIALES EN QUESTION. ILLUSTRATION A TRAVERS DEUX OUTILS D'EVALUATION EN EDUCATION PHYSIQUE ET SPORTIVE

Lenzen Benoît*, Dénervaud Hervé**, Poussin Bernard***

* Institut des sciences du mouvement et de la médecine du sport & Institut universitaire de formation des enseignants, Université de Genève, Benoit.Lenzen@unige.ch

** Institut des sciences du mouvement et de la médecine du sport, Université de Genève, Herve.Denervaud@unige.ch

*** Institut universitaire de formation des enseignants & Institut des sciences du mouvement et de la médecine du sport, Université de Genève, Bernard.Poussin@unige.ch

Mots-clés : compétences personnelles, compétences sociales, éducation physique et sportive, évaluation

Résumé. Dans cette contribution, nous questionnons la notion de compétences personnelles et sociales, avatar de la notion de compétence que l'on retrouve fréquemment aujourd'hui dans les documents professionnels, instructions officielles et autres référentiels de compétences. En nous appuyant sur une littérature professionnelle et scientifique conséquente, nous démontrons d'abord la fragilité conceptuelle de cette notion. Nous illustrons ensuite cette (re)mise en question à travers la confrontation de deux outils d'évaluation en éducation physique et sportive (EPS). Cela nous conduit finalement à proposer de substituer à cette notion insatisfaisante de compétences personnelles et sociales la notion plus consistante de compétences dans et autour de l'action.

1. Introduction

Depuis une trentaine d'années, l'approche par compétences tend à s'imposer dans les systèmes d'éducation et de formation professionnelle de nombreux pays francophones, suscitant moult débats quant à sa pertinence, sa fragilité conceptuelle, sa légitimité, etc. (Bronckart, 2009 ; Crahay, 2006). Il est généralement admis que la notion de compétence a d'abord été introduite en linguistique, dans le cadre des premiers travaux de Chomsky (Bronckart, 2009 ; Rey, 1996). Elle désignait alors un pouvoir d'engendrement et d'adaptation des actions, une puissance générative (Rey, 1996). Mais une autre vision de la compétence s'est développée parallèlement dans les milieux du travail et de la formation des adultes (Bronckart, 2009 ; Rey, 1996). Dans ce cadre, la compétence est plutôt attachée aux conditions particulières d'une situation donnée, spécifique à la tâche ou au type de tâches qu'elle permet d'accomplir avec efficacité (Rey, 1996). C'est, nous semble-t-il, sur cette dernière acception que s'est construite l'approche par compétences actuelle.

Très vite, la notion de compétence s'est accompagnée de son lot d'avatars : e.g., compétences transversales ; compétences spécifiques, démultiplicatrices, stratégiques et dynamiques (Leclercq, 1987). Les compétences personnelles et sociales s'inscrivent dans cette lignée. Comme Rey l'a remarquablement fait il y a une quinzaine d'années pour la notion de compétences transversales, il nous a semblé utile de questionner, de manière évidemment moins approfondie ici, la solidité conceptuelle de cette notion de compétences personnelles et sociales, pour ensuite illustrer cette

(re)mise en question à travers la confrontation de deux outils d'évaluation en éducation physique et sportive (EPS).

2. Les compétences personnelles et sociales...

Les compétences personnelles et sociales sont souvent couplées, voire même associées avec d'autres « familles » de compétences. C'est par exemple le cas dans l'illustration construite par Jourdan (2010, p. 84) sur la base de la circulaire n° 98-237 du 24 novembre 1998 qui concerne l'éducation à la santé et à la citoyenneté en France et sur les contenus du socle commun de connaissances et de compétences (décret n° 2006-830 du 11 juillet 2006). Selon cette illustration, éduquer à la santé et à la citoyenneté, c'est permettre à l'élève de (a) développer ses *compétences personnelles, sociales* et civiques (pôle plutôt centré sur le développement d'attitudes), (b) acquérir les moyens d'un regard critique vis-à-vis de son environnement (pôle plutôt centré sur l'apprentissage de capacités), et (c) connaître son corps, sa santé, les comportements et leurs effets (pôle plutôt centré sur l'apprentissage de connaissances). C'est également le cas dans une brochure pour les enseignants d'EPS intitulée « *Compétences personnelles et sociales* en éducation physique et sportive », développée par l'Office fédéral du sport suisse (OFSP, 2010) dans le cadre du projet « Qualité de l'éducation physique et sportive » (Qeps.ch)¹.

Dans la littérature anglophone, le concept de *life skills* (Jones & Lavalley, 2009 ; Petitpas, Van Raalte, Cornelius & Presbrey, 2004) nous semble proche de la notion de compétences personnelles et sociales. En effet, les *life skills* sont définies par l'Organisation mondiale de la santé (WHO, 1999, cité par Jones & Lavalley, 2009, p. 159) comme « *la capacité à adopter un comportement positif et adaptatif qui permet aux individus d'affronter efficacement les exigences et les challenges de la vie quotidienne* » (traduit par nous). Par ailleurs, dans une étude explorant les besoins d'athlètes adolescents britanniques en matière de *life skills*, Jones et Lavalley (2009) ont catégorisé inductivement les besoins exprimés en distinguant *personal life skills* et *interpersonal life skills*. On y retrouve bien la distinction, mais également l'apparente complémentarité entre compétences personnelles et sociales. A l'inverse, dans d'autres publications, les notions de compétences personnelles et sociales sont traitées séparément, avec des acceptions pouvant être très différentes ainsi que nous allons le montrer dans les deux sections suivantes.

2.1 Les compétences personnelles

Les compétences personnelles sont définies par Kanning (2004, cité par OFSP, 2010, p. 9) comme « *l'ensemble des capacités, des aptitudes et des connaissances nécessaires à une personne pour s'autogérer* ». Pour l'EPS, elles sont déclinées en deux facettes : l'*autoréflexion* et l'*autogestion*. La première est elle-même déclinée en (a) prise de conscience de ses propres forces et faiblesses, (b) fixation d'objectifs individuels réalistes, et (c) identification et expression de ses besoins, tandis que la seconde est décomposée en (a) régulation de l'attention et de la concentration, (b) régulation de la motivation et de la volonté, et (c) régulation des émotions (compétences émotionnelles). C'est cette même vision des compétences personnelles qui est à la base de certains modèles et instruments d'évaluation dans la littérature anglophone. Ainsi, pour évaluer les compétences personnelles d'adolescents en vue de les mettre en relation avec leur consommation de drogues, Griffin, Scheier, Botvin et Diaz (2001) ont combiné les scores de différents questionnaires pour obtenir trois indicateurs représentatifs selon eux des compétences personnelles : la prise de décision, l'autocontrôle et l'autorégulation.

¹ Qeps.ch est un outil d'évaluation de la qualité de l'enseignement de l'EPS, développé par l'Office fédéral du sport suisse à destination des acteurs de l'EPS des différents cantons. Le canton de Genève a opté pour une version hybride de qeps.ch, en cours de développement par un groupe de travail dont nous faisons partie, afin de tenir compte des spécificités cantonales en matière d'EPS.

En revanche, d'autres modèles et instruments d'évaluation révèlent une toute autre vision de la notion de compétences personnelles. Ainsi, le modèle discuté par Greenspan et Driscoll (1997) associe à cette notion les quatre éléments suivants : (a) la compétence académique (incluant l'intelligence conceptuelle) ; (b) la compétence quotidienne (intelligence pratique et sociale) ; (c) la compétence physique ; et (d) la compétence affective. Dans ce modèle, la frontière ne semble pas très nette entre compétences personnelles et compétences sociales. Le modèle de McGrew et Bruininks (1990), qui inclut les dimensions de (a) compétence physique, (b) intelligence pratique, (c) intelligence conceptuelle et (d) compétence émotionnelle, est très proche du précédent. Dans ces deux modèles, on remarquera l'amalgame entre les notions de compétence et d'intelligence, notions que nous retrouverons associées dans d'autres modèles relatifs aux compétences sociales. Autres exemples : lorsque Pender, Walker, Sechrist et Frank-Stromberg (1990, cités par Wallston, 1992) ont repris et adapté la *Perceived Competence Scale* de Wallston (1989, cité par Wallston, 1992), ils l'ont rebaptisée *Personal Competence Rating Scale*. Des items de cet instrument tels que « *je réussis dans les projets que j'entreprends* » et « *normalement, mes plans ne fonctionnent pas bien* » (traduits par nous) témoignent du fait que la notion de compétence personnelle y est considérée comme une sorte d'auto-efficacité généralisée, en référence aux travaux de Bandura (1977). Dans l'étude de Ahmed, Minnaert, van der Werf et Kuyper (2010), la compétence personnelle est référée à un domaine disciplinaire, en l'occurrence les mathématiques. La *Personal competence beliefs measure*, un des instruments utilisés dans cette recherche, évalue en fait (a) à quel point les participants s'estiment compétents en mathématiques, (b) à quel point ils s'attendent à être performants à l'avenir en mathématiques, et (c) à quel point ils s'estiment capables d'apprendre efficacement quelque chose de nouveau en mathématiques. De la même façon, la notion de compétence personnelle est donc considérée par ces chercheurs comme une sorte d'auto-efficacité, circonscrite dans ce cas au domaine des mathématiques. En conclusion, malgré son usage fréquent dans les documents professionnels, les instructions officielles (programmes, plans d'études...) et autres référentiels de compétences, il apparaît que la notion de compétences personnelles est une notion pour le moins floue et ambivalente.

2.2 Les compétences sociales

Les compétences sociales constitueraient quant à elles « *l'ensemble des connaissances, des capacités et des aptitudes qui permettent à une personne d'améliorer la qualité de son comportement social* » (OFSPO, 2010, p. 9). Pour l'EPS, elles sont déclinées en quatre facettes : (a) la *communication*, elle-même déclinée en expression claire et adéquate, écoute et compréhension, et expression non verbale appropriée ; (b) la *coopération*, décomposée en disposition à jouer des rôles, participation active, et maintien de la distance nécessaire par rapport au groupe pour conserver son identité personnelle ; (c) la *gestion des conflits*, déclinée en reconnaissance et approche des conflits, prise en compte des intérêts et objectifs des deux parties, et recherche d'une solution commune ; et (d) la *morale*, subdivisée en connaissance et application des règles, respect des règles même en présence d'autre besoins, et fair-play.

Dans un modèle de référence pour le diagnostic du retard mental, Greenspan et Grandfield (1992) subdivisent la compétence générale en compétence instrumentale d'une part, et en *compétence sociale* d'autre part. Chacune de ces composantes est elle-même subdivisée en aspects intellectuels et non intellectuels. Pour la *compétence sociale*, les aspects non intellectuels sont subdivisés en deux dimensions de la personnalité : le *tempérament* et le *caractère* ; les aspects intellectuels renvoient quant à eux d'une part à l'*intelligence pratique*, typiquement évaluée à l'aide de checklists de comportements adaptatifs à la vie quotidienne, et d'autre part à l'*intelligence sociale*, désignant l'habileté à comprendre et gérer efficacement les objets et événements de la vie interpersonnelle et sociale, et incluant des dimensions telles que la prise de rôle, l'empathie, la perception d'autrui, le jugement moral, la communication, etc. La composante *compétence sociale* de ce modèle s'appuie sur un important corpus de littérature théorique et empirique sur la

personnalité et le développement social. Cette littérature est le lieu de tensions entre deux conceptions de la compétence sociale : une conception *orientée vers le résultat* et une conception *orientée vers les traits*. La première se focalise sur le succès des individus dans les rôles sociaux importants (vie amoureuse, emploi, indépendance...). La seconde consiste à définir et évaluer la compétence sociale à partir de traits comportementaux particuliers supposés pouvoir prédire le succès dans les rôles sociaux fondamentaux. Il y a 20 ans, cette dernière était la plus répandue. Nous reviendrons sur ce constat dans la discussion.

Le modèle SOCIAL (*socio-cognitive integration of abilities model*) de Beauchamp et Anderson (2010), beaucoup plus récent, cherche à rendre compte de l'importance des influences biologiques et environnementales sur le développement de compétences sociales. Il souligne notamment le rôle de médiateurs internes (tempérament, personnalité, attributs physiques) et externes (environnement familial, statut socio-économique, culture) dans l'émergence de la fonction sociale. Il détaille également les facteurs cognitifs et affectifs qui déterminent la présence et l'intégrité des habiletés sociales : (a) fonctions d'attention et d'exécution (contrôle attentionnel, flexibilité cognitive, fixation de buts) ; (b) communication ; et (c) habiletés socio-émotionnelles (perception de l'émotion, attribution, empathie, raisonnement moral). Ce modèle ne contribue pas véritablement à un éclaircissement de la notion de compétence sociale en tant que telle. En revanche, il pointe quelques leviers sur lesquels un enseignant pourrait agir en vue de développer des compétences qualifiées de sociales chez ses élèves (facteurs cognitifs voire affectifs), tout en mettant en exergue un certain nombre de variables médiatrices sur lesquelles un enseignant a peu de prises, relativisant d'autant l'enseignabilité de compétences sociales à et par l'école. En conclusion, si la notion de compétences sociales, tout aussi répandue dans la littérature scientifique et professionnelle que la notion de compétences personnelles, apparaît moins floue et ambivalente que cette dernière, elle n'en reste pas moins abstraite et peu opérationnelle.

3. ...Une notion en contradiction avec le concept de compétence

Comme nous le rappelions en introduction, la notion de compétence n'en finit pas d'alimenter débats et controverses dans les milieux scolaire et académique. Il semble toutefois qu'un consensus soit en train de se dégager – et les échanges observés et vécus durant ce colloque renforcent cette impression – autour de définitions de la compétence proches de celle-ci : « *une compétence repose sur la mobilisation, l'intégration et la mise en réseau d'une diversité de ressources [internes et externes] [...] Cette mobilisation des ressources s'effectue dans une situation donnée, dans le but d'agir : la compétence est nécessairement située ; pour autant, elle s'exerce dans une diversité de situations, à travers un processus d'adaptation et pas seulement de reproduction de mécanisme [...] il est difficile de définir la notion de compétence indépendamment de son cadre d'apprentissage et des situations prévues pour son évaluation* » (IGEN, 2007, pp. 11-12, cité par Mottier Lopez & Tessaro, 2010, p. 32). De cette définition, nous retenons deux caractéristiques majeures. La première, c'est que la compétence est inévitablement située. La seconde, c'est qu'elle est difficilement définissable en dehors des situations conçues pour l'enseigner et l'évaluer. Ces deux caractéristiques renvoient à la notion de compétence adaptative², définie par Kolb (1984) comme les habiletés requises pour accomplir efficacement une tâche spécifiée, autrement dit comme la congruence entre les habiletés d'un individu et les exigences de la tâche à accomplir.

Par conséquent, considérer isolément des compétences personnelles et/ou sociales revient à les couper des situations spécifiques à la tâche ou au type de tâches qu'elles permettent d'accomplir avec efficacité, autrement dit à conférer à ces compétences un caractère de transversalité dont Rey

² Ce qualificatif apparaît aujourd'hui redondant, puisqu'il figure explicitement dans la définition-même de la compétence selon l'IGEN.

(1996) a précisément démontré l'inconsistance. Prenons un exemple pour illustrer cette objection. Pour fustiger la norme de complexité inédite inhérente à la notion de compétence, Crahay (2006) prenait l'exemple d'un chirurgien qui, réussissant pour la quarantième fois une transplantation cardiaque, ne ferait dès lors pas preuve de compétence. Sans nous prononcer sur la pertinence de ce raisonnement, envisageons maintenant ce même chirurgien sous l'angle de ses compétences qualifiées de sociales. Pour avoir réussi quarante fois une transplantation cardiaque, il est vraisemblablement habile à diriger son équipe dans le bloc opératoire. Il communique très clairement ses directives, collabore efficacement avec ses assistants, fait preuve d'empathie et de respect à leur égard, etc. En revanche, lorsqu'il s'agit d'informer ses patients et leur famille des suites post-opératoires, il se montre peu agréable, semble toujours pressé, fait preuve de peu d'écoute et d'empathie à leur égard, etc. Cet exemple, bien que fictif, n'en est pas moins plausible. Dans ce cas de figure, comment doit-on qualifier les « compétences sociales » de ce chirurgien ? Est-il « compétent socialement » ou ne l'est-il pas ?

4. Illustration en EPS

Pour prolonger et illustrer de façon moins caricaturale cette objection purement théorique à l'encontre de la notion de *compétences personnelles et sociales*, nous allons nous appuyer sur deux outils d'évaluation en EPS. Le premier est issu de la version fédérale du système Qeps.ch. Le second est issu de la version hybride genevoise actuellement en cours de développement.

4.1 L'évaluation des compétences sociales et disciplinaires dans des tests différents

Dans l'« espace qualité » du site Internet www.qeps.ch, nous pouvons trouver des exemples d'applications concrètes destinées aux enseignants et visant à évaluer (a) les « compétences personnelles et sociales » des élèves et (b) les « compétences disciplinaires »³. Dans le répertoire disponible, nous commencerons notre illustration par un exemple pratique nommé « radeau⁴ », en précisant que les différents tests soumis à l'analyse sont tous destinés à des élèves de 15 à 18 ans qui sont engagés dans une scolarité post-obligatoire (lycée ou école professionnelle). Dans ce test, il est annoncé que l'évaluation porte sur les compétences sociales des élèves, plus spécifiquement sur la compétence de coopération et ceci à travers deux de ses facettes : (a) la disposition des élèves à jouer des rôles ; et (b) leur participation active. Dans ce jeu, les élèves sont répartis par groupes de six joueurs et deux groupes s'affrontent par partie. Chaque groupe dispose de deux tapis et les élèves ont pour tâche de traverser la salle sur les tapis sans qu'aucun élève ne pose le pied par terre. Le but du jeu est de traverser la salle le plus vite possible. Après chaque passage, les élèves sont invités à faire le point sur la technique utilisée pour chercher à l'améliorer et être de plus en plus rapide. Les critères retenus pour observer et évaluer l'activité des élèves sont : (a) la répartition des tâches au sein du groupe ; (b) l'efficacité du rôle au sein du groupe ; (c) le fair-play ; (d) la participation active de tous à la recherche de solutions ; (e) la capacité à faire des compromis ; (f) la capacité à coopérer ; et (g) l'aide mutuelle.

Incontestablement, cette tâche engage les élèves à communiquer entre eux autant pendant l'activité qu'entre les différents duels contre les autres équipes. L'écoute et le soutien mutuel, le respect des différents niveaux d'habileté, la répartition des rôles entre les élèves, la recherche active de solutions seront alors observés pour évaluer les « compétences sociales » des élèves. Mais communiquer, échanger et coopérer pour quoi ? Certainement pour être plus efficace – en l'occurrence plus rapide comme le précise le but de la tâche – dans le déplacement, c'est-à-dire dans l'habileté motrice mise en jeu dans ce test. Toutefois, aucun critère n'est prévu pour évaluer

³ Dans cette partie, nous reprenons la terminologie telle qu'elle est utilisée dans qeps.ch, en la mettant entre guillemets.

⁴ <http://www.sport.admin.ch/q-shop/selbstsozialkompetenzen.php?ll=1>, page consultée le 7.2.2012.

le résultat de l'action (le temps pour traverser la salle). Or, s'il faut donner un sens à l'activité sociale de ce jeu, nous pensons que c'est bien pour aller plus vite. Ainsi, les habiletés motrices, au cœur des enjeux d'apprentissage de l'EPS, ne font l'objet d'aucune évaluation. Par conséquent, nous pensons que ce test évalue les élèves dans une situation de collaboration et de coopération factice, car non reliée à des enjeux de savoirs disciplinaires, d'habiletés motrices en l'occurrence. L'enjeu d'efficacité (en termes de performance ou de maîtrise) étant occulté, les élèves vont alors pouvoir s'engager dans une activité sociale à moindre frais car déconnectée de savoirs consistants et d'enjeux propres aux activités physiques, sportives et artistiques (APSA) enseignées traditionnellement en éducation physique et sportive.

Qeps.ch ne propose pas uniquement des exemples pratiques pour évaluer les « compétences personnelles et sociales ». D'autres tests sont proposés pour évaluer les « compétences disciplinaires ». Là, la référence à la culture sportive et à la pratique sociale est bien plus visible que pour l'exemple du « radeau » présenté précédemment. Par exemple, pour la course d'endurance, le test de Cooper de 12 minutes cherche à tester les élèves sur leur capacité à utiliser et améliorer leur potentiel aérobie. Le but du test est de courir la plus grande distance possible sans s'arrêter ni marcher. L'épreuve se déroule sur un stade ou sur une piste balisée et l'organisation prévoit que la moitié des élèves courent pendant que les autres relèvent les distances et les arrêts. En toute logique, les critères d'observation portent sur ces mêmes aspects : nombre d'arrêts et distance accomplie. Ils serviront à évaluer la compétence des élèves à « appliquer les principes de bases de l'endurance pour entraîner cette capacité⁵ ». A l'inverse du test précédent qui évaluait de manière exclusive les « compétences sociales », ce test se centre exclusivement sur les « compétences disciplinaires ». Cela ne nous satisfait pas plus tant il est vrai que les enjeux éducatifs de l'enseignement de la course d'endurance sont absents de cette évaluation et que celle-ci est alors enseignée pour elle-même et non pas pour les valeurs qu'elle peut véhiculer et qui constituent un des facteurs décisifs pour la programmation d'une APS (au profit ou au détriment d'une autre) en milieu scolaire. Ces enjeux éducatifs pour la course d'endurance peuvent porter sur des aspects d'« autoréflexion » sur ses propres capacités aérobies, sur des aspects de coopération entre les élèves pour échanger des informations sur la course ou sur des aspects de « morale » liés au respect des règles de l'activité.

4.2 L'évaluation « dans » et « autour » de l'action dans un même test

Dans sa version hybride, qeps.ch offre la possibilité aux cantons d'adapter la démarche « qualité » aux contextes locaux de l'enseignement de l'EPS. A Genève, la lecture que nous faisons du concept de compétence, développée en début d'article, nous a amenés à postuler que les compétences « disciplinaires » et « personnelles/sociales » ne pouvaient être dissociées et devaient faire l'objet d'une évaluation dans un même test. Parmi les tests développés dans cette optique, nous retiendrons un test d'endurance consistant à accomplir trois courses de 600 mètres avec une période de repos (équivalente à la moitié de la durée de la course précédente) entre chaque course. L'épreuve se déroule sur un stade ou un terrain balisé et les élèves sont regroupés en binômes (un coureur et un « observateur », qui changeront de rôle une fois la course terminée). Les compétences évaluées dans cette épreuve sont de (a) préparer un projet de course incluant des temps de récupération passive pour reproduire une performance au maximum de ses possibilités aérobies et (b) prendre les temps de passage et gérer les temps de récupération sur la course d'un camarade pour valider sa performance. Les critères d'observation retenus sont : (a) le temps cumulé des trois courses ; (b) le pourcentage de vitesse maximum aérobie (VMA⁶) du coureur auquel la course a été accomplie ; (c) l'écart entre le projet de course formalisé avant la course et

⁵ <http://www.sport.admin.ch/q-shop/kompetenzraster.php>, page consulté le 7.2.2012.

⁶ La VMA est la vitesse à laquelle la consommation du métabolisme en oxygène est maximale. Selon les sujets, cette vitesse peut être tenue de manière continue entre 4 et 7 minutes.

le projet réalisé ; et (d) la fiabilité de la prise des temps de passage et la gestion de la durée des temps de repos entre chaque course.

Cette épreuve vise donc à évaluer de manière simultanée des compétences « dans l'action » (compétences « disciplinaires » dans qeps.ch) et « autour de l'action » (compétences « personnelles et sociales » dans qeps.ch). Ce type d'évaluation nous semble plus cohérent autant d'un point de vue théorique que pratique. En effet, dans ce test les savoirs disciplinaires sont premiers. Cette épreuve confronte les élèves à des enjeux de savoirs spécifiques à la course d'endurance. Ainsi, l'épaisseur culturelle de cette APS est conservée, en veillant toutefois à ce que cette culture ne soit pas seulement une culture en mot mais bien une culture en acte, adaptée au public et à une pratique scolaire (Dhellemmes et al., 2007). A travers cette composante culturelle, nous pensons que la question du sens de l'action et du mobile pour agir trouve des éléments de réponse. Pratiquement, cela se traduit pour l'élève par l'exigence de performance et donc de modification de sa motricité. En course d'endurance, cela se mesure par le temps accompli pour effectuer la distance prévue et le pourcentage de VMA auquel la course est réalisée. Ceci est évalué dans l'épreuve. Toutefois, dans cette épreuve, il y a d'autres enjeux de formation, indissociables des enjeux disciplinaires. En effet, la réussite et donc la compétence passe aussi par l'apprentissage d'autres ressources : (a) des connaissances (sur le règlement de la course, sur son potentiel aérobie par exemple) ; (b) des habiletés méthodologiques (pour utiliser les fonctions d'un chronomètre par exemple) ; et (c) des attitudes (pour tenir à jour son carnet d'entraînement par exemple). Ainsi, ces compétences « autour de l'action » devront se développer pour permettre une efficacité « dans l'action ». Pour ce faire, différents rôles sociaux doivent être mis en scène afin d'engager les élèves à collaborer, à échanger, à coopérer pour devenir des coureurs d'endurance compétents dans cette épreuve. Ces compétences « autour de l'action » s'évaluent via deux indicateurs : (a) l'écart entre le projet de course et le temps effectivement réalisé ; et (b) la fiabilité de la prise des temps de passage et la gestion des temps de repos. Pourquoi ces critères ? Etablir un plan de course réaliste (temps de course et pourcentage de VMA) engage les élèves à une autoréflexion sur leurs capacités aérobies et à une analyse des informations collectées lors des leçons précédentes durant leurs entraînements. Relever les temps d'un camarade et gérer ses temps de repos engagent à une attention envers l'activité de l'autre ainsi qu'à une habileté méthodologique spécifique à la course d'endurance. De plus, cela va permettre de communiquer avec son camarade sur un objet précis, l'efficacité de son action, un peu comme pourrait le faire un entraîneur avec son athlète.

Ainsi, les compétences « dans l'action » et « autour de l'action » se conjuguent dans un même projet, finalisé par une épreuve unique. Cette perspective nous semble plus prometteuse pour mettre en œuvre en EPS l'approche par compétences.

5. Discussion

Dans cette section, nous tenterons d'abord de mettre en relation l'origine de la notion de compétences personnelles et sociales avec le caractère transversal inhérent à cette notion, que nous (re)mettons en question à la suite de Rey (1996). Nous discuterons ensuite des conséquences de l'utilisation de cette notion dans l'enseignement de l'EPS sur le type d'activités et d'objets d'apprentissage proposés aux élèves.

5.1 L'origine de la notion de compétences personnelles et sociales

D'après plusieurs auteurs (Bronckart & Dolz, 2002 ; Dugué, 1994 ; Hirtt, 1996 ; Ropé, 2002, cités par Crahay, 2006), le concept de compétence est externe aux sciences de l'éducation⁷. Il ne proviendrait pas non plus du champ de la psychologie scientifique, mais plutôt du monde de l'entreprise, d'où le manque de scientificité qui lui est reproché (Crahay, 2006). Qu'en est-il de la notion de compétences personnelles et sociales ? Il est plausible que cette dernière constitue une émergence du concept-mère « compétence » et que son statut scientifique souffre pareillement de cette maternité extrascientifique. En revanche, à l'inverse du concept de compétence dont le cheminement semble à présent bien établi depuis sa naissance dans le monde de l'entreprise jusqu'à sa prise en charge par les sciences de l'éducation en passant par sa reprise par l'OCDE et sa diffusion parmi les décideurs des systèmes éducatifs (Crahay, 2006), il est difficile de savoir à quel stade de ce parcours et dans quels lieux a émergé la notion de compétences personnelles et sociales. Toujours est-il que dans la littérature scientifique, on la retrouve aujourd'hui majoritairement dans le champ de la psychologie, des neurosciences voire de la psychiatrie.

Comme nous l'avons déjà relevé, c'est une conception de la compétence sociale *orientée vers les traits* qui a dominé la littérature dans ces champs, en tout cas avant les années 90 (Greenspan & Granfield, 1992). Certaines acceptions de la compétence personnelle – notamment celles qui associent cette notion à une sorte d'auto-efficacité générale mesurable à l'aide de questionnaires – semblent également relever d'une telle conception. En vertu de celle-ci, la congruence entre les habiletés d'un individu et les exigences de la tâche à accomplir n'a pas à être prise en compte, puisque les traits comportementaux à partir desquels est définie la compétence personnelle et/ou sociale sont supposés pouvoir prédire le succès dans les domaines les plus variés. La discussion de Crowe, Beauchamp, Catroppa et Anderson (2011) à l'issue d'une revue systématique des outils d'évaluation de la fonction sociale⁸ pour les enfants et les adolescents est à cet égard éloquent. La majorité des méthodes d'évaluation basées sur des techniques d'observation y sont en effet mésestimées, en raison notamment du fait qu'elles sont spécifiques au contexte dans lequel elles sont employées. Ce caractère situé est pourtant constitutif de la définition du concept de compétence.

5.2 Les conséquences sur l'enseignement de l'EPS

Ainsi que la confrontation des deux outils d'évaluation en EPS l'a bien montré, la diffusion de la notion de compétences personnelles et sociales en EPS a comme principale conséquence la perte de représentativité culturelle des activités et objets d'apprentissage dans cette discipline scolaire. Il faut toutefois convenir que la finalité culturaliste⁹ assignée à l'EPS ne constitue qu'une finalité parmi d'autres (naturaliste¹⁰, citoyenne¹¹) qui peuvent s'avérer contradictoires (Delignières & Garsault, 2004). Ainsi, certains courants de l'EPS (courant psychomoteur, pédagogie des conduites motrices...) ne se préoccupent pas particulièrement de la dimension culturelle des activités et objets d'apprentissage proposés aux élèves. Pourtant, cette dimension est considérée comme essentielle par bon nombre d'auteurs, qui en font une condition sine qua non à une EPS porteuse de sens pour les élèves (Patte, 1997 ; Ubaldi, 2006). Ce risque d'une déculturation de l'EPS

⁷ Nous ne faisons pas ici référence au concept scientifique de compétence tel que développé en linguistique par Chomsky (voir introduction).

⁸ Parmi les tests répertoriés, certains concernent bien la compétence sociale (*interpersonal/social competence*)

⁹ EPS se donnant pour objectif principal l'appropriation par l'élève d'objets culturels sélectionnés.

¹⁰ EPS principalement organisée autour d'objectifs de développement et d'enrichissement des ressources de l'élève (physiques, psychomotrices, cognitives...).

¹¹ EPS se donnant pour objectif principal de former les futurs citoyens, c.à.d. de préparer les élèves à jouer un rôle actif et positif dans la société.

inhérent à la mise en œuvre, dans cette discipline scolaire, d'une conception orientée vers les traits – donc transversale – de la notion de compétences personnelles et sociales, est comparable au phénomène de « décatégorisation » des questions étudiées à l'école observé par Schneider-Gilot (2006) dans d'autres disciplines scolaires suite à la généralisation de l'approche par compétences. D'après cette auteure en effet, « *il semble qu'on mésestime là le fait que l'exercice dans une discipline d'une compétence 'candidate' à la transversalité ne favorise pas forcément son exercice dans une autre discipline tant sont différents et spécifiques les 'espaces didactiques' (au sens de Leutenegger et al., 2003) que constituent les multiples disciplines abordées par un élève lors de sa scolarité, mobilisant des approches, des normes, des codes... particuliers, non réductibles les uns aux autres sauf à sacrifier l'épistémologie propre de ces différentes disciplines et risquer ainsi de favoriser des apprentissages 'creux' chez les élèves au sein de chacune d'elles* » (p. 94). Notre proposition de remplacer la notion de compétences personnelles et sociales, ainsi bien entendu que les pratiques d'enseignement et d'évaluation que celle-ci sous-tend, par les notions de compétences « dans l'action » et « autour de l'action » – même si cette terminologie est critiquable et pourrait faire l'objet de remaniements ultérieurs – participe de cette volonté de réintroduire du disciplinaire et de la vigilance épistémologique à l'école.

6. Conclusion

Dans cette contribution, nous (re)mettons clairement en question la notion de *compétences personnelles et sociales*, tant pour son manque de solidité conceptuelle que pour ses conséquences fâcheuses sur le type d'activités et d'objets d'apprentissage proposés aux élèves en EPS. Nous proposons de lui substituer la notion de *compétences dans l'action et autour de l'action*, plus en accord avec le concept-mère de compétence et plus garante du respect de l'épistémologie propre aux différentes disciplines scolaires, dont la nôtre.

Références

- Ahmed, W., Minnaert, A., van der Werf, G. & Kuyper, H. (2010). The role of competence and value beliefs in students' daily emotional experiences : A multilevel test of a transactional model. *Learning and Individual Differences*, 20, 507-511.
- Bandura, A. (1977). Self-efficacy : Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Beauchamp, M.H. & Anderson, V. (2010). SOCIAL : An integrative framework for the development of social skills. *Psychological Bulletin*, 136(1), 39-64.
- Bronckart, J.-P. (2009). La notion de compétences est-elle pertinente en éducation ? *L'école démocratique*. Consulté le 21.06.2011 sur Internet : <http://www.skolo.org/spip.php?article1124>.
- Crahay, M. (2006). Dangers, incertitude et incomplétude de la logique de compétence. *Revue française de pédagogie*, 154, 97-110.
- Crow, L.M., Beauchamp, M.H., Catroppa, C. & Anderson, V. (2011). Social function assessment tools for children and adolescents : A systematic review from 1988 to 2010. *Clinical Psychology Review*, 31, 767-785.
- Delignières, D. & Garsault, C. (2004). *Libres propos sur l'éducation physique*. Paris : Editions Revue EP.S.
- Dhellemmes, R., Barbot, A., Bonnefoy, G., Coltice, M., Coston, A., Grasset, L., Mascaret, N., Metzler, J., Portes, M., Tribalat, T. & Ubaldi, J.-L. (2007). Pour une culture scolaire des APSA en EPS. *Revue EPS*, 328, 61-67.
- Greenspan, S. & Driscoll, J. (1997). The role of intelligence in a broad model of personal competence. In D.P. Flanagan, J.L. Genshaft & P.L. Harrison (Eds.), *Contemporary intellectual assessment : Theories, tests, and issues* (pp. 131-150). New York, NY : Guilford Press.

- Greenspan, S. & Granfield, J.M. (1992). Reconsidering the construct of mental retardation : Implications of a model of social competence. *American Journal of Mental Retardation*, 96(4), 442-453.
- Griffin, K.W., Scheier, L.M., Botvin, G.J. & Diaz, T. (2001). Protective role of personal competence skills in adolescent substance use : Psychological well-being as a mediating factor. *Psychology of Addictive Behaviors*, 15(3), 194-203.
- Jones, M.I. & Lavalley, B. (2009). Exploring the life skills needs of British adolescent athletes. *Psychology of Sport and Exercise*, 10, 159-167.
- Jourdan, D. (2010). *Éducation à la santé. Quelle formation pour les enseignants ?* Saint-Denis : Inpes.
- Leclercq, D. (1987). L'ordinateur et le défi de l'apprentissage. *Horizons, Philips Professional Systems*, 13, 29-32.
- McGrew, K.S. & Bruininks, R.H. (1990). Defining adaptive and maladaptive behavior within a model of personal competence. *School Psychology Review*, 19(1), 53-73.
- Mottier Lopez, L. & Tessaro, W. (2010). Evaluation des compétences à l'école primaire genevoise : entre prescriptions et pratiques. *Mesure et évaluation en éducation*, 33(3), 29-53.
- Office fédéral du sport suisse (OFSP) (2010). *Compétences personnelles et sociales en éducation physique et sportive*. Macolin : OFSP
- Patte, C. (1997). Comment l'enseignant d'EPS pourrait-il ne pas prendre le sport comme référence culturelle ? *Contrepied*, 1, 48-50.
- Petitpas, A.J., Van Raalte, J.L., Cornelius, A.E. & Presbrey, J. (2004). A life skills development program for High School student-athletes. *The Journal of Primary Prevention*, 24(3), 325-334.
- Rey, B. (1996). *Les compétences transversales en question*. Paris : ESF.
- Schneider-Gilot, M. (2006). Quand le courant pédagogique « des compétences » empêche une structuration des enseignements autour de l'étude et de la classification de questions parentes. *Revue française de pédagogie*, 154, 85-96.
- Ubaldi, J.-L. (Ed.) (2006). *L'EPS dans les classes difficiles*. Paris : Editions Revue EP.S.
- Wallston, K.A. (1992). Hocus-Pocus, the focus isn't strictly on locus : Rotter's social learning theory modified for health. *Cognitive Therapy and Research*, 16(2), 183-199.

EVALUATION DE LA COMPÉTENCE À COMMUNIQUER À L'ÉCRIT ET À L'ORAL EN FORMATION INITIALE DES MAÎTRES : ÉTAT DE LA SITUATION AU QUÉBEC

Christian Dumais* et Nancy Goyette**

* Doctorant à l'Université du Québec en Outaouais, chargé de cours à l'Université du Québec à Montréal et à l'Université de Montréal, dumais.christian@uqam.ca

** Doctorante et chargée de cours à l'Université du Québec à Trois-Rivières, nancy.goyette@uqtr.ca

Mots-clés : compétence, écrit, oral, formation des maîtres, Québec

Résumé. Depuis 2001, le Québec s'est doté d'un référentiel composé de douze compétences professionnelles que doit développer tout étudiant évoluant dans un baccalauréat en enseignement. Parmi ces compétences à développer, l'une concerne la communication écrite et orale dans divers contextes liés à l'enseignement. Cette compétence doit faire l'objet d'une évaluation. Pour obtenir son brevet d'enseignement, l'étudiant doit faire la preuve de la maîtrise de cette compétence. Pendant plusieurs années, les universités ont pu choisir la façon de l'évaluer par divers outils. Toutefois, afin d'harmoniser et d'améliorer les pratiques d'enseignement et d'évaluation du français, et pour valoriser la qualité de la langue écrite, une entente interuniversitaire adopta la création d'une épreuve unique de certification en français écrit, le TECFÉE. À l'oral, aucune mesure particulière n'a été prise par l'ensemble des universités québécoises pour permettre l'amélioration de cette compétence. L'implantation du TECFÉE ne s'est pas fait sans heurt et le taux d'échec élevé a engendré la mise en place de différentes mesures d'aide à l'écrit et aussi à l'oral dans quelques universités. Malgré cela, nous nous interrogeons à propos de l'efficacité de ces mesures.

1. Introduction

Depuis le début des années 2000, le Québec s'est doté d'un référentiel composé de douze compétences professionnelles que doit développer tout étudiant évoluant dans un baccalauréat en enseignement (Ministère de l'Éducation du Québec, 2001). Pour chacune des compétences, un niveau de maîtrise est attendu afin que l'étudiant puisse obtenir le droit d'enseigner. Parmi ces compétences à développer, l'une concerne la communication écrite et orale dans divers contextes liés à l'enseignement. Cette compétence doit faire l'objet d'une évaluation.

2. Problématique

La problématique privilégiée lors de notre communication était celle de l'articulation entre le référentiel de compétences, les pratiques d'enseignement/formation des étudiants en enseignement et les pratiques d'évaluation. En effet, en 2001, les États généraux sur la situation et l'avenir de la langue française au Québec dénonçaient le fait que la moitié des futurs enseignants québécois ne possédait pas une maîtrise de la langue française suffisante pour exercer leur fonction (Gouvernement du Québec, 2001a). Devant ce constat, une collaboration entre le ministère de l'Éducation du Québec et les universités permit la mise en place de normes particulières en matière de langue. Dès 2003, une réflexion fut amorcée par l'Association des doyens et directeurs

pour l'enseignement et la recherche en éducation du Québec au sujet d'une politique et de modalités communes entre les universités pour harmoniser les pratiques d'enseignement et d'évaluation du français et, aussi, valoriser la qualité de la langue dans la formation. Cette réflexion mena, en 2008, à une entente interuniversitaire rendant possible l'utilisation d'une épreuve unique, dans toutes les universités québécoises, qui évalue la compétence langagière écrite en français de tous les futurs enseignants. Ce test, le TECFÉE (test de certification en français écrit pour l'enseignement), s'il est réussi, ouvre la porte à l'entrée dans la profession, en permettant l'émission du brevet officiel d'enseignement si l'étudiant réussit tous les cours de son programme de formation en enseignement. Toutefois, la mise en place de ce nouveau dispositif d'évaluation ne se fait pas sans heurt, puisque les modalités règlementaires relatives à ce test contraignent l'étudiant à le réussir dans les deux premières années du baccalauréat, sans quoi il ne peut s'inscrire à son troisième stage (à la troisième année de formation). Ajoutons que le contenu surchargé des programmes de formation ne permettent pas ou très peu d'intégrer des cours de français afin d'aider les étudiants à développer cette compétence au cours de leur cheminement. Enfin, après trois ans de mise en œuvre, les résultats mitigés de ce test et le taux d'échec à la première passation (jusqu'à 60 % dans certaines universités) démontrent clairement que les étudiants ont besoin de développer leur compétence langagière écrite et que les apprentissages effectués en vue du TECFÉE ne semblent pas toujours efficaces et durables.

Pour ce qui est de l'évaluation de la compétence langagière orale des étudiants, aucune disposition officielle n'a encore été prise contrairement à l'évaluation de l'écrit. Dans certaines universités, seuls les stages en classe permettent d'évaluer cette compétence, ce qui est nettement insuffisant. Dans d'autres universités, en plus des stages, une seule évaluation est effectuée. Cette dernière est faite en début de formation et consiste seulement en une présentation de cinq minutes. Certaines universités offrent des cours de mise à niveau et des cours spécifiques dans le cursus scolaire ou par l'intermédiaire d'un centre d'aide en français qui relève d'une base volontaire des étudiants, alors que d'autres universités n'offrent aucun cours spécifique de la langue orale. L'évaluation de la compétence langagière orale en formation des maîtres varie d'une institution à une autre et ne semble pas toujours être une priorité. Pourtant, elle nécessite d'être balisée et une place particulière doit être accordée à l'enseignement de l'oral dans la formation des maîtres afin que les étudiants puissent développer cette compétence (Plessis-Bélair, 2006 ; Fisher, 2007).

3. Des questions à répondre

Cette communication a porté un regard sur la situation vécue dans les universités québécoises en ce qui concerne l'articulation entre les pratiques d'enseignement et de formation d'une part, et, d'autre part, l'évaluation des compétences. Nous nous sommes interrogés à propos de l'efficacité d'une épreuve unique d'évaluation de l'écrit dans toutes les universités québécoises et à propos de la place restreinte accordée à l'oral et à son évaluation. Plusieurs questions ont été posées. Le contexte actuel relié aux évaluations certificatives des futurs enseignants n'est-il pas dénaturé au sens où l'on évaluerait plus les résultats de « bachotages » que le développement réel des compétences ? Peut-on croire que les étudiants ayant réussi le test TECFÉE sont compétents dans la maîtrise de la langue française ? Comment développer un outil représentatif du développement réel de cette compétence de communication, et ce, à l'écrit comme à l'oral ? Quelles sont les mesures possibles à mettre en place dans les universités pour assurer une réelle formation à l'oral ainsi qu'une évaluation adéquate de l'oral des étudiants ?

4. Des pistes de solution

Dans le but de favoriser le développement de cette compétence chez les étudiants et d'assurer une évaluation plus adéquate, certaines universités ont emboîté le pas pour créer des mesures d'aide. À l'écrit, on compte entre autres la création de centres de soutien à la formation des maîtres, des tests diagnostiques en français spécialement conçu pour les étudiants en enseignement et des ajustements au cheminement scolaire pour les étudiants ne répondant pas aux exigences. À l'oral, des cours et ateliers spécifiques ont été créés et diverses interventions avec les acteurs concernés

ont été réalisées pour permettre aux élèves de développer leur langue orale. Ces différentes mesures ont été analysées et discutées lors de la communication et plusieurs questions ont été soulevées. Malgré ces mesures d'aide innovantes mises en place par les universités pour améliorer le développement de cette compétence et son évaluation, les mesures sont-elles efficaces et aident-elles réellement les étudiants ? De plus, la mise en œuvre d'un test unique de certification de l'écrit nous questionne à plusieurs égards, notamment par sa validité et le paradoxe qu'il engendre face à l'esprit du Programme de formation de l'école québécoise. En effet, l'évaluation des apprentissages est axée sur le développement de compétences dans une vision socioconstructiviste qui « se caractérise par les objets sur lesquels elle porte, par le rapport qu'elle entretient avec le processus d'apprentissage, par les manières d'interpréter les résultats de l'évaluation sommative, par les méthodes employées et par les valeurs qui la sous-tendent » (Gouvernement du Québec, 2001b, p. 91). Pour ce qui est de l'aspect oral de cette compétence, comment pouvons-nous nous assurer de son développement et de la validité de son évaluation étant donné le peu de place accordé à son enseignement et à son évaluation ? Autant en ce qui concerne l'aspect écrit que l'aspect oral, comment réussir à développer une compétence durable ?

Références

- Fisher, C. (2007). Le développement des compétences langagières à l'oral dans le contexte de la formation à l'enseignement. Dans É. Falardeau, C. Fisher, C. Simard et N. Sorin (Dir.), *Didactique du français : Les voies actuelles de la recherche*. Québec : Presses de l'Université Laval.
- Gouvernement du Québec (2001a). *Commissions des états généraux sur la situation et l'avenir de la langue française au Québec : Le français, une langue pour tout le monde - Une nouvelle approche stratégique et citoyenne* (Rapport de la commission présidée par Gérard Larose). Récupéré le 7 juillet 2011 de <http://www.spl.gouv.qc.ca/documentation/rapportssondagesstatistiques/>
- Gouvernement du Québec (2001b). *La formation à l'enseignement. Les orientations. Les compétences professionnelles*. Ministère de l'Éducation. 210 p.
- Ministère de l'Éducation, du Loisir et du Sport (2009). *Politique et modalités relatives au test de certification en français écrit pour l'enseignement (TECFÉE)*. Récupéré le 2 juillet 2011 de http://www.cpe.uqam.ca/PDF/PolitiqueModalites_relatives_TECFEE%20ADEREQ.pdf
- Ministère de l'Éducation du Québec (2001). *La formation à l'enseignement : Les orientations, les compétences professionnelles*. Gouvernement du Québec.
- Plessis-Bélair, G. (2006). La didactique de l'oral dans le programme actuel de formation : Au cœur de l'enseignement et de l'apprentissage. Dans J. Loiselle, L. Lafortune et N. Rousseau (Dir.), *L'innovation en formation à l'enseignement : Pistes de réflexion et d'action*. Québec : Presses de l'Université du Québec.

Evaluer les compétences à l'ère du numérique :

L'E-portfolio comme outil

Résumé

Le présent article se penche sur la question liée au développement d'outils de présentation, d'analyse et d'évaluation des compétences en milieu universitaire au profit d'un public en formation continue. Il s'agit de montrer tout l'intérêt pédagogique et professionnel de l'adoption de l'e-portfolio comme outil tirant particulièrement profit des possibilités offertes par les technologies d'information et de communication, notamment celles du Web2.0. En effet, se référant aux nouvelles approches pédagogiques centrées sur les apprenants, la mise en œuvre d'outils d'aide à l'apprentissage, d'autoévaluation, de régulation et de métacognition, s'avère d'une grande importance pour soutenir les parcours de formation d'adultes, surtout que l'expérience confirme leur caractère à la fois fiable et équitable.

Conçu de façon modulaire, l'e-portfolio proposé à Paris 3, Sorbonne Nouvelle, dans le cadre du Master2 professionnel AIGEME (Mention ingénierie de l'e-learning), permet d'intégrer l'ensemble des activités réalisées dans le cadre des cours et à travers l'expérience du stage. D'une part, cet outil favorise chez l'apprenant la créativité, la réflexion, le partage des réalisations et la rétrospection. D'autre part, il s'agit de faciliter à l'enseignant la tâche de suivi, d'encadrement, de régulation, d'évaluation et de certification des compétences acquises. Dans cet article, nous présenterons d'abord le contexte de l'introduction de l'e-portfolio ainsi que le scénario de son utilisation en vue de soutenir l'apprentissage et évaluer les compétences. Nous ne manquerons pas de passer en revue quelques définitions de l'outil et d'étudier son impact positif sur le processus d'apprentissages et son intérêt pour l'évaluation des compétences. Par la suite, une réflexion portant sur ses forces et ses limites permettra de proposer quelques recommandations d'usage.

Mots-clés : E-portfolio, pratique réflexive, développement professionnel

Introduction

Certes, l'avènement des TIC a doublement impacté le développement des compétences et leur mode d'évaluation. En effet, la technologie se met de plus en plus au service de l'évaluation des compétences tant dans le milieu scolaire que professionnel. Toutefois, pour vivre en intelligence avec ce nouvel environnement à la fois diversifié, numérique et mondial, la seule maîtrise des compétences de spécialité n'est plus suffisante, il faut également s'approprier certaines compétences transversales dites parfois « clés » impliquant la mobilisation de savoir-faire cognitifs et pratiques, ainsi que des qualités psychosociales telle que la motivation, la créativité et l'autonomie.

Ainsi, l'évaluation des compétences suppose désormais l'adoption et le développement de nouveaux dispositifs de mesure susceptibles de relever les défis actuels, à la fois technologiques et organisationnels. Dans ce sens, parmi les outils technologiques existants, l'e-portfolio permet bien de décrire et d'illustrer la carrière d'une personne, son expérience ainsi que ses réussites, à travers une sorte de réflexion/réflexivité permanente et un regard critique sur les acquis et réalisations. Le choix de l'e-portfolio se justifie simplement par le fait qu'il représente à la fois un outil d'apprentissage continu, de communication, de développement professionnel et d'évaluation.

1. Description du dispositif pédagogique dans lequel s'insère l'outil E-portfolio

Le contexte pédagogique d'implantation du e-portfolio, dans le cadre du Master AIGEME à Paris 3 Sorbonne Nouvelle, se caractérise par un public majoritairement engagé dans le monde professionnel, mais qui cherche à parfaire ses connaissances dans le domaine de l'ingénierie du e-learning.

1.1. Contexte de la formation AIGEME

1.1.1. Objectifs de la formation

Le Master AIGEME *Ingénierie de la formation à distance en sciences humaines* est conçu comme une démarche professionnalisante appuyée sur la recherche. Il vise à valoriser, par un enseignement informatique pédagogiquement adapté, les compétences d'étudiants non-

scientifiques pour les préparer à des emplois-métiers liés au multimédia, à la FOAD en sciences humaines.

1.1.2. Contenu

La formation s'est organisée autour de trois principaux axes :

- ▶ Approche théorique et méthodologique,
- ▶ Formation technique aux langages et aux logiciels,
- ▶ Réalisation de projets de type professionnel (mémoire et stage) et d'un portfolio,

1.1.3. Modalités

Fondé sur un partenariat université-entreprise, le diplôme *Applications Informatiques, Gestion, Etudes, Multimédia, E-formation (AIGEME)* s'effectue essentiellement à distance via une plate-forme de formation à distance (Blackboard/WebCT)¹.

Des ressources multimédia (cédéroms, vidéos, DVD) et la participation à des visioconférences sont également proposées selon les enseignements.

1.1.3. Compétences visées

- Maîtriser les principaux contextes pédagogiques et techniques de l'eFormation dans une démarche d'autonomisation : les classes virtuelles par visioconférence et Internet, les campus universitaires ou universités virtuelles d'entreprise.
- Utiliser Internet pour actualiser et organiser des contenus pédagogiques et accéder à des ressources extérieures pertinentes, des contenus diffusés, les outils de collaboration et d'interaction sur Internet, le partage d'applications à distance, les outils collaboratifs, des serious games...
- Connaître les techniques du système d'information, les réseaux, la structuration de l'information, la visioconférence, la création sur supports diversifiés (plates-formes et solutions mobiles) et les principes d'ergonomie.
- Comprendre les enjeux stratégiques, organisationnels, politiques, juridiques et éthiques de la eFormation
- Être capable de développer un projet multimédia en équipes multiculturelles et multilingues
- Produire des énoncés de qualité professionnelle en français et en anglais dans le domaine du Elearning.

1.2. Les modalités de contrôle des connaissances

¹ En cours de migration vers Moodle

Dans ce contexte de formation continue, l'évaluation des apprenants se fait le long de la formation via le contrôle continu notamment à travers la résolution de situations problèmes, des réflexions, des recherches, etc. Toutefois, Les responsables de la formation ont mis en œuvre un scénario du e-portfolio intégré dans les modalités d'encadrement et de rétroaction, de manière à ce qu'il contribue largement au développement des connaissances et compétences visées par la formation, sans oublier pour autant les autres compétences complémentaires telles que apprendre à apprendre.

2. De la problématique de l'évaluation des compétences

D'après le cadre conceptuel du programme DeSeCo, les compétences clés regroupent principalement trois catégories dont celles de se servir, de manière interactive, d'outils tels que les langues et les technologies d'information et de communication, d'interagir dans des groupes hétérogènes et multiculturels et enfin d'agir en autonomie. Or, la capacité d'un individu à réfléchir et à agir de façon réflexive est au cœur de ce cadre de compétences : *« cette pratique réflexive n'implique pas seulement d'appliquer une formule ou une méthode de manière routinière dans une situation donnée, mais aussi de faire face au changement, de tirer des enseignements des expériences et de réfléchir et d'agir avec esprit critique. »*²

Rappelons qu'à l'origine, l'évaluation assure trois fonctions clés à savoir, sociale relative à l'organisation d'examens, de concours ainsi qu'à l'insertion sociale et professionnelle des jeunes, institutionnelle, universitaire et pédagogique.

Au-delà de la dimension sociale et institutionnelle de la compétence, la prise en compte du « Sentiment d'Efficacité Personnelle » (SEP)³ semble essentielle vu que la notion est d'abord relative à l'image de soi et à la motivation intrinsèque. Dans ce sens, P.-Henri François associe compétence et motivation : *« Le sentiment de compétence est le résultat de l'auto-évaluation par l'individu de ses capacités à mettre en œuvre des comportements ou des ensembles de comportements identifiés comme efficaces pour obtenir ou éviter certains résultats. Il joue un rôle important dans les processus de motivation, dans la sélection et la régulation des conduites. »*. [P. Henri François 97].

² LA DÉFINITION ET LA SÉLECTION DES COMPÉTENCES CLÉS, consultable sur : <http://www.oecd.org/dataoecd/36/55/35693273.pdf> (Inspiré du *Key competencies for a successful life and a well-functioning society*, (2003), D.S. Rychen et L.H. Salganik (éd).

³ P. Henri François 97

3. L'introduction d'un outil évaluatif: l'e-portfolio

3.1. Qu'est-ce qu'un portfolio?⁴

Il existe une multitude de définitions du terme «portfolio», de la plus usuelle : «*Collection d'œuvres propre à refléter le talent de son auteur*» à la plus éducative : «*Collection structurée des travaux d'un étudiant et des commentaires qui leur sont attachés, qui fait foi de ses compétences montrant des traces pertinentes de ses réalisations. Les travaux conservés, et parfois affichés, sont sélectionnés en fonction de critères établis par l'étudiant et par l'enseignant*».

La littérature récente propose plusieurs types de portfolios allant de la simple présentation de travaux ou de réalisations jusqu'à la recherche. Nous en retiendrons la taxonomie de l'ASCD⁵, qui se limite à trois catégories de portfolios : d'apprentissage, de présentation et d'évaluation. Dans une autre définition, un même portfolio peut réunir quatre fonctions (Weiss, 2000) :

- d'information – le portfolio de présentation
- de formation – le portfolio d'apprentissage/
- de développement professionnel
- d'attestation – le portfolio d'évaluation

3.2. Du simple portfolio au portfolio numérique

Le passage du simple Portfolio au E-portfolio représente plusieurs avantages notamment ceux d'avoir accès au portfolio en tout lieu et à tout moment, de classer et de réorganiser facilement et rapidement des documents, d'assurer la sécurité et la confidentialité du contenu, particulièrement des renseignements personnels et de créer des hyperliens d'un document à un autre. Plusieurs qualificatifs peuvent définir un e-portfolio. Nous en retiendrons celles qui nous semblent rappeler parfaitement la finalité de l'outil, amené à être expérientiel, réflexif, identitaire, d'apprentissage, évolutif, valorisant et professionnel.

Il s'agit d'un dossier personnel « *documenté et systématique. Il est constitué par la personne pour reconnaître ses acquis ou pour les faire reconnaître sur le plan institutionnel ou professionnel* »⁶.

⁴ Voir le document « Le portfolio sur support numérique » de Robert Bibeau sur <http://ntic.org/guider/textes/portfolio.doc>

⁵ Association for Supervision and Curriculum Development

⁶ Voir le site : www.educnet.education.fr

Pour souligner l'intérêt du e-portfolio, nous nous référerons à la citation suivante de Serge Ravet (Directeur de l'association Eifel et responsable du module sur le ePortfolio dans le Master 2 AIGEME) : « *On utilise la technique de l'apprentissage réflexif et attestatif. On réfléchit sur ce qu'on a fait, on restitue la trame de ses expériences, on en extrait ses compétences, et on rassemble toutes les preuves de ses qualités sociales et professionnelles, acquises pendant sa formation, des stages ou diverses activités, mêmes associatives.* »⁷

3.3. Scénarisation du Eportfolio et perception des utilisateurs

Comme signalé plus haut, le e-Portfolio multilingue crée et alimenté régulièrement par chaque apprenant, doit permettre au moment de la soutenance, une évaluation globale de son parcours, de ses réalisations tant au niveau du mémoire que du stage professionnel. Pratiquement tourné vers l'apprentissage tout au long de la vie et s'appuyant sur les technologies de l'information et de la communication, ce travail réflexif personnel présente les compétences acquises dans le domaine de la e-formation et mises en application en contexte professionnel lors du stage.

Le e-portfolio, tel qu'il est programmé dans la formation AIGEME, doit servir de support pour chaque apprenant afin de décrire à la fois le processus d'apprentissage (découverte, analyse, assimilation, réflexion) et le fruit de ce processus (collection des compétences acquises).

C'est cette double dimension qui fait que cet outil pédagogique se transforme en un véritable miroir favorisant la rétrospection, la remise en question et la régulation ; d'où le développement d'un apprentissage réflexif consistant pour chaque apprenant à marquer un temps dans son apprentissage où il observe le chemin parcouru, confronte alors les connaissances et compétences nouvellement acquises, celles maîtrisées avec ses objectifs de formation. Cela doit lui permettre d'améliorer son apprentissage, de réorienter son auto-formation, son organisation, lui donner des pistes de recherche, le pousser à revenir sur certains points mal assimilés.

L'accent est donc mis sur trois dimensions :

- a. le **produit** : L'apprenant s'appuie sur les outils du web 2.0 pour publier en ligne le contenu. Il choisit notamment des solutions de blogging du type wordpress.

⁷ Voir un exemple de production d'étudiante : <http://dinaelkassas.wordpress.com/>

- b. le **processus** : Il est évolutif puisque l'élaboration de l'e-portfolio s'étale sur toute la durée de la formation et permet à l'apprenant de publier au fur et à mesure ses travaux et réalisations, de les annoter et de les commenter.
- c. le **progrès** : L'apprenant est sollicité dès le démarrage de la formation pour alimenter régulièrement le portfolio afin de favoriser le progrès, d'où le caractère évolutif de l'outil. En effet, le travail mené sur son e-portfolio doit mener l'apprenant, en fin de formation, à la présentation d'un dossier complet qui fera l'objet de sa soutenance devant un jury. Parallèlement au travail de l'apprenant, le feed-back de l'enseignant-encadrant est fourni en vue de permettre à l'apprenant de réguler son travail et de développer les capacités d'auto-évaluation.

A la consultation des dizaines d'e-portfolios élaborés par des étudiants ayant suivi la formation AIGEME, deux notions sont souvent abordées, à savoir l'apprentissage réflexif et la métacognition. Dans les paragraphes qui suivront, nous tenterons d'expliquer en quoi et pourquoi elles doivent être associées au portfolio.

D'une part, la pratique réflexive consiste à porter une sorte de regard critique sur son propre fonctionnement et ajuster ses schèmes d'apprentissage afin d'apprendre plus efficacement. R.BIbeau en souligne l'intérêt : *"Pour jeter un regard critique sur ses processus d'apprentissage et son mode de développement des compétences il faut conserver des traces de ses travaux, de sa démarche, de ses processus, de ses méthodes, de ses stratégies et documenter ces traces...expliquer et écrire pourquoi on sélectionne tel document que l'on dépose dans son portfolio. Comment on règle tel problème ? C'est l'objectif du portfolio d'apprentissage."*⁸

D'autre part, la métacognition, selon la définition de Flavell (1976), considéré comme le pionnier dans ce domaine, «se rapporte à la connaissance qu'on a de ses propres processus cognitifs, de leurs produits et de tout ce qui touche, par exemple, les propriétés pertinentes pour l'apprentissage d'informations et de données... La métacognition se rapporte entre autre chose, à l'évaluation active, à la régulation et l'organisation de ces processus en fonction des objets cognitifs ou des données sur lesquelles ils portent, habituellement pour servir un but ou un objectif concret. » Ce qui renvoie à la connaissance qu'un apprenant peut avoir sur ses stratégies cognitives.

⁸ http://www.fse.ulaval.ca/Gerard.Scallon/valise_BEP/portfolioguide.pdf

Conclusion : Limites et perspectives

Au vu de ce qui précède, les limites de l'e-portfolio sont minimales par rapport à ses nombreux atouts. Il faudrait peut-être souligner le premier frein lié à la connectivité et à l'accès aux équipements informatiques. De même que certains apprenants ont besoin d'être particulièrement encadrés afin que leurs réflexions soient pertinentes et plus profondes. Par ailleurs, il faudrait penser à alléger les contenus des activités programmées surtout celles qui exigent de la rédaction.

Malgré ces inconvénients, l'e-portfolio permet d'assurer un suivi régulier des apprentissages de l'apprenant et d'en mesurer la qualité. Ce dernier prend en charge sa propre formation, mène une réflexion continue sur les succès et difficultés rencontrées, passe de la théorie à la pratique, élabore la synthèse de ses apprentissages et se motive davantage face à ses performances. Cette autonomie et cette responsabilisation de l'étudiant vis-à-vis de son propre processus d'apprentissage font enfin que « *Le portfolio s'inscrit dans un courant d'évaluation qualifiée d'«authentique», puisque cette forme d'évaluation exige d'illustrer les connaissances ou les habiletés dans des situations réelles d'apprentissage (Meyer, 1992).* »⁹

Références bibliographiques :

- Bandura, A. (1995). Self-efficacy in changing societies. New York, Cambridge University Press. p.11.
- De ketele, J.-M. (2000). En guise de synthèse: Convergences autour des compétences, in Bosman, C., Gerard, F.-M, Roegiers, X. (Éds). *Quel avenir pour les compétences?* Bruxelles: De Boeck.
- Donnay, J. et Charlier, É. (2006). Apprendre par l'analyse de pratiques: initiation au compagnonnage réflexif. Namur: Presses universitaires de Namur, Sherbrooke: Éditions CRP.
- Gerard, Fs-M. *Évaluer des compétences : guide pratique*. De Boeck, 2008
- Kolb, D.(1984). *Experiential Learning*.Englewood Cliffs, N.J.: Prentice-Hall.
- Michaud, G., Dionne, P., Beaulieu, G.. *Le bilan de compétence : regards croisés entre la théorie et la pratique*. Septembre éd., 2007
- Pacteau, Chantal. L'évaluation aujourd'hui : comment mesurer les compétences ?. *Sciences humaines*, décembre 2001, Hors série n° 35, p. 10-13.
- Perrenoud, Ph. (1997). *Construire des compétences dès l'école*. Paris: ESF.

⁹ Le portfolio: vers une pratique réflexive de l'enseignement Georgette Goupil professeure au département de psychologie Université du Québec à Montréal Extraits de Vie Pédagogique, 107, 1998

- Paulson, F. L., Paulson, P. R., & Meyer, C. A. (1991). What makes a portfolio a portfolio ? *Educational Leadership*, 48(5), 60-63. Consultable en ligne:
<http://faculty.milken school.org/sperloff/edtech/portfolioarticle2.pdf>
- Berthiaume, D., Daele, A. (2010). Faire appel au portfolio pour évaluer les apprentissages des étudiant-e-s ? Consultable en ligne:
http://www.unil.ch/webdav/site/cse/shared/brochures/memento_portfolio.pdf

**LA DEMARCHE EPORTFOLIO DANS L'ENSEIGNEMENT SUPERIEUR FRAN  AIS :
PREMIERS RETOURS D'EXPERIENCES ET PROPOSITION DE CLARIFICATION**

Jean Heutte*, Annie J  z  gou**

** Universit   d'Artois et CREF (EA 1589) de l'Universit   Paris Ouest Nanterre – La D  fense
jean.heutte@univ-artois.fr*

***   cole des Mines de Nantes et CREAD (EA 3875) de l'Universit   de Rennes 2
annie.jezegou@mines-nantes.fr*

Mots-cl  s : *accompagnement, agentivit  , comp  tences, ePortfolio, motivation*

R  sum  . *La d  marche ePortfolio participe    la construction identitaire de l'  tudiant, en tant qu'apprenant et futur professionnel. Elle lui permet notamment d'identifier, d'explicit  r, de valoriser et de capitaliser ses comp  tences, cela tout au long de son parcours universitaire mais aussi de sa vie professionnelle ult  rieure, tout en soutenant ses efforts d'insertion dans le monde du travail. La pr  sente communication contribue    apporter quelques premiers   l  ments de clarification    une telle d  marche. Cette clarification r  sulte des investigations d'un groupe de travail national, institu      l'initiative du Minist  re fran  ais de l'Enseignement Sup  rieur et de la Recherche. Ici, les auteurs pr  sentent, du moins en partie, un   tat de l'avancement du travail r  alis   par ce collectif.*

La d  marche ePortfolio constitue un enjeu important pour le syst  me universitaire. Elle renvoie    la fois aux probl  matiques de l'  ducation et    celle de la formation « tout au long de la vie », au travers de la formalisation des parcours et des r  f  rentiels, de nouvelles modalit  s formelles et informelles d'acquisition ou encore de mise en exergue de comp  tences construites. En m  me temps, elle s'inscrit dans une logique de reconnaissance et de valorisation des acquis acad  miques et exp  rientiels, notamment dans la perspective de l'insertion professionnelle. Elle se mat  rialise par un objet num  rique : le ePortfolio.

Cet objet est la partie visible de trois d  marches pouvant, selon les cas, exister ind  pendamment les unes des autres ou s'articuler entre-elles. La premi  re renvoie    la progression r  alis  e par l'  tudiant dans ses apprentissages (ePortfolio d'apprentissage). Par cons  quent, elle s'envisage sous l'angle du « processus », tandis que les deux autres portent sur le « produit » issu d'un tel processus. Ainsi, la deuxi  me d  marche porte sur l'  valuation formative des acquis, en termes de comp  tences, r  sultant de ses apprentissages (ePortfolio d'  valuation) tandis que la troisi  me est centr  e sur une explicitation d  taill  e, de la part de l'  tudiant, de ses exp  riences et r  alisations, de ses capacit  s et comp  tences (ePortfolio de pr  sentation). Ces d  marches ont en commun de s'appuyer sur une analyse r  flexive de la part de l'  tudiant ; cette analyse pouvant, si besoin,   tre accompagn  e par des acteurs   ducatifs.

L'impulsion donn  e    la d  marche globale ePortfolio par les pouvoirs publics, dont le Minist  re de l'enseignement sup  rieur et de la recherche (MESR), n'est pas une exception fran  aise. Ainsi, dans la plupart des pays europ  ens ainsi que ceux de l'ensemble de l'Organisation de coop  ration et de d  veloppement   conomique (OCDE), les organismes gouvernementaux, en charge de la mise en   uvre des politiques publiques dans l'enseignement sup  rieur, organisent et structurent, au niveau national, la visibilit   et la promotion de projets et d'initiatives concernant la d  marche ePortfolio.

Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

En France, de nouveaux textes règlementaires font référence de façon plus ou moins directe à la démarche ePortfolio. Pour exemple : concernant la certification Informatique et Internet (C2i)¹, les principes et les modalités de certification reposent sur la constitution, par le candidat, d'un dossier numérique de compétences. Dans ce dossier, il doit faire état des savoirs qu'il a acquis ainsi que de la manière dont il a mobilisé et combiné ces acquis en situation, cela au regard des compétences requises pour l'obtention du certificat. D'une manière plus spécifique, « *ce dossier, constitué par le candidat, rassemble des éléments apportant la preuve des savoirs acquis, des aptitudes développées et des compétences maîtrisées en regard d'un référentiel C2i®. Ces éléments peuvent être des productions résultant des activités proposées au candidat et intégrées, autant que faire se peut, dans son cursus ; les résultats de contrôle de connaissances ; des productions externes commentées résultant d'activités du candidat en dehors de son cursus.* » (circulaire du 9 juin 2011).

Autre exemple : l'article 2 de l'arrêté du 1er août 2011, relatif à la Licence, précise que ce diplôme atteste l'acquisition d'un socle de connaissances et de compétences dans un champ disciplinaire ou pluridisciplinaires. Il annonce, dans son article 3, la production de référentiels de compétences définis pour une discipline ou un ensemble de disciplines. Actuellement, des référentiels sont en cours d'élaboration et seront prochainement publiés. *In fine*, dans l'article 13 de cet arrêté, il est précisé : « *Un processus dématérialisé de suivi des crédits acquis par chaque étudiant est mis en place* ».

Ces mouvements devraient converger et participer à la création d'un support dématérialisé, de type ePortfolio, qui capitalise et valorise l'ensemble des acquis d'apprentissage et d'expérience de l'étudiant. Dès son entrée dans un établissement de formation, et dans une mise en perspective accompagnée du projet de formation et professionnel qu'il construit, l'étudiant devrait pouvoir s'appuyer sur une démarche à l'analyse réflexive tout en disposant, si besoin, d'un accompagnement de la part d'acteurs éducatifs ainsi que d'un outil numérique pour mettre en lisibilité ces acquis.

Moyennant certaines conditions de transparence et d'interopérabilité, un tel dispositif (accompagnement et/ou outil) pourra garantir la mobilité de l'étudiant et l'inscription dans la durée de la démarche. Dans cette perspective, il est crucial d'éviter un certain nombre de risques inhérents aux initiatives individuelles et qui peuvent aboutir à des impasses pour les étudiants et à un coût injustifié pour les établissements en termes de moyens humains et financiers.

Afin d'apporter des réponses à nombre de questions liées à la démarche ePortfolio, un groupe de travail national (GTN) "ePortfolio" a été mis en place par la direction générale de l'enseignement supérieur (DGESIP). Son but est de produire des éléments de clarification et des recommandations à l'attention des établissements d'enseignement supérieur français. Ce GTN est supporté par la mission numérique pour l'enseignement supérieur (MINES), au sein du service de la stratégie pour l'enseignement supérieur et l'insertion professionnelle. Il a pour objectifs spécifiques de :

- Dresser un état des lieux de la mise en œuvre de la démarche ePortfolio dans les établissements d'enseignement supérieur ;
- sensibiliser et informer la gouvernance des établissements, les services d'insertion professionnelle et les enseignants sur la démarche ePortfolio, tout en les outillant pour sa mise en œuvre notamment pour son pilotage politique et pédagogique ;
- proposer un cahier des charges fonctionnel afin d'aider à la mise en place des dispositifs de gestion de ePortfolios (SGeP) intégrés dans les systèmes d'information (SI) des établissements d'enseignement supérieur. Ces dispositifs doivent donner la possibilité aux étudiants de disposer de ePortfolios sécurisés, pérennes, personnalisables, tout en leur permettant de présenter les données à partir de points de vue différents et des cibles visées au cours de leur formation initiale, en vue de leur insertion professionnelle comme de leur parcours professionnel tout au long de la vie.

¹ Circulaire du 9 juin 2011 ou arrêté du 14 décembre 2010 relatif au C2i niveau 2 "enseignant"

Actes du 24^e colloque de l'Admée-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

Ces éléments seront rassemblés dans un document de clarification et de recommandations de type "livre blanc" à paraître d'ici le printemps 2012. Appelé à être révisé périodiquement en fonction de nouveaux besoins ou d'avancées technologiques majeures, ce document permettra également d'orienter d'éventuels développements :

- De standards (notamment processus de traduction d'échanges et de passerelles) ;
- d'applications informatiques (notamment de systèmes de gestion d'ePortfolios) ;
- de connecteurs entre différents systèmes d'information (système d'Information d'organismes de formation et/ou Système d'Information d'employeurs...).

1. Premiers retours d'expériences

La synthèse des premières recommandations d'un grand nombre d'acteurs, actuellement impliqués dans des projets ePortfolio, fait apparaître trois préoccupations majeures :

La première préoccupation consiste à se centrer principalement sur la démarche ePortfolio plutôt que sur l'outil ePortfolio en tant que tel, tout en repérant les conditions institutionnelles et éducatives permettant de soutenir cette démarche. De telles conditions renvoient notamment à l'individualisation des parcours s'appuyant sur l'élaboration préalable de référentiels de compétences, à la mise en place de portefeuilles destinés à mettre en évidence et à valoriser les acquis de l'apprentissage et ceux issus de l'expérience des étudiants, tout accompagnant si besoin ces derniers dans la démarche ePortfolio et en les plaçant dans la perspective d'un projet personnel et professionnel.

Cette démarche implique de concilier des logiques parfois contradictoires tels que, entre autres, la prise en compte des programmes (nationaux, académiques, universitaires), d'éléments constitutifs des référentiels en vigueur (CV EuroPass, Portefeuille européen des langues (PEL, les C2i...), cela dans un contexte de massification (ou globalisation) et d'individualisation (ou personnalisation) des parcours. Cependant, les premières initiatives constatées (souvent liées à des établissements isolés) sont hétérogènes. Elles sont *a priori* peu compatibles les unes avec les autres. Elles ne prennent pas en compte les recommandations concernant les normes et standards européens ou internationaux. La plupart des dispositifs techniques mis en place par les établissements ne permettent pas la circulation fluide des données destinées à valoriser les compétences des étudiants d'un établissement à l'autre, d'une structure d'aide à l'insertion professionnelle à l'autre, auprès des recruteurs et employeurs potentiels, en France comme en Europe. Cette situation peut, à terme, constituer une impasse coûteuse et préjudiciable pour les établissements comme pour les étudiants.

La deuxième préoccupation interroge une inversion de paradigme culturellement et institutionnellement délicate : d'une logique d'enseignement plutôt centrée sur « l'activité de transmission de savoirs par les enseignants » à « une logique de construction de compétences centrée sur l'apprentissage et la construction de connaissances par l'étudiant ». En effet, la démarche nécessite de mieux considérer autrement les acquis de la formation, au-delà des contenus des curricula, pour davantage raisonner « compétences » et valorisation des acquis de l'expérience. Cette inversion de paradigme, à laquelle participe la démarche ePortfolio, peut nécessiter une formation et un accompagnement non seulement en direction des enseignants mais aussi des étudiants. D'une manière générale, il s'agit de développer un enseignement davantage centré sur la construction de connaissances et de compétences par l'étudiant lui-même, tout en lui mettant à disposition un ensemble de ressources éducatives qu'il devra s'approprier avec l'aide de l'enseignant. Une telle évolution induit la décentration d'un enseignement régi par la logique de transmission de contenus pour s'orienter vers un modèle centré sur l'aide à l'apprentissage. Ainsi, ce déplacement traduit une mutation pédagogique à laquelle les enseignants doivent être sensibilisés et préparés.

La troisième préoccupation est liée au fait que le ePortfolio est un objet personnel appartenant à son auteur, en l'occurrence l'étudiant. De plus, il doit être connectable à des systèmes institutionnels d'information, de formation et de valorisation (formation, certification,

dipl  mation, insertion professionnelle...). L'espace de valorisation et de d  veloppement personnel doit pouvoir se connecter avec d'autres espaces (institutionnels, formels ou informels) ou r  seaux sociaux. D  s lors, il s'agit de penser la mall  abilit   et la porosit   dans l'ensemble des dispositifs techniques afin de permettre la cr  ation d'une cha  ne entre la personne, les entreprises, les organismes de formation dans le contexte de la FTLV.

D  s lors se pose la question de la p  rennit   et de l'authenticit   et la propri  t   des donn  es ainsi que la dur  e du support et de l'engagement de service garanti par les   tablissements. Il convient de se pr  occuper des normes et standards, en prenant en compte que le fait que l'interop  rabilit   se base maintenant sur la normalisation des processus de traduction d'  changes et de passerelles (notamment pour ce qui concerne la description des comp  tences).

2. D  finition et clarification

Afin de faciliter les   changes entre l'ensemble des parties prenantes au sein des   tablissements, il est n  cessaire de stabiliser les termes employ  s autour des propositions de clarification qui suivent.

1.1. Principes

Le ePortfolio est la propri  t   intellectuelle de son auteur, en l'occurrence de l'  tudiant. Son contenu lui appartient et rel  ve de sa vie priv  e. Il en a la ma  trise d'usage ; lui seul d  cide des donn  es qu'il souhaite publier.

L'  tablissement s'assure du pilotage de la globalit   de la d  marche ePortfolio ainsi que de la qualit   du dispositif technique institutionnel, support    cette d  marche. Int  gr   dans le syst  me d'information de l'  tablissement, ce dispositif garantit l'interop  rabilit   : il prend en compte les normes et standards internationaux existants ou en cours d'  laboration, ainsi que la s  curit   des donn  es.

1.2. La d  marche ePortfolio

Comme soulign   pr  c  demment, est consid  r  e comme relevant d'une "d  marche ePortfolio", toute d  marche d'analyse r  flexive d'un   tudiant sur son parcours, ses apprentissages, ses exp  riences, ses comp  tences ou encore ses r  alisations. La d  marche, pouvant si besoin   tre accompagn  e, vise    identifier,    expliciter et    formaliser l'ensemble de ces dimensions biographiques - tout en les valorisant et les capitalisant - dans un environnement num  rique (le ePortfolio). Cet outil n'est que la partie visible d'une telle d  marche.

Rappelons qu'il existe trois types de d  marches pouvant, si besoin, s'articuler entre-elles :

- *D  marche ePortfolio d'apprentissage*, avec la vis  e d'identification de la progression dans les apprentissages et des acquis en termes de savoirs ;
- *d  marche ePortfolio d'  valuation*, avec la vis  e d'  valuation de connaissances ou de comp  tences ;
- *d  marche ePortfolio de pr  sentation*, avec la vis  e de la mise en valeur des savoirs ou comp  tences acquises, du parcours, des exp  riences v  cues, des productions.

Le dispositif technique ePortfolio    vocation    supporter, ind  pendamment ou en ensemble, ces trois d  marches.

Le pilotage de la d  marche ePortfolio consiste    organiser et    maintenir un ensemble de dispositifs (notamment de formation et techniques) coh  rents et coordonn  s au sein de l'  tablissement, en vue d'accompagner les   tudiants dans la d  marche ePortfolio. La d  marche prend tout son sens si l'  tudiant l'associe    la question de son projet professionnel et personnel, garantissant ainsi que, quel que soit le moment o   il entre ou quitte l'  tablissement, ses acquis soient capitalis  s. Il revient    l'  tablissement d'inscrire cette d  marche dans sa politique de mise en   uvre de la formation tout au long de la vie (FTLV) et, de fait, de s'interroger sur l'importation, l'exportation des donn  es ainsi que la dur  e d'acc  s aux services num  riques supportant la d  marche ePortfolio.

Actes du 24^e colloque de l'Adm  e-Europe
L'  valuation des comp  tences en milieu scolaire et en milieu professionnel

1.3. Le ePortfolio

Le ePortfolio (Cloutier, Fortier, & Slade, 2006; De Rozario, 2005; Endrizzi, Gausse, & Leclercq, 2005; Gauthier, 2008; Ravet, 2009) est un ensemble   volutif de documents et de ressources   lectroniques capitalis  s dans un environnement num  rique d  crivant et illustrant l'apprentissage, l'exp  rience, les comp  tences ou le parcours de son auteur. Accessible    distance via une technologie interop  rable, un ePortfolio s'appuie sur une base de donn  es personnelles (informations, documents ou liens accessibles via internet) et un (ou plusieurs) espace(s) collectif(s) de publication s  lective.

L'auteur du ePortfolio doit pouvoir ma  triser le contenu et les services associ  s du ePortfolio qu'il veut partager, sous son contr  le, avec des tierces personnes, notamment pour :

- Capitaliser ses exp  riences tout en apportant les preuves de la ma  trise de comp  tences (scientifiques, d'ing  nierie, sociales, etc.) ;
- permettre la validation, la certification ou la valorisation de ses acquis de l'  ducation, de la formation ou de l'exp  rience ;
- favoriser l'autod  termination de son parcours de formation (initiale et continue) ;
- accompagner son insertion professionnelle ou son d  veloppement personnel et professionnel tout au long de la vie ;
- obtenir une lisibilit   professionnelle sur Internet et se d  marquer par la singularit   de ses exp  riences, de son projet, de son parcours.

1.4. Le dispositif technique ePortfolio

Le dispositif technique ePortfolio doit   tre r  fl  chi en tenant compte de ces trois niveaux d'organisation (Ravet, 2009) :

- *Le syst  me de gestion de ePortfolios (SGeP)* est un dispositif technique permettant    une organisation de g  rer un ensemble de ePortfolios en ad  quation avec le syst  me d'information et la gestion du dispositif, telle qu'elle sera d  finie au pr  alable par l'organisation.
- *Le syst  me ePortfolio (SeP)* est un ensemble de services num  riques institutionnels permettant    une personne (ou une organisation) d'archiver les r  sultats de ses apprentissages, de les relier entre eux et    d'autres sources d'informations (autres documents, bases de donn  es de comp  tences) et de publier des portfolios adapt  s aux besoins d'audiences particuli  res. C'est   galement dans le SeP que s'organisent les   changes entre les acteurs (  tudiants, enseignants, tuteurs, ma  tres de stage, administration...).
- *Le ePortfolio (eP)* est un espace personnel de l'  tudiant/apprenant avec un ensemble de services lui permettant d'organiser ses donn  es.

Ce dispositif technique support    la d  marche ePortfolio (Ravet 2009) comprend diff  rents outils qu'il est n  cessaire de distinguer, tout en soulignant l'importance de leurs articulations :

- *Une archive personnelle* contenant les preuves de comp  tences et acquis de l'  ducation, de la formation, de l'exp  rience personnelle ou professionnelle (permettant notamment d'  laborer un dossier de comp  tences).
Exemple de fonctionnalit  s : Consultation du dossier scolaire, universitaire, du parcours professionnel et personnel.
- *Un syst  me d'  dition* permettant de s  lectionner des   l  ments de cette archive, de les relier entre eux ou avec des sources ext  rieures (par exemple, lier une preuve de comp  tence avec un r  f  rentiel de comp  tences acad  mique ou professionnel).
Exemple de fonctionnalit  s : Edition de CV, questionnaire d'artefacts/de production, acc  s    un espace r  flexif (blogs, wiki...), acc  s    des communaut  s d'appartenance, exportation des donn  es et acc  s aux ressources externes d  di  es    l'orientation et    l'insertion

professionnelle (ONISEP, CIDJ, APEC, Organisations professionnelles...), import et exports de données, système de recherche multicritères.

- *Un système de publication* permettant de communiquer le résultat de ce travail d'édition, à un formateur ou un tuteur (dans le cadre d'une formation), à un évaluateur (dans le cadre d'une certification/diplomation), à un employeur potentiel (pour une recherche d'emploi), ou à toute autre personne ou entité à qui l'étudiant voudrait le communiquer.
Exemple de fonctionnalités : publication de CV, envoi de lien et de documents.
- *Un système de gestion* permettant à une organisation (une université ou un employeur) de gérer un ensemble de ePortfolio en fonction de l'objectif de l'organisation : suivi de l'évaluation, gestion des compétences, gestion des carrières.
Exemple de fonctionnalités : Saisie de notes, d'appréciation, d'évaluation, consultation du dossier scolaire et universitaire, du parcours professionnel.

Le dispositif technique support à la démarche ePortfolio doit permettre l'exploitation des flux de données et d'information entre les espaces personnels et les espaces institutionnels (principes de porosité et de malléabilité).

3. Conclusion provisoire : à propos de la professionnalisation des enseignants du supérieur au regard de la démarche ePortfolio

Différentes tensions induites par les inversions de paradigmes sous-tendues par la démarche ePortfolio interrogent à la fois l'identité professionnelle, la professionnalité et la professionnalisation des enseignants du supérieur :

- La formation "orientée compétences" vs la formation "académique disciplinaire" ;
- l'identité professionnelle "enseignant-accompagnateur" vs l'identité professionnelle "enseignant-chercheur" ;
- le dispositif pédagogique "centré sur l'étudiant qui co-construit ses savoirs et ses compétences" avec l'enseignant vs le dispositif pédagogique "centré sur l'enseignant qui transmet son savoir et ses connaissances".

L'analyse de ces tensions permettra de mettre en évidence certains paramètres (notamment modalités d'accompagnement, de formation, de reconnaissance...) qu'il conviendra de prendre en compte afin d'initier collégialement la démarche ePortfolio dans l'établissement.

Il apparaît essentiel de veiller à ancrer l'engagement des enseignants-chercheurs dans la démarche ePortfolio par une approche pédagogique et de reconnaissance professionnelle adaptée afin d'éviter qu'elle ne se limite à une exécution d'une exigence institutionnelle ou qu'un petit nombre d'universitaires ne deviennent les seuls représentants de ce changement. Les démarches ayant abouti dans ce domaine démontrent que les changements ne s'opèrent que si l'enseignant est rapidement convaincu de leur légitimité. La prise en compte des représentations des enseignants face aux usages du numérique dans leur pratique et celles de leurs étudiants et la mise en évidence des exigences sociétales actuelles dans un échange sont des leviers pour amorcer les changements à opérer. Par ailleurs, une reconnaissance de ces activités dans le référentiel des missions des enseignants-chercheurs de l'établissement devra être associée, ainsi qu'un travail d'aménagement éventuel des volumes horaires des programmes de formation si, auparavant, ils n'incluaient pas d'heures d'accompagnement de l'étudiant.

La mise en œuvre d'un projet ePortfolio impactera aussi d'une part le travail des ingénieurs, techniciens ou personnels administratifs en charge de l'organisation et de la mise en œuvre du dispositif pour ce qui concerne les aspects technique et fonctionnel, d'autre part le travail des personnels en charge de l'accompagnement des étudiants dans la construction de leur projet (conseiller d'orientation, documentalistes, responsables de l'orientation et de l'insertion professionnelle...)

Il conviendra de s'assurer que les interactions entre les différentes fonctionnalités du ePortfolio (formation, évaluation, insertion professionnelle...) et entre les accompagnateurs se fassent simplement et de manière transparente pour les étudiants.

Les établissements devront mettre en place :

- Des actions de communication et d'information auprès des ingénieurs, techniciens ou personnels administratifs pour les informer et les impliquer dans la mise en œuvre du projet ;
- des actions de formation à l'accompagnement et à la prise en main des outils et applications.

Ainsi, au moment de l'élaboration du projet stratégique fondateur de la démarche ePortfolio, il conviendra de bien prendre en compte l'état des forces et des complémentarités en présence, ainsi que les attentes et les besoins de l'ensemble des parties prenantes afin de construire collégialement les dispositifs d'accompagnement et de formation de l'ensemble des personnels de l'établissement.

Enfin, dans le contexte spécifique de l'enseignement supérieur, il est nécessaire, pour des raisons culturelles spécifiques au corps des enseignants chercheurs, de justifier la conception et le déploiement de ces projets sur des modèles et des références scientifiques relatifs au développement des compétences, à l'ingénierie de formation et de l'acceptation sociale des innovations dans les contextes professionnels. Ainsi, si à l'évidence de nombreux types d'acteurs doivent être mobilisés (un seul corps professionnel ne peut prétendre à lui seul avoir une compétence globale et exclusive), le pilotage, la conception, la mise en œuvre et l'évaluation des dispositifs de formation et d'accompagnement doivent s'appuyer sur des modèles et des références scientifiques explicites (Heutte, Lameul, & Bertrand, 2010).

Le développement d'un lien organique étroit entre le déploiement de la démarche ePortfolio et la recherche, notamment celle relative à la pédagogie universitaire, sera certainement l'une des clés de la conception de dispositifs en mesure de favoriser objectivement la réussite des étudiants.

4. Références et bibliographie

- Cloutier, M., Fortier, G., & Slade, S. (2006). *Le portfolio numérique, un atout pour le citoyen apprenant*. Bibliothèque nationale du Canada.
- De Rozario, P. (2005). *Portfolio numérique dans La Formation*. Présenté au Séminaire national "Le portfolio numérique dans la formation : le cas du C2i enseignant", organisé par la Conférence des Directeurs d'IUFM et Ministère de l'éducation nationale, de l'enseignement supérieur et de la recherche. Dijon. Consulté de <http://www.dijon.iufm.fr/static/tice/sem-port/>
- Endrizzi, L., Gaussel, M., & Leclercq, F. (2005). *Les publications électroniques*. (Veille scientifique et technologique No. 8). INRP. Consulté de <http://ife.ens-lyon.fr/vst/LettreVST/08-mars-2005.php>
- Gauthier, P.-D. (2008). *Exploiter son portfolio numérique: construire son identité professionnelle numérique pour valoriser ses compétences/The use of a personal digital portfolio: how to build its own professional digital identity and enhance its competences*. Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie, 34(3).
- Heutte, J., Lameul, G., & Bertrand, C. (2010). *Dispositifs de formation et d'accompagnement des enseignants du supérieur: point de situation et perspectives française concernant le développement de la pédagogie universitaire numérique*. Actes du 7e Colloque Technologies de l'Information et de la Communication pour l'Enseignement (TICE 2010) (p. 6–8).
- Ravet, S. (2009). *Leçon sur le ePortfolio — EIfEL*. Consulté mars 14, 2011, de <http://www.eife-l.org/publications/eportfolio/eportfolio-lessons>

Évaluer le portfolio pour évaluer des compétences

Une grille d'évaluation doublement critériée

Evelyne Vanden Eynde

Résumé

Le service d'aide à la réussite de la Haute Ecole Francisco Ferrer organise depuis quelques années déjà un séminaire méthodologique obligatoire visant à donner aux étudiants de première année de bachelier les bases d'une méthode de travail efficace et d'une réflexivité accrue quant à leur choix professionnel. A cette occasion, il a semblé opportun de leur faire élaborer un portfolio d'apprentissage qui non seulement témoignerait de leur évolution au cours du premier quadrimestre et de l'acquisition des compétences visées, mais servirait également de base à l'évaluation du séminaire en question. Afin d'objectiver au maximum cette évaluation, l'élaboration d'une grille critériée était indispensable.

La particularité des modalités d'évaluation auxquelles l'équipe pédagogique du séminaire a recours se situe donc à deux niveaux.

D'une part, les items d'évaluation sont fixés de manière précise, à priori. Et d'autre part, l'équipe a travaillé la critérisation des niveaux de satisfaction de ces items, sous forme de degré de maîtrise de chacune des compétences visées.

Cet outil d'évaluation facilite le travail des différents enseignants en charge du séminaire et permet d'assurer, au mieux, une certaine homogénéité dans la correction et l'évaluation. De plus, les étudiants peuvent s'en servir comme un outil d'auto-évaluation afin d'améliorer leur portfolio tout au long de son élaboration et jusqu'à sa présentation.

Notons que la compétence terminale du séminaire en question est d'apprendre à apprendre, et pour cela quatre compétences transversales à la fois cognitives et socio-affectives sont mises en avant : organiser, communiquer, analyser et s'engager.

Afin de préciser la nature de ces compétences, nous avons mis en évidence cinq ressources – jouant le rôle d'indicateur – pour chacune d'elles. Puisqu'en effet, une compétence se définit comme un savoir-agir complexe prenant appui sur une variété de ressources. Comme précisé ci-avant, pour chacune de celles-ci, l'analyse et la cotation sont découpées en quatre niveaux selon le degré de satisfaction aux critères.

Mots clés

Aide à la réussite – portfolio – compétences – grille – (auto-)évaluation.

Introduction

Depuis la création de la HEFF en 1995, un service d'aide à la réussite et d'orientation est à la disposition des étudiants de première année.

Tout au long de ces années passées et plus récemment avec la réforme de Bologne, les préoccupations de nos autorités politiques et académiques, tant en Belgique qu'à l'étranger, se sont multipliées en faveur de la promotion de la réussite des étudiants de 1^{ère} année de l'enseignement supérieur au sein des universités et des Hautes Ecoles. En effet, un taux de réussite avoisinant les 40% est devenu une réalité insupportable. De nouveaux moyens pour favoriser la réussite en 1^{ère} année et améliorer la qualité de l'Enseignement supérieur sont, dès lors, alloués aux Hautes Ecoles.

Dès lors, en 2007, la Haute Ecole Francisco Ferrer a mis sur pied un dispositif d'aide à la réussite visant l'intégration de sa population étudiante, dont une proportion certaine était issue de milieux socio-économiques défavorisés, dans l'enseignement supérieur. Celui-ci a pour ambition d'offrir aux étudiants de première génération les clés d'une méthodologie de travail efficace, des pistes de réflexion quant à leur choix d'orientation étudiante et professionnelle, ainsi que des occasions de s'intégrer dans leur nouvel environnement.

Afin, d'éviter toute discrimination entre les étudiants qui auraient besoin de ce dispositif d'aide, la Haute Ecole a rendu celui-ci obligatoire pour tous et de ce fait, a valorisé le séminaire par des crédits ECTS.

C'est pourquoi, il a fallu trouver un moyen juste et efficace d'évaluer l'acquisition des compétences méthodologiques visées durant ce séminaire. L'équipe des conseillers pédagogiques et des promoteurs de la réussite s'est mise d'accord sur le portfolio comme outil d'évaluation des compétences visées par le séminaire. Restait alors à évaluer ce portfolio et à élaborer une grille d'évaluation critériée.

Cadre théorique

Naccache, Samson et Jouquan (2006) stipulent que le « portfolio est une collection organisée et cumulative de travaux et de réflexions d'un étudiant qui rassemble des informations sur les compétences qu'il a développées au cours d'une période plus ou moins longue d'apprentissage »

Quant à lui, Tardif (2006) définit le portfolio comme « un dossier apte à fournir les documents nécessaires pour porter un jugement sur la qualité des performances. C'est un dossier évolutif rassemblant les travaux des apprenants ».

Toujours en se basant sur Jaques Tardif (2006), nous définissons les compétences – ici évaluées – comme un savoir-agir complexe prenant appui sur une variété de ressources.

Analyse de l'outil

Evaluer des compétences méthodologiques et réflexives au travers du portfolio.

Dans le contexte du séminaire d'*Activités d'Intégration Liées à l'Enseignement Supérieur (AILES)*, il fallait trouver un moyen d'évaluer et de valoriser le travail accompli par les étudiants tout au long du séminaire ainsi que l'acquisition des compétences transversales visées. Les étudiants seraient en mesure de mobiliser celles-ci dans chacun de leurs enseignements d'abord, et dans leur vie professionnelle ensuite.

En effet, la compétence terminale des AILES est d' apprendre à apprendre, et pour cela quatre compétences transversales à la fois cognitives et socio-affectives sont mises en avant : organiser, communiquer, analyser et s'engager.

Afin de préciser la nature de ces compétences, nous avons mis en évidence cinq ressources¹ – jouant le rôle d'indicateur – pour chacune d'elles. Puisqu'en effet, une compétence se définit comme un savoir-agir complexe prenant appui sur une variété de ressources.

Afin d'attester de leur parcours, de leurs questionnements et de leurs progrès, l'outil portfolio avait été retenu comme témoin et accompagnateur de leur apprentissage. Celui-ci peut être utilisé comme outil d'évaluation des apprentissages puisqu'il permet de recueillir et conserver des échantillons des réalisations de l'élève. Le portfolio s'inscrit dans une démarche d'évaluation formative continue et permet une évaluation en collaboration avec l'étudiant. Il constitue également un outil de communication interactive entre l'apprenant et l'enseignant et est une source complémentaire d'informations sur l'apprenant, qui serait utilisée pour nuancer un jugement ou démontrer d'un projet. (Baron et Bruillard, 2003).

Il s'agissait dès lors, de trouver un moyen valide d'évaluer.

Les avantages que présente le portefeuille étudiant se situent dans la motivation que les étudiants peuvent retirer de cette implication nouvelle dans leur formation. Il est le témoin de leur progression et se contente de mettre l'accent sur le meilleur. En effet, les étudiants doivent en quelque sorte gérer les preuves de leur apprentissage et choisir ce qu'ils y démontrent. Le portfolio offre la possibilité d'une actualisation permanente des traces du processus d'apprentissage de l'étudiant. Et facilite une ouverture sur le monde permettant l'intervention (pédagogique) d'une tierce personne. Le portfolio étant, avant tout, un outil de présentation, il implique également une structuration de son travail optimale et démultiplie les possibilités d'expression.

D'un point de vue purement pédagogique, le portfolio offre la possibilité aux étudiants de prendre conscience de leur progression dans l'apprentissage, ainsi que du rôle de l'écrit dans la réflexivité sur l'apprentissage et les stratégies qu'ils déploient à cet effet. De plus, de par le fait qu'il constitue également un outil d'évaluation, les étudiants ont également l'occasion de développer

¹ En annexe 2, le référentiel des compétences et ressources du dispositif d'Activités d'Intégration Liées à l'Enseignement Supérieur

des facultés d'auto-évaluation et de recevoir une rétroaction sur leurs productions. Et par là, il responsabilise l'étudiant face à ses apprentissages.

De plus, en termes d'évaluation, le portfolio donne lieu à une évaluation sommative des réalisations de l'apprenant, où l'on n'évalue pas un travail à un moment donné, mais un ensemble de travaux collectés au cours d'une phase d'apprentissage. En effet, la fonction principale de l'évaluation est bien d'aider l'apprenant dans son cheminement et de permettre à l'enseignant d'ajuster ses interventions pédagogiques. Le portfolio pourrait de ce fait être utilisé pour jauger les compétences des apprenants indépendamment de tout système de formation. Et enfin, son utilisation permet l'élaboration de critères d'évaluation communs à l'apprenant et à l'enseignant.

En somme, le portfolio permet tout d'abord de développer les capacités métacognitives des apprenants en leur demandant de s'auto-évaluer, d'observer leurs progrès en réfléchissant à leurs productions, en les évaluant et en prenant conscience de leurs forces et faiblesses

Le portfolio de l'apprenant serait donc une collection de travaux sélectionnés par l'apprenant et/ou l'enseignant comme preuve de sa maîtrise des compétences requises ou de ses progrès. Il est, dans un second temps, commenté par l'enseignant autour de ses réalisations, de son cheminement et de ses réflexions.

Evaluer un portfolio au travers d'une grille d'évaluation

Après avoir répondu « le portfolio » à l'interrogation première « Quelle évaluation pour ces compétences ? » ; l'équipe responsable du séminaire s'est donc réunie à nouveau, autour de la question « Quelle évaluation pour le portfolio ? »

Il paraissait trivial pour les membres de l'équipe pédagogique d'opter pour un portfolio afin de rendre compte de leurs apprentissage, de leur réflexivité et de leur parcours. Un examen n'aurait qu'insuffisamment pu rendre compte de tout cela. Cependant, le portfolio présentait un inconvénient non négligeable en ce qu'il comporte de subjectivité. Là où la subjectivité règne, il est évident que l'évaluation est délicate. Il importe donc, afin de rendre compte de la maîtrise des compétences visées et du développement de l'étudiant, une grille d'évaluation répondant à des critères objectifs.

En ce qui concerne la partie écrite portfolio, il fallait distinguer la forme du fond. Pour la forme, sont donc valorisées l'organisation du document et l'expression écrite. Plus précisément, il faut que le lecteur soit capable de repérer et distinguer chacune des parties du portfolio et que la syntaxe, l'orthographe et le vocabulaire soient adéquats. En ce qui concerne le fond, il a été réparti selon les différentes activités. En effet, il est requis que l'étudiant dispose des traces et des analyses de chacune d'elles, ainsi qu'une synthèse globale sous la forme d'un texte intégrateur.

La présentation orale était également sujette à une évaluation en regard de la forme et des contenus abordés. Concernant la forme, le support powerpoint devait répondre aux critères de nombre et de qualité et l'expression écrite et orale était également retenues, ainsi que la posture

de l'étudiant et sa capacité à répondre de manière pertinente aux questions. Sur le fond, les informations doivent être représentatives et pertinentes et l'étudiant doit démontrer d'un esprit d'analyse et de synthèse.

Pour chacun des items d'évaluation, l'analyse et la cotation sont découpées en quatre niveaux selon le degré de satisfaction aux critères définis en extension².

² Exemple de critères définis en extension pour l'item « organisation du document écrit » :

- niveau 1 : les parties ne sont pas identifiées et sont présentées dans le désordre
- niveau 2 : toutes les parties ne sont pas identifiées mais présentées dans l'ordre prescrit
- niveau 3 : toutes les parties sont identifiées mais présentées dans le désordre
- niveau 4 : toutes les parties sont identifiées et présentées dans l'ordre prescrit

Grille d'évaluation synthétique³

PRÉSENCE AUX ACTIVITÉS		6
PORTFOLIO		30
Forme	Organisation du document	3
	Expression écrite	3
Fond	Introduction	3
	Traces et analyse	24
	PPP	4
	Interviews	4
	Team Building	4
	MT	4
	Bibliothèque	4
	Maîtrise du discours	4
	Synthèse finale (texte intégrateur)	6
PRÉSENTATION ORALE		24
Forme	Support électronique	3
	Expression écrite	3
	Expression orale	3
	Présentation & posture	3
Fond	Pertinence des informations	3
	Analyse des informations	3
	Capacité de synthèse	3
	Ouverture au dialogue	3

³ La grille d'évaluation complète et critériée se situe en Annexe p.8

Conclusion

Quand des portfolios sont utilisés à des fins de contrôle du développement personnel et académique d'un étudiant, il est important de savoir comment il s'évalue et s'analyse. Par conséquent, il est crucial que les portfolios utilisés de cette façon contiennent des réflexions écrites. Celles-ci sont généralement organisées autour des compétences visées par le dispositif en place. En l'occurrence dans le cas des « Ailes », des compétences transversales cognitives et socio-affectives. Et le portfolio est alors un lieu de réflexion sur la façon dont leurs accomplissements sont reliés aux objectifs fixés. Ce type d'évaluation diffère des évaluations traditionnelles puisque les étudiants ne sont pas examinés mais sont invités à prouver leur compétence au travers de la réalisation et de la présentation de leur portfolio. Pour être valides, ces évaluations doivent répondre à des critères à la fois précis, partagés et communiqués.

La particularité des modalités d'évaluation auxquelles l'équipe pédagogique des « Ailes » a recours se situe à deux niveaux.

D'une part, les items d'évaluation sont fixés *de manière précise*, à priori. Et d'autre part, l'équipe a travaillé la critérisation des niveaux de satisfaction de ces items, sous forme de *degré de maîtrise* de chacune des compétences visées.

Cet outil d'évaluation facilite le travail des différents enseignants en charge du séminaire et permet d'assurer, au mieux, une certaine homogénéité dans la correction et l'évaluation. De plus, les étudiants peuvent s'en servir comme un outil d'auto-évaluation afin d'améliorer leur portfolio tout au long de son élaboration et jusqu'à sa présentation.

Plus qu'un simple outil, le portfolio est une démarche, un dispositif complexe qui nécessite une organisation pédagogique adaptée et un investissement (en temps) conséquent de la part des élèves comme des enseignants.

Bibliographie

ALLAL Linda et WEGMULLUER Edith. 1998. *Le portfolio dans la dynamique de l'entretien tripartite*. Mesure et évaluation en éducation. Vol. 20 n°3 pages 5 à 31.

BEHRENS Mathis. 1998. *Le portfolio, un moyen d'améliorer le dialogue entre formateur et apprenant*. Mesure et évaluation en éducation. Vol 20 n°3 pages 55 à 83.

DESJARDINS Richard. 2002 *Le portfolio de Développement Professionnel Continu*. Chenelière McGraw-Hill, 108, p.

FORGETTE-GIROUS Renée et SIMON Marielle. 1998. *L'application du dossier d'apprentissage à l'université*. Mesure et évaluation en éducation. Vol 20 n°3 pages 85 à 103.

GOUPIL Georgette, LUSSIER-DESROCHERS Dany, BERTHIAUME Caroline, LEGAULT Anik et HENAULT. 2000. *Le portfolio : un exemple d'application dans un cours de cycle supérieur*. Res Academica Vol. 18 n°1 et 2.

McMULLAN Mirjan, ENDACOTT Ruth, GRAY Morag, JASPER Melanie, MILLER Carolyn M.L., SCHOLLES Julie et WEBB Christine. 2003. *Portfolios and assessment of competence : a review of the literature*.

PAULSON F. Leon, PAULSON Pearl R et MEYER Carol. 1991. *Whar makes a Portfolio a Portfolio ?* Educational Leadership

PAULSON F. Leon et PAULSON Pearl R. 1994. *Assessing Portfolios Using The Constructivist Paradigm*. Paper presented at the annual meeting of American educational Research Association, New Orleans, L.A.

SCALLON Gerard. 2004. *L'évaluation des apprentissages dans une approche par compétences*, Ed. De Boeck, coll. Pédagogies en développement, Bruxelles

TARDIF Jacques, 2006, *L'évaluation des compétences*. Ed. Chenelière Education, coll. Chenelière didactique, Montréal

Annexe 1

Grille d'évaluation

Nom :

Classe :

Présence aux activités		3	2	1	0	/6	
		L'étudiant est présent à toutes les activités et y participe activement	L'étudiant est présent à toutes les activités sans y participer activement ou est présent à la moitié des activités en participant activement	L'étudiant est présent à moins de la moitié des activités	L'étudiant n'a aucune activité	/3	
Portfolio		3	2	1	0	/30	
FORME	Organisation du document	Toutes les parties sont identifiées et présentées dans l'ordre prescrit	Toutes les parties sont identifiées mais présentées dans le désordre	Toutes les parties ne sont pas identifiées mais présentées dans l'ordre prescrit	Les parties ne sont pas identifiées et présentées dans le désordre	/3	
	Expression écrite	Pas de faute de syntaxe ni d'orthographe. Le vocabulaire est élaboré et varié	Jusqu'à 10 fautes d'orthographe et de syntaxe. Le vocabulaire est varié	Jusqu'à 10 fautes d'orthographe et de syntaxe. Le vocabulaire est restreint	Au-delà de 10 fautes d'orthographe ou de syntaxe	/3	
FOND	Introduction	Introduction complète : contexte, contenus et moyens	Introduction partielle : contexte et contenus	Introduction partielle : contexte	Absence d'introduction	/3	
	Traces et analyse des activités	Trace et analyse réflexive de l'activité en regard du métier de l'étudiant	Trace de l'activité et analyse de celle-ci décontextualisée par rapport aux objectifs du séminaire	Trace de l'activité, sans analyse	Pas de trace de l'activité	ppp	/3
						itv	/3
						TB	/3
						MT	/3
						bib	/3
						fr	/3
	Texte intégrateur Synthèse finale	Le texte témoigne de l'évolution, des progrès réalisés et de la capacité à porter un regard sur soi-même	Le texte ne témoigne pas de l'évolution et des progrès réalisés mais il apparaît une capacité à porter un regard sur soi-même	Le texte témoigne de l'évolution et des progrès réalisés mais la capacité à porter un regard sur soi-même n'apparaît pas	Le texte ne témoigne pas de l'évolution ni de la capacité à porter un regard sur soi-même	/3	

Présentation orale		3	2	1	0	/24
FORME	Support électronique	Le diaporama de maximum 5 dias répond aux critères qualité	Le diaporama de maximum 5 dias ne répond pas aux critères de qualité	Le diaporama comprend plus de 5 dias	Absence de diaporama	/3
	Expression écrite	Pas de faute de syntaxe ni d'orthographe. Le vocabulaire est élaboré et varié	Jusqu'à 5 fautes d'orthographe et de syntaxe. Le vocabulaire est varié	Jusqu'à 5 fautes d'orthographe et de syntaxe. Le vocabulaire est restreint	Au-delà de 5 fautes d'orthographe ou de syntaxe	/3
	Expression orale	La construction du discours et la diction sont très bonnes	La construction du discours et la diction sont bonnes	La construction du discours et la diction sont satisfaisantes	La construction du discours et la diction sont insatisfaisantes	/3
	Ouverture au dialogue	L'étudiant est ouvert au dialogue et rebondit sur les questions et les commentaires	L'étudiant est ouvert au dialogue mais ne rebondit pas sur les questions et les commentaires	L'étudiant rebondit sur les critiques mais extériorise les causes	L'étudiant n'accepte pas le dialogue et extériorise les causes	/3
	Présentation / Posture	L'étudiant est dynamique. Il regarde et s'adresse à tous les membres du groupe.	L'étudiant regarde et s'adresse au groupe et présente une attitude figée	L'étudiant ne regarde pas l'ensemble du groupe et présente une attitude figée	L'étudiant est statique. Il ne s'adresse pas à l'ensemble du groupe	/3

CONTENUS	Pertinence des informations	Les informations sont pertinentes pour comprendre le fonctionnement de l'étudiant dans le cadre de sa formation	Les informations ne permettent de comprendre que partiellement le fonctionnement de l'étudiant dans le cadre de sa formation	Les informations ne permettent pas de comprendre le fonctionnement de l'étudiant dans le cadre de sa formation	Le choix des informations n'est pas pertinent et n'est pas en relation avec la situation de l'étudiant dans le cadre de sa formation	/3
	Analyse des apprentissages	Forces et besoins d'apprentissage présents et identifiés	Forces et faiblesses partiellement identifiés	Pas de réflexion et peu d'identification des besoins d'apprentissage	Aucune réflexion et aucune identification des besoins d'apprentissage	/3
	Capacité de synthèse	La présentation témoigne de l'évolution, des progrès réalisés et de la capacité à porter un regard sur soi-même	La présentation ne témoigne pas de l'évolution et des progrès réalisés mais il apparaît une capacité à porter un regard sur soi-même	La présentation témoigne de l'évolution et des progrès réalisés mais la capacité à porter un regard sur soi-même n'apparaît pas	La présentation ne témoigne pas de l'évolution ni de la capacité à porter un regard sur soi-même	/3

TOTAL	/60
Note finale	/20

Apprendre à apprendre

```
graph TD; A[Apprendre à apprendre] --> B[ORGANISER]; A --> C[COMMUNIQUER]; A --> D[ANALYSER]; A --> E[S'ENGAGER]; B --> B1[Organiser un plan de travail]; B --> B2[Organiser un planning]; B --> B3[Organiser un document écrit]; B --> B4[Organiser une communication orale]; B --> B5[Organiser un travail d'équipe]; C --> C1[Communiquer à l'écrit]; C --> C2[Communiquer dans un vocabulaire correct]; C --> C3[Communiquer à l'oral]; C --> C4[Communiquer devant un groupe]; C --> C5[Communiquer ses compétences : méta-communiquer]; D --> D1[Analyser des ressources]; D --> D2[Analyser une situation]; D --> D3[Analyser et synthétiser une information]; D --> D4[Analyser une progression]; D --> D5[Analyser son travail : auto-évaluation]; E --> E1[S'engager dans un groupe]; E --> E2[S'engager dans un projet]; E --> E3[S'engager dans une démarche professionnelle]; E --> E4[S'engager dans une réflexion personnelle]; E --> E5[S'engager dans l'E.S.];
```

ORGANISER

Organiser un plan de travail

Organiser un planning

Organiser un document écrit

Organiser une communication orale

Organiser un travail d'équipe

COMMUNIQUER

Communiquer à l'écrit

Communiquer dans un vocabulaire correct

Communiquer à l'oral

Communiquer devant un groupe

Communiquer ses compétences : méta-communiquer

ANALYSER

Analyser des ressources

Analyser une situation

Analyser et synthétiser une information

Analyser une progression

Analyser son travail : auto-évaluation

S'ENGAGER

S'engager dans un groupe

S'engager dans un projet

S'engager dans une démarche professionnelle

S'engager dans une réflexion personnelle

S'engager dans l'E.S.

LE PORTFOLIO DE DEVELOPPEMENT PROFESSIONNEL DU BACHELOR EN SCIENCES DE L'EDUCATION DU L'UNIVERSITE DU LUXEMBOURG : UN MEDIA ET UN PROCESSUS EN CONSTANTE EVOLUTION

Débora Poncelet, Michelle Brendel, Viviane Bourg, et Sylvie Kerger*

* Université du Luxembourg

debora.poncelet@uni.lu

michelle.brendel@uni.lu

sylvie.kerger@uni.lu

viviane.bourg@uni.lu

Mots-clés : Portfolios, Process, Practice, Pre-service teacher, Professional Development

Résumé. Depuis la création du Bachelor en sciences de l'Education lors de l'année académique 2005-2006, un portfolio de développement professionnel a été mis en œuvre au sein de la formation. Une équipe d'enseignants a été chargée d'améliorer le travail avec ce portfolio afin qu'il devienne un outil susceptible d'aider les étudiants à réfléchir sur mais aussi à autoréguler leurs propres apprentissages en lien au développement professionnel. Ce groupe de travail est notamment responsable de l'organisation et de la structuration des séminaires relatifs à la thématique « Becoming a teacher » au sein de laquelle le portfolio occupe une place centrale, de la façon dont les étudiants développent et structurent leur portfolio, du suivi des étudiants par les tuteurs et évidemment de l'évaluation du portfolio. Après avoir identifié un certain nombre de difficultés durant les premières années de fonctionnement et avoir réfléchi à des voies possibles d'amélioration tant du processus que de l'outil lui-même, il semble important aujourd'hui d'évaluer le dispositif actuel afin d'assurer la qualité des interventions éducatives auprès des étudiants. Cette communication vise à présenter le contexte du Bachelor en Sciences de l'éducation de l'Université du Luxembourg, les modifications suggérées ainsi que les perspectives envisagées.

1. Introduction théorique

1.1 Le métier d'enseignant au cœur du processus de changement

Depuis de nombreuses années, un débat a lieu pour savoir si les enseignants peuvent réellement être considérés comme des professionnels ou simplement des techniciens (Foster, Walker & Hyunsook Song, 2007, p. 30). Selon les auteurs, il semble évident que les enseignants ont besoin de plus qu'un manuel, que des guides pour enseignants ou de quelques astuces pédagogiques pour être en mesure de développer des activités significatives d'apprentissage et d'enseignement pour les élèves et les étudiants. "The balance has mostly recently shifted in favor of the view of teachers as professionals, especially highlighting at least one essential behavior that separates the professional from the technician: reflexive action" (p. 30).

Comme l'avance Perrenoud (1999a), le métier d'enseignant(e) évolue et se transforme. Pour faire face à ces changements, nos étudiant(e)s doivent donc apprendre à faire preuve d'adaptabilité et partant, à développer un certain nombre de compétences qui leur permettront de répondre favorablement aux situations nouvelles qu'ils rencontreront inévitablement au niveau de l'organisation et de l'animation des situations d'apprentissage, de la gestion de la progression des apprentissages, de la conception et de l'évolution des dispositifs de différenciation, de l'implication des élèves dans leurs apprentissages du travail en équipe, de la gestion de l'école, de l'information et de l'implication des parents, de l'utilisation des nouvelles technologies, des questions d'éthique professionnelle et de gestion de leur formation continue (Perrenoud, 1999a).

Pour les y préparer, le programme de formation du BScE mise notamment sur le portfolio en tant qu'outil de développement professionnel.

Le monde dans lequel nous vivons aujourd'hui génère des changements de plus en plus rapides que chaque citoyen est amené à décoder, à interpréter et auxquels il doit s'adapter. Si nous ne dotons pas aujourd'hui les enfants de clés de compréhension de l'environnement dans lequel ils vivent, il est clair qu'ils éprouveront d'importantes difficultés à trouver une place dans ce nouveau contexte de société. Des leçons sont dès lors nécessairement à tirer pour la formation des enseignants. Face à ces transformations sociétales, il semble inévitable que l'école change, qu'elle progresse elle aussi. En ce sens, la professionnalisation de l'enseignant est au cœur des débats et le développement des compétences de réflexivité, d'innovation et de coopération en sont ses corollaires. « Les enseignants, s'ils ne sont pas des intellectuels à part entière, sont au moins des médiateurs et des interprètes actifs de cultures, de valeurs, de savoir en train de se transformer » (Perrenoud, 1999b, p. 2). Il s'agit donc pour les enseignants d'être capables de construire leurs propres démarches sur la base du contexte dans lequel ils seront amenés à travailler : les besoins de leurs élèves, l'environnement scolaire, l'intérêt de faire du team-teaching, les ressources disponibles, Il est en effet clair que la formation initiale ne peut pas leur donner une fois pour toutes les acquis nécessaires et exploitables durant toute leur carrière professionnelle.

Comme l'avancent Vanhulle, Déum, Mattar et Schillings (2004), il est possible, dès la formation initiale, de pourvoir les enseignants d'habiletés réflexives en regard de leur future profession, des contenus-matières, de la méthodologie qu'ils adoptent pour faire réussir au mieux les élèves de leur classe et de façon plus générale, des valeurs et des missions assignées à l'enseignement. La réflexivité, selon Perrenoud (2001), « *postule implicitement que l'action fait l'objet d'une représentation. L'acteur est supposé savoir ce qu'il fait, il peut donc s'interroger sur les mobiles, les modalités et les effets de son action* ». Dans l'enseignement, contrairement à d'autres métiers, la réflexivité ne peut se limiter à chercher des solutions aux problèmes qui se posent. Il faut davantage la voir comme un fonctionnement stable et régulier dont l'intensité peut varier en fonction de l'objet des situations à gérer (de forte en cas de problèmes aigus à moins intense, mais toujours nécessaire, face à la gestion des faits classiques du quotidien). Un enseignant réflexif, selon Perrenoud (2005), va développer trois types d'habiletés. Il est capable d'articuler son savoir théorique à sa pratique. Il développe un « savoir-analyser » qui lui permet de cibler les informations pertinentes parmi l'ensemble des données reçues et de les interpréter de façon adéquate. Et enfin, il est capable de mener à bien une réflexion a posteriori qui regroupe l'analyse de l'expérience vécue et sa transformation en connaissances.

Les questions liées à la formation des enseignants sont nombreuses. La formation vise à développer une identité, des compétences, une attitude réflexive, mais aussi des connaissances. Par conséquent, les éducateurs et les étudiants ont à explorer différentes façons d'enseigner et d'apprendre basée sur une compréhension partagée de l'organisation du travail et les tâches à effectuer. Ainsi, l'étudiant est vraiment au cœur du processus de formation et le principal responsable de sa formation.

Les cours, les séminaires, les tutorats, ... constituent autant d'opportunités susceptibles d'influencer les représentations des étudiants, de placer les apprenants dans une situation de recherche, d'encourager le travail en groupe et la collaboration, d'assurer les processus d'évaluation formative et d'autoévaluation, etc. Les besoins des étudiants, leurs représentations, leurs questions sont réellement prises en considération en vue de développer des situations d'apprentissage appropriées au sein desquelles ces apprenants deviennent de réels acteurs de leurs apprentissages. Il est évident qu'une telle approche s'ouvre à la diversité des parcours individuels et encourage le pluralisme des contributions culturelles.

1.2 Un portfolio de développement professionnel : un moyen de faire face au changement

Depuis l'année scolaire 2005-2006, dès la création du Bachelor en Sciences de l'Education, un portfolio de développement professionnel a été implanté au sein de la formation. Tout d'abord, suite à la réforme de l'enseignement supérieur, divers projets ont vu le jour en vue de réformer la

formation initiale des enseignants. L'idée principale était de former des enseignants capables de développer une attitude réflexive et d'adopter une attitude de recherche. Ensuite, en 2009, la loi scolaire de 1912 a été réformée. Dans les classes primaires, une nouvelle approche de l'évaluation est maintenant appliquée dans laquelle le portfolio occupe une place particulière. Cette logique vise à comprendre, à suivre et à documenter les réalisations et les progrès des élèves. Dans ce contexte, pour les enseignants de formation initiale, l'approche basée sur le portfolio prend tout son sens. Si le portfolio de développement professionnel diffère dans ses objectifs du portfolio primaire, le processus de collecte, de sélection, les échanges (avec les pairs et avec les tuteurs) ainsi que le processus réflexif se retrouvent dans les deux approches. En développant leur propre portfolio, nos étudiants peuvent élaborer des stratégies de réflexion et développer une pratique qu'ils utiliseront plus tard avec leurs futurs élèves.

Il apparaît par conséquent que le portfolio de développement professionnel, par ses caractéristiques intrinsèques liées au monitoring comme à l'évaluation formative, rencontre pleinement les objectifs de formation. Il suffit pour s'en rendre compte de voir à quel point la littérature foisonne sur le sujet (for example: Klenowski, 2000; Darling, 2001; Deum and Vanhulle, 2005; Harland, 2005; Gearhart and Osmundson, 2009); Jorro, 2005; Parsons and Stephenson, 2005; Stolle, Watkins and Goerss, 2005; Vanhulle, 2005; Mansvelder-Longayroux, and Beijaard Verloop, 2007; Imhof and Picard, 2009).

Le portfolio choisi par le BScE pour être le témoin et l'accompagnateur de l'étudiant durant sa formation pédagogique universitaire est le portfolio de développement professionnel. Ce type de portfolio, destiné à accompagner la construction des compétences professionnelles, en est la trace et la mémoire. Vanhulle (2000) souligne qu'il devient ainsi un outil à partir duquel on peut entamer un dialogue avec soi-même (auto-évaluation), avec ses pairs (co-évaluation) et avec l'équipe de formation (co-évaluation et évaluation formative). Cet instrument autorise de cette façon, une forme de réflexivité chez les futurs enseignants, un souci d'articuler des questions d'expérience de classe à des questions plus théoriques, qui s'appuient tant sur le dialogue avec autrui, que sur le retour sur soi, ses croyances, ses compétences et ses décisions. Ainsi au fil du temps, le portfolio de développement professionnel devient, selon Perrenoud (1999a), un support qui aide le (futur) enseignant à apprendre "méthodiquement" de l'expérience et à transformer sa pratique d'année en année. Il semble donc que l'utilisation du portfolio en formation initiale puisse apporter des éléments de réponse à différentes intentions pédagogiques, à savoir : la réflexion et la structuration de la pensée, l'apprentissage et la connaissance de soi comme apprenant, la prise en charge de son propre cheminement professionnel et le développement de sa créativité (UQTR, 2003). Retenons également, qu'en accord avec Goupil (1998), les principes de ce type de portfolio sont en cohérence avec ceux de la formation d'adultes qui privilégient le contrôle de l'apprentissage par l'apprenant(e) lui-même. Ainsi, pour répondre à tous ces critères, les caractéristiques du portfolio de développement professionnel devraient être les suivantes : *« avoir une fonction précise, ce qui suppose que la personne doit savoir exactement pourquoi elle l'élabore ; permettre l'intégration des expériences, par exemple, en formation des enseignants, il devrait comprendre des contenus théoriques et des expériences en classe ; faciliter l'évaluation des acquis expérimentiels et celle reliée à des activités réelles en classe ; être dynamique, c'est-à-dire refléter la croissance professionnelle qui a eu lieu au fil des années et contribuer au développement de l'appropriation de la démarche d'élaboration et d'évaluation du portfolio »* (Barton et Collins, 1993, cité dans Goupil, 1998).

Trois objectifs principaux transcendent ainsi la finalité de développement professionnel de l'outil portfolio utilisé dans le cadre de la formation universitaire luxembourgeoise :

- 1) faciliter le travail d'articulation théorique-pratique et donc, aider les étudiant(e)s dans la réalisation de liens entre la formation reçue à l'université et la pratique de terrain ;
- 2) permettre de témoigner de l'état de leur cheminement ainsi que de leur identité professionnelle émergente et
- 3) être sensibilisé à une pratique d'évaluation authentique innovante qui, comme le suggère la nouvelle loi scolaire luxembourgeoise, sera un jour prochain généralisée dans les classes du Grand-Duché de Luxembourg.

Ce troisième point est particulièrement important pour nous. En effet, depuis la réforme de l'école fondamentale, comme nous l'avons déjà souligné, une nouvelle approche de l'évaluation doit désormais être développée dans toutes les classes primaires du pays et le portfolio y occupe une place importante. L'enseignement et l'évaluation par compétences se tournent davantage sur la documentation, les réalisations et les progrès de l'étudiant. Dans ce contexte, l'approche par portfolio au sein du Bachelor prend tout son sens pour les futurs enseignants. Si le portfolio de développement professionnel est différent dans ses objectifs du portfolio de l'école primaire, la pratique de la collection, les échanges entre pairs et avec l'enseignant-tuteur ainsi que l'aspect réflexif du processus de sélection se retrouvent dans les deux approches. En développant leur propre portfolio, nos étudiants développent et acquièrent des habitudes qu'ils devront favoriser plus tard chez leurs futurs élèves.

Soutenant le développement et la mise en œuvre du portfolio de développement professionnel au sein du Bachelor en Sciences de l'Education de l'Université du Luxembourg, nous avons soumis une proposition de projet de recherche afin d'évaluer la pertinence et l'impact de notre action dans le cadre de ce programme. En effet, pour assurer la qualité du module « Becoming a teacher » (dont le travail avec le portfolio est premier), il est essentiel d'évaluer et d'analyser les effets de l'outil et la démarche réflexive sur les enseignants en formation initiale.

2. L'évaluation du dispositif « portfolio » au sein du Bachelor en Sciences de l'Education de l'Université du Luxembourg

2.1 Description de l'approche par portfolio telle que pratiquée actuellement

Une équipe d'enseignants du BScE a été chargée par la direction de réfléchir sur des pistes d'amélioration du travail avec le portfolio de développement professionnel. L'objectif poursuivi est que ce portfolio devienne un outil susceptible d'amener les étudiants à réfléchir et à réguler leurs propres apprentissages durant les quatre années que dure la formation mais aussi au-delà. Ce groupe de réflexion s'occupe tout à la fois de la structuration et de l'organisation du portfolio des étudiants, de la façon dont ils sont guidés dans leur démarche d'apprentissage par leur tuteur, la construction de séminaires portant spécifiquement sur le développement professionnel et l'utilisation du portfolio et évidemment, du suivi et de l'évaluation du portfolio. Après avoir identifié diverses difficultés durant les premières années d'utilisation du portfolio au sein du Bachelor et conceptualisé des voies possibles d'amélioration tant du processus que de l'outil lui-même, il a semblé important de procéder à l'évaluation du degré d'implémentation de l'outil au sein de la formation (tant au niveau du processus que de l'outil) et d'en mesurer les effets réels sur le développement professionnel de nos étudiants. C'est pourquoi, une proposition de recherche a été soumise au FNR dans le but d'obtenir des subsides pour effectuer ce travail d'amélioration du dispositif.

Les trois buts principaux que nous nous sommes alors fixés sont les suivants :

1. Améliorer les cours (séminaires, tutorats et évaluation) en lien avec l'approche portfolio telle que développée au sein du Bachelor.

De façon à faire face aux difficultés que nous avons pu observer durant les premières années de vie du Bachelor, nous avons réalisé une critique et une remise en question profonde de nos pratiques.

D'abord, il est apparu que beaucoup d'étudiants éprouvaient des difficultés à développer comme à rédiger des réflexions de haut niveau en ce qui concerne leur propre processus d'apprentissage (tant au niveau des apprentissages théoriques effectués à l'université qu'en ce qui concerne les expériences de terrain) (voir aussi le point 2 à ce sujet).

Deuxièmement, nous avons observé un manque d'engagement des étudiants dans l'utilisation du portfolio. Certains d'entre eux ne perçoivent pas, par exemple, l'utilité de réaliser des exercices réflexifs ou encore de mener des tutorats (entre pairs et avec le tuteur).

Troisièmement, une majorité d'étudiants voit un paradoxe entre l'évaluation qui est faite du portfolio (nécessaire dans le cadre d'une validation institutionnelle) et les finalités annoncées

du portfolio. Or, le portfolio peut tout à la fois être un objet de développement professionnel tout en subissant une évaluation institutionnelle. Selon Tucker, Stronge & Gareis (2002), “when portfolio are used for evaluation, there is an expectation for the educator to provide evidence and document competence. Consequently, the portfolio often becomes a collection of the teacher’s finest work and his or her students’ best products. Although all of us would like to be judged by our most impressive work, a narrow focus on strengths seldom assists us in improving our skills and knowledge.” (p.73).

Et quatrièmement, certains collègues n’adhèrent pas toujours complètement aux valeurs, aux finalités défendues par le portfolio et ne comprennent pas en profondeur les mécanismes liés à son utilisation. Ce dernier point pourrait expliquer, dans une certaine mesure, les difficultés précédemment exposées.

Il semble important de travailler sur ces différents aspects parce que le portfolio a été introduit dans la formation avec cette visée particulière d’amélioration du développement professionnel mais aussi en vue de favoriser les apprentissages coopératifs, de documenter et de structurer la communication entre les étudiants et leurs tuteurs (universitaires ou issus du temps de terrain), de modifier les façons traditionnelles de réaliser des synthèses réflexives des apprentissages réalisés et d’engager profondément les étudiants tout comme les enseignants universitaires dans la formation initiale.

2. Améliorer la qualité de la formation initiale en ce qui concerne la capacité des étudiants à devenir des acteurs réflexifs.

Selon Mansvelder-Longayroux, Beijaard and Verloop (2007), pour les enseignants, apprendre à devenir des professionnels réflexifs, constitue une des conditions majeures susceptible de les aider à poursuivre leur développement professionnel une fois engagé sur le terrain. Les auteurs précisent encore que cette habileté réflexive pourrait permettre aux enseignants de faire face au changement engendré par les réformes menées dans le monde de l’éducation. « *The goal of the professional development portfolio can support and emphasize this continuous learning process of a “reflective practitioner”* » (Imhof and Picard, 2009, p. 149).

3. Aider nos étudiants à faire l’expérience d’une pratique d’évaluation innovante qu’ils devront nécessairement mettre en œuvre une fois devenus enseignants.

Depuis la réforme de 2009 en effet, une nouvelle pratique d’évaluation a vu le jour au sein des classes de l’enseignement primaire au sein de laquelle le portfolio occupe une place de choix. Le processus mis en œuvre dans l’approche portfolio vise la compréhension, le suivi et la documentation des acquis ainsi que la progression de l’élève. Dans ce contexte, l’approche portfolio telle que menée au sein du BScE prend tout son sens pour nos étudiants. Si le portfolio de développement professionnel diffère dans ces objectifs du portfolio de l’enseignement primaire, la collection, la sélection, les échanges (entre pairs et avec le tuteur) ainsi que l’approche réflexive sont communes aux deux approches. En développant leur propre portfolio, nos étudiants peuvent développer des stratégies réflexives et faire l’expérience d’une pratique qu’ils devront mettre en œuvre plus tard dans leur classe, avec leurs futures élèves. Selon Adams (1995), « *teacher educators have a unique opportunity to advance assessment reform by exemplifying portfolio assessment models (and other models) in preservice education programs. More importantly, teacher educators can use portfolio to collect more detailed and in depth assessment information than that provided by most traditional forms of assessment* » (p. 570).

Comme l’affirment Lacourse, Hensler and Dezutter (2007, p. 186), il est réellement important d’aller au-delà des difficultés rencontrées et d’élaborer des pistes de solution possibles. Etant donné les potentialités qu’offre le portfolio de développement professionnel en ce qui concerne la formation à la pratique réflexive censée aider les futurs enseignants à être capables d’analyser leurs propres expériences professionnelles et à traduire ces analyses en savoirs professionnels communicables, les auteurs affirment que le portfolio constitue une approche et un outil à développer dans le cadre de la formation initiale.

2.2 Une révision de l'approche par portfolio : révision suggérée

2.2.1 Introduction

L'étude sera conduite à l'Université du Luxembourg et plus précisément au sein du Bachelor en Sciences de l'Education. Pour faciliter la compréhension du contexte dans lequel s'inscrit cette formation, il semble important de préciser certains détails. Le Bachelor est organisé sur quatre années. Le programme de formation inclut des cours de base en éducation, psychologie de l'éducation, méthodologie générale, didactique (mathématiques, sciences, langues, arts, éducation physique et musique), une approche de la recherche en éducation et sciences sociales ainsi que des temps de terrain où les étudiants mettent en pratique les savoirs vus théoriquement. En ce qui concerne cette importante part pratique, les étudiants, placés en dyades, réalisent des stages dans les quatre cycles de l'enseignement fondamental (maternel et primaire) mais aussi en éducation différenciée ou dans les classes préparatoires. Durant cette période de temps de terrain, les étudiants travaillent de concert avec les enseignants en place qui planifient et co-évaluent les compétences professionnelles du jeune avec un tuteur universitaire. Ce tuteur universitaire réalise des visites sur le terrain (au minimum une fois durant les trois semaines de pratique) en vue d'observer et d'évaluer le travail réalisé par l'étudiant. Le Bachelor en Sciences de l'Education accueille 480 étudiants (120 étudiants par année). Le diplôme obtenu permet, après réussite du concours organisé par le Ministère de l'Education luxembourgeoise, aux étudiants d'enseigner au maternel, au primaire, dans les classes préparatoires ou en éducation différenciée. Depuis sa création en 2003, le Bachelor a subi d'importants changements et est encore en cours de révision. Ces changements ont d'inévitables répercussions sur le travail que nous réalisons avec le portfolio. Le paragraphe suivant a pour objectif de préciser quelque peu la façon dont nous avons planifié le travail avec le portfolio de développement professionnel à l'heure actuelle.

2.2.2 Le portfolio de développement professionnel

Pour faciliter le travail avec le portfolio et tracer les lignes directrices qui semblent essentielles à nos yeux, nous proposons à nos étudiants, dès la première année, de développer leur portfolio autour de 4 aspects différents : 1) Qu'est-ce que pour vous l'apprentissage ? ; 2) L'articulation entre la théorie issue des séminaires universitaires et la pratique mobilisée durant le temps de terrain ; 3) Les moments marquants retenus par les étudiants comme ayant influencé leur regard sur l'enseignement et les apprentissages, sur la façon d'envisager le métier d'enseignant, sur la façon de se voir en tant que professionnel et finalement, 4) Les perspectives des étudiants (pour le semestre à venir, pour leur métier futur, ...). Ces quatre éléments constituent le point de départ de la construction du portfolio et la façon de les envisager peut être différente pour chaque étudiant. Les activités réflexives proposées dans le cadre des séminaires dédiés au portfolio ont pour objectif d'aider les étudiants à définir leur propre structure, à développer leurs propres idées concernant la façon de collecter leurs propres traces mais aussi de les guider dans l'apprentissage de l'écriture réflexive.

Les finalités ainsi poursuivies par les séminaires "Becoming a teacher" sont organisées autour de trois composantes dynamiques et interreliées qui sont :

1. Le développement professionnel : nous cherchons à organiser des tâches susceptibles d'améliorer les skills, attitudes, connaissances et stratégies relatives au développement professionnel.
2. Le développement académique : des tâches d'écriture réflexives sont organisées en vue d'améliorer la réflexivité des étudiants. Différentes activités sont proposées de façon à développer différents types d'écrits (de l'écriture biographique à l'écriture scientifique). A l'instar de Vanhulle (2005), nous pensons que « reflective writing can function as another mediation tool, extending social interactions and debates so that real transformations can be consolidated » (p. 292).

3. Le développement personnel : nous essayons d'organiser des activités réflexives susceptibles d'amener les étudiants d'un statut d'étudiant (lycéen) à un statut de professionnel de l'éducation réflexif.

Durant les quatre premières années de la formation, ces trois composantes sont également évaluées à travers des transpositions esthétiques dont l'ultime aboutissement est, en quatrième année, la réalisation d'une exposition esthétique. Ces transformations esthétiques ont pour objectif de montrer et d'apprendre aux étudiants combien il est possible de diversifier et d'enrichir ses points de vue sur un objet d'apprentissage. Les étudiants apprennent ainsi peu à peu à se décentrer grâce à d'autres voies, d'autres supports de réflexion.

Depuis l'année académique 2009-2010, nous avons organisé les séminaires « Becoming a teacher » de la façon suivante :

1^{ère} année:

Six séminaires et deux tutorats individuels sont organisés.

Durant ces séminaires, nous proposons aux étudiants de travailler sur :

- L'introduction du concept de portfolio et le portfolio comme outil de développement professionnel.
- Quelle structure les étudiants peuvent-ils adopter et quels types d'artefacts peuvent-ils collecter pour leur portfolio ?
- Les compétences professionnelles et la profession d'enseignant : introduction.
- De l'écriture anecdotique à l'écriture scientifique.

2^{ème} année:

Durant la deuxième année, un semestre est dédié à la mobilité. Chaque étudiant doit partir dans une université étrangère pour le semestre de son choix. Les étudiants en mobilité doivent bien évidemment poursuivre le travail avec le portfolio et un tutorat individuel ou en petits groupes est organisé une fois revenus au Luxembourg.

Pour les étudiants qui restent à l'Université du Luxembourg, un séminaire réflexif ainsi qu'un tutorat sont organisés. Par ailleurs, deux séminaires portant sur la transformation esthétique sont également organisés.

Durant ces trois séminaires, les étudiants travaillent sur :

- Les compétences professionnelles : qu'est-ce qu'une compétence et qu'entend-on par compétences professionnelles ? Que disent les documents officiels luxembourgeois sur cet aspect ?
- Le recueil d'artefacts en lien avec ces compétences professionnelles.
- La sélection d'objets pour le portfolio.
- L'identification et processus réflexif centré sur les artefacts à travers différents moyens d'expression (introduction à la transformation esthétique).

3^{ème} année:

Durant la troisième année, deux séminaires réflexifs et deux tutorats individuels sont proposés. Quatre séminaires sont également organisés et portent sur la transposition esthétique. Nous proposons enfin quatre séminaires relatifs aux méthodes pédagogiques innovantes (Reggio, Freinet et Latour). L'évaluation des skills professionnels (ainsi que l'auto-évaluation) comme l'évaluation en général sont analysés. Un focus particulier est mis sur l'évaluation formative et narrative.

L'approche esthétique est approfondie. Les étudiants apprennent à utiliser des approches multimodales dans la compréhension des artefacts utilisés pour leur biographie esthétique. Ils essaient également d'implémenter ces concepts dans leur temps de terrain en vue de créer des liens avec leur portfolio.

4^{ème} année:

Durant la dernière année, deux séminaires réflexifs et deux tutorats individuels sont proposés. Le second tutorat individuel porte sur l'évaluation finale des quatre années de développement professionnel. Trois séminaires sont offerts en vue de préparer l'exposition esthétique finale qui doit rendre compte du développement professionnel et personnel des étudiants au cours des quatre années de formation.

Durant les séminaires, nous proposons aux étudiants de travailler sur deux aspects distincts :

- Les étudiants doivent réfléchir sur une transposition didactique du portfolio à l'école primaire. Quelle méthodologie semble la plus appropriée pour encourager les élèves à développer une attitude réflexive en regard de leur propre processus d'apprentissage ?
- Les étudiants doivent préparer une exposition esthétique en lien avec leur développement professionnel et personnel.

Tous les éléments du contexte nous ont aidés à définir les questions de recherche, l'échantillon et ont aussi guidé le choix de notre approche de recherche (méthodologie et sélection des instruments les plus appropriés).

2.2.3 Questions de recherche

Deux composantes sont principalement ciblées dans notre recherche et feront l'objet d'un recueil de données.

1. L'influence du portfolio sur le développement professionnel

- Ce portfolio a-t-il un réel impact sur le développement professionnel des étudiants lors de leur formation initiale ?
- Quels sont exactement les comportements et les perceptions qui seraient significativement affectées par cette approche basée sur le portfolio ?
- Quels éléments seraient-ils susceptibles de mettre ces aspects en évidence et l'influence est-elle de longue durée ?

2. Le degré d'acceptation de l'instrument lui-même tant du point de vue que des étudiants que des enseignants universitaires.

- Quelles sont les représentations de tous les acteurs du Bachelor en Sciences de l'Education (élèves, professeurs, direction et équipe de direction du BScE) sur l'approche basée sur le portfolio : sa mise en œuvre, son utilisation, son évaluation, son développement ?

Les réponses nous aideront à clarifier le cœur de notre problème :

- Quels sont les aspects de l'approche portfolio qui devraient être modifiés, intégrés au dispositif ou encore abandonnés afin d'assurer la meilleure qualité du module « Becoming a teacher » au sein du Bachelor en Sciences de l'Education ?

2.2.4 Sampling

- Tous les étudiants inscrits au Bachelor en Sciences de l'Education; soit 120 étudiants par an et un total de 480 élèves pour les quatre ans de formation.
- Les professeurs d'université et de l'équipe de direction

2.2.5 Instruments et recueil de données

Après avoir recueilli les observations des intervenants concernant l'approche portfolio, son objet, structure, son organisation et son suivi au moyen d'un questionnaire, nous utiliserons ces données pour élaborer des scénarios alternatifs liés à la mise en œuvre du portfolio et de son fonctionnement. Ces scénarios seront soumis à l'évaluation par le biais de groupes de discussion.

1) Le questionnaire (seulement pour les étudiants)

L'objectif que nous poursuivons par la mise en œuvre d'un questionnement des étudiants vise à évaluer le degré d'intégration du portfolio au sein du BScE mais aussi d'appropriation du dispositif par les étudiants. A l'instar de Coen et Schumacher (2006), nous faisons référence au modèle d'intégration d'une innovation en trois niveaux (Depover et Strebelle, 1997, cité dans Coen et Schumacher, 2006) pour juger de ce processus :

- a. **L'adoption** : c'est la décision de changer quelque chose dans sa pratique par conviction personnelle ou sous une pression externe qui peut s'exercer au départ du microsystème (ici en

l'occurrence les demandes du Bachelor matérialisées sous la forme des consignes générales, des tutorats et des séminaires).

- b. **L'implantation** : c'est la concrétisation de la volonté affirmée, lors de la phase d'adoption, de s'engager dans un processus conduisant à une modification des pratiques classiques et habituellement reconnues comme légitimes.
- c. **La routinisation** ; c'est le recours aux nouvelles pratiques qui s'opère sur une base régulière et intégrée aux activités habituelles sans manifestation d'une intervention externe.

Le tableau ci-dessous reprend les caractéristiques essentielles des différents niveaux présentées selon trois aspects : la structure institutionnelle, la forme et le fond. Au niveau de la structure institutionnelle, on retrouve : la résistance au changement, les caractéristiques psychologiques et les caractéristiques sociales. En ce qui concerne le fond, nous avons quelque peu adapté la classification de Michaud et Alin pour rendre compte des caractéristiques-clés de l'évolution des écrits des étudiants au sein du portfolio.

Table 1 : Description des caractéristiques évaluées en lien avec le modèle de Depover et Strebelle (1997)

	L'adoption	L'implantation	La routinisation
LA STRUCTURE INSTITUTIONNELLE			
La résistance au changement	Usage limité du PF Tâtonnement personnel, essai-erreur Recours au PF au cours de périodes fortement accompagnées, guidées et structurées Pas de prises d'initiatives	Utilisation régulière du PF Commencement de l'autonomie Manque de décentration Travaux effectués au cas par cas	Le PF devient un outil au service du développement professionnel Décentration et réflexivité Auto-évaluation et autorégulation du dispositif Inscription du PF dans le long terme
Le psychologique	Sentiment d'incertitude (frustration) Enthousiasme Autoquestionnement de l'étudiant vers le changement Décision de s'approprier les principes du PF	Investissement personnel Motivation, mobilisation Efforts nets de la part des étudiants dans l'appropriation du PF	Clairvoyance Investissement de temps moindre
Le social	Grande dépendance vis-à-vis des formateurs Réseau social pas ou peu construit	Dépendance vis-à-vis des formateurs Tissage progressif d'un réseau social entre pairs, avec les tuteurs et avec les formateurs de terrain	Indépendance vis-à-vis des formateurs
LA FORME			
L'aspect technique	Premiers essais Tentatives de compréhension de ce que pourrait être un PF au niveau de la forme	Centration sur les apports technologiques du PF Centration sur la maîtrise du PF Le PF est une fin en soi	Maîtrise de l'outil Le PF n'est plus une fin en soi mais un moyen.
LE FOND			
Les indices réflexifs et de scientificité des écrits			
Type d'écrits	Type 0 Romancé	Type 1 Anecdotique	Type 2 Description commentée
Degré de réflexivité	Non réflexif	Non réflexif à peu réflexif	Peu réflexif
Indicateurs sur le fond	Structure narrative du récit. Description de situation sur la base de faits réels qui ont été adaptés, romancés par l'auteur.	Description de situation. Utilisation de « recette ». La solution s'applique dans réflexion.	Les choix sont justifiés sur des jugements personnels sans véritable analyse. Commentaires personnels.
Indicateurs sur la forme	Structure et temporalité Il raconte un fait, un événement en situant son déroulement dans le temps et dans l'espace. Il en retrace les étapes et en fixe la durée. Le texte est souvent entrecoupé de passages descriptifs, explicatifs ou argumentatifs.	Structure Juxtaposition de données (style protocolaire, proposition infinitive, suite d'énumération sans lien). Temporalité Peu ou pas d'ajout de locution	Structure Dépôt émotionnel et présentation neutre sans analyse argumentative des faits. L'auteur fait part de sentiments et d'émotions introduits souvent par une structure entre « ... ». Temporalité Discours impersonnel et forme passive.
Type d'écrits	Type 0	Type 1	Type 2
Place du « je » dans le récit	Peut être importante	Faible	Important
			Type 3 Peut être importante « Je » en « nous » éventuellement Connecteurs argumentatifs
			Type 4 La prise de distance n'implique pas le « je » obligatoirement.

Pour évaluer le niveau d'intégration et d'appropriation du portfolio sur l'échelle présentée ci-dessous allant de l'adoption à la routinisation en passant par l'implantation, il semble que la construction de situations en lien avec la structure institutionnelle, la forme et le fond soit le mode d'évaluation le plus efficace.

Dans chacune des vignettes soumises aux étudiants, les différents aspects en lien avec la structure institutionnelle, la forme et le fond seraient décrits et proposés selon un ordre croissant ou décroissant allant de l'adoption à la routinisation ou de la routinisation à l'adoption (l'implantation est toujours présentée au milieu de l'échelle). Après lecture des trois situations, le répondant doit se situer – sur une échelle de cinq points – en estimant son degré de proximité par rapport aux trois situations présentées (Coen et Schumacher, 2006, p. 13).

Voici un exemple de vignette qu'on pourrait leur proposer. L'exemple a été rédigé sur la base du point : « Résistance au changement ».

SITUATION 1

« Mon portfolio ? Oui, je l'utilise de façon régulière surtout pour collecter des documents ou des traces qui me paraissent intéressantes. Pour moi, il s'agit d'un dossier qui m'aide à organiser et à conserver des éléments utiles en lien avec mon futur métier. Toutefois, les échéances établies en classe ont été une aide précieuse pour assurer une certaine continuité dans l'élaboration du portfolio. »

« Le portfolio permet de faire un retour sur notre cheminement. C'est un document qui nous représente bien et qui sera très utile en entrevue. C'est un outil très important pour moi parce qu'il me permet de faire un retour sur ce que je suis et sur mon cheminement professionnel jusqu'à présent. Bref, je peux dire que j'ai travaillé très fort à l'élaboration de mon portfolio, de façon régulière, afin qu'il reflète qui je suis par mon organisation et par ma capacité de réfléchir et de me remettre en question. C'est un outil très utile sur le plan professionnel, alors je l'ai construit de manière à pouvoir y apporter des changements au fur et à mesure que j'évoluerai en enseignement. »

☐

Je me sens proche de cette situation

☐

Je me situe entre ces deux situations

☐

Je me sens proche de cette situation

« J'utilise vraiment très peu mon portfolio. J'y pense et je travaille avec lui uniquement lorsque je sais que je vais bientôt avoir un tutorat. Il faut vraiment qu'il y ait une échéance proche ou alors une consigne précise donnée dans un séminaire pour que je rédige des textes ou que je réfléchisse à des traces possibles à intégrer à mon dossier. Pour moi, le portfolio ne représente rien d'autre qu'un travail universitaire parmi d'autres. »

De laquelle de ces trois situations vous sentez-vous le plus proche ?

Remarques :

.....

.....

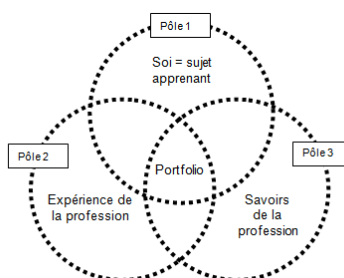
.....

A ces caractéristiques liées au degré d'implantation de l'innovation et d'appropriation du média par les étudiants, nous devons également juger de l'atteinte des finalités visées par le portfolio. Nous devons par conséquent nécessairement tenir compte des caractéristiques inhérentes au contexte de l'implantation, à savoir la formation elle-même. Le Bachelor en Sciences de l'Education poursuit la finalité de former des professionnels de l'enseignement fondamental. Pour ce faire, nous avons retenu le portfolio de développement professionnel. Le portfolio comme outil de développement professionnel est au croisement de trois types de foyers d'attention à savoir :

Pôle 1 : Le développement personnel de l'étudiant,

Pôle 2 : Le développement des compétences liées à la profession au départ de la pratique de terrain ;

Pôle 3 : Les savoirs liés à la profession elle-même.



Ces trois rubriques ont une influence sur les valeurs à la base de l'approche portfolio, sur les méthodes privilégiées dans les séminaires « Becoming a teacher » ainsi que sur les objectifs ciblés pour l'approche portfolio. Ces trois pôles déterminent par conséquent les sujets couverts par le survey :

- 1) Les finalités éducatives de l'outil
- 2) Les compétences transversales visées par la formation
- 3) Les compétences propres au champ de l'éducation
- 4) Le soi comme apprenant
- 5) La gestion quotidienne de l'outil

En vue d'évaluer le degré d'appropriation de ces aspects par les étudiants et la place qui leur est accordée dans l'approche portfolio, les sujets devront répondre à des items sur la base d'une échelle de fréquence ou d'accord. Il existe différents instruments scientifiques validés pour évaluer ces aspects comme, par exemple :

- Teacher Survey : Perceived Value of Portfolios (Tucker, Stronge and Gareis, 2002, p. 173-174)
- The assessment of portfolio of professional development, reflective analysis and support to professional development (Lacourse, Hensler and Dezutter, 2007, p. 198-199).
- Implementing Portfolios in a Teacher Education Program (Stolle, Goerss & Watkins, 2005, p. 34-36).

2) Les focus groups (pour les étudiants et les acteurs principaux du BScE)

Un groupe de discussion est une méthodologie de recherche qualitative. Différentes personnes ayant une caractéristique commune (culturelle, sociales, idéologique...) sont réunies dans un même groupe.

Nous recueillerons leur réponse et les attitudes à l'égard l'approche portfolio telle qu'ils la connaissent au sein du BScE.

Les axes définis pour le questionnaire seront utilisés pour développer les entretiens semi-structurés nécessaires à la réalisation des entrevues au sein des groupes de discussion.

À la fin de la collecte de données, les données devraient nous permettre d'élaborer des scénarios possibles d'utilisation et de gestion du portfolio au sein du BScE.

3. Conclusion

La présente communication avait pour objectifs de présenter le contexte d'utilisation du portfolio de développement professionnel au sein du Bachelor en Sciences de l'Education de l'Université du Luxembourg, les difficultés rencontrées jusqu'ici et les perspectives de changement.

La principale perspective a consisté en la rédaction d'un projet de recherche que nous avons soumis au FNR dans le courant du mois de mai dernier. L'obtention ou non de ce budget de recherche sera connue dans le courant du mois de septembre prochain.

C'est une synthèse des principaux aspects décrits dans cette proposition que nous avons proposés ici : théorie, description du contexte actuelle et difficultés rencontrées, questions de recherche, échantillon et méthodologie.

Si nous obtenons ce financement, les résultats de cette recherche devraient nous permettre d'améliorer le dispositif actuel et l'approche que nous privilégions en ce moment avec nos étudiants. Il s'agit pour nous d'améliorer encore le processus d'apprentissage, de réflexivité et d'évaluation mené jusqu'ici et partant, d'assurer une qualité accrue de la formation elle-même.

4. Bibliographie

- Adams, T. L. (1995). A paradigm for portfolio assessment in teacher education, *Education*, 115(4), 568-570.
- Coen, P-F & Schumacher, J. (2006). Construction d'un outil pour évaluer le degré d'intégration des TIC dans l'enseignement, *Revue Internationale des Technologies en Pédagogie Universitaire*, 3(3), 7-17.
- Darling, L. F. (2001). Portfolio as practice: the narrative of emerging teachers, *Teaching and Teacher education*, 17, 107-121.
- Déum, M. & Vanhulle, S. (2005). Portfolio et démarches socio-réflexives en enseignement et formation. *Les cahiers de la section des Sciences de l'Education*. Université de Genève. Faculté de psychologie et des sciences de l'éducation, 116, 150 pages.
- Foster, B. R., Walker, M. L. & Song, K. H. (2007). *A beginning teaching portfolio handbook. Documenting and reflecting on your professional growth and abilities*, New Jersey: Pearson Prentice hall.
- Gearhart, M. & Osmundson, E. (2009). Assessment Portfolios as Opportunities for Teacher Learning, *Educational Assessment*, 14, 1-24.
- Goupil, G. (1998). Le portfolio : vers une pratique réflexive de l'enseignement, *Vie Pédagogique*, 107, 4 pages.
- Harland, T. (2005). Developing a portfolio to promote authentic enquiry in a teacher education, *Teaching in Higher Education*, 10(3), 327-337.
- Imhof, M. & Picard, C. (2009). Views on using portfolio in teacher education, *Teaching and Teacher education*, 25, 149-154.
- Jorro, A. (2005). Réflexivité et auto-évaluation dans les pratiques enseignantes; *Mesure et évaluation en éducation*, 27(2), 33-47.
- Klenowski, V. (2000). Portfolios: promoting teaching, *Assessment in Education*, 7(2), 215-236.
- Lacourse, F., Hensler, H. & Dezutter, O. (2007). Le portfolio de compétences professionnelles. Un miroir singulier pour un regard évaluatif sur la réflexion professionnelle en enseignement, in Jorro, A. (Ed.). *Evaluation et développement professionnel* (pp. 185-199), Paris : L'Harmattan.
- Mansvelder-Longayroux, D. D., Beijaard, D. & Verloop, N. (2007). The portfolio as a tool for stimulating reflection by student teachers, *Teaching and Teacher Education*, 23, 47-62.
- Michaud, C. et Alin, C. (2008). *Processus portfolio – La comparaison du développement de l'identité professionnelle de deux groupes de stagiaires à l'IUFM : les stagiaires en formation initiale et les stagiaires en situation*, Actes du colloque organisé par la conférence des directeurs d'IUFM « La formation des enseignants en Europe : approche comparative », Paris, 9-10 Décembre 2008.
- Parsons, M. & Stephenson, M. (2005). Developing reflexive practice in student teachers: collaboration and critical partnerships, *Teachers and Teaching: theory and practice*, 11(1), 95-116.
- Perrenoud, P. (1999a). *Dix nouvelles compétences pour enseigner. Invitation au voyage*. Paris : ESF.

- Perrenoud, P. (1999b). *Former les enseignants dans des contextes sociaux mouvants : pratique réflexive et implication critique*, Actes du congrès de l'Anped (Association nationale pour la recherche en éducation), Caxambu (Brasil), 26-30 septembre 1999.
- Perrenoud, P. (2001). De la pratique réflexive au travail sur l'habitus, *Recherche et Formation*, 36, 131-162.
- Perrenoud, P. (2005). Assumer une identité réflexive, *Educateur*, 2, 30-33.
- Stolle, C., Goerss, B. & Watkins, M. (2005). Implementing Portfolios in a Teacher Education program, *Issues in Teacher Education*, 14(2), 25-43.
- UQTR (2003). *Formation pratique et stage. Le portfolio*. Consulté en décembre 2008, sur le site de l'Université du Québec à Trois-Rivières, http://www2.uqtr.ca/hee/site_1/index.php?no_fiche=96.
- Tucker, P. D., Stronge, J. H. & Gareis, C. R. (2002). *Handbook on teacher portfolios for evaluation and professional development*, Larchmont, NY: Eye on Education.
- Vanhulles, S., Déum, M., Mattar, C. & Schillings, A. (2004). Des pratiques réflexives en formation initiale, pour une intégration optimale de la théorie et de la pratique, *Caractères*, 15(2), 19-29.
- Vanhulle, S. (2000). La littératie dans la formation des futurs instituteurs : interactions sociales et portfolio personnel. *Les cahiers du Service de Pédagogie expérimentale. Université de Liège. Faculté de Psychologie et des Sciences de l'Éducation*, 47-66.
- Vanhulle, S. (2005). How future teachers develop professional knowledge through reflective writing in a dialog frame, *Educational studies in language and Literature*, 5, 287-314.

**EVALUATION DU PROCESSUS DE MISE EN PLACE DE DEUX OUTILS
PEDAGOGIQUES (PORTFOLIO ET GRILLE D'EVALUATION DES STAGES)
EN LIEN AVEC LE REFERENTIEL DE COMPETENCES INTEGRE EN
KINESITHERAPIE.**

Helyett Wardavoir*, Karin Van Loon, Catherine Romanus*, Dominique Peeters**,
Patrick Parmentier*, Florence Parent*****

** Haute Ecole Libre de Bruxelles I. Prigogine, Bruxelles, Belgique –
helyett.wardavoir@helb-prigogine.be*

*** Haute Ecole P.H. Spaak, catégorie paramédicale, section kinésithérapie, Bruxelles,
Belgique - vanloon@he-spaak.be*

**** Département d'Epidémiologie et de Promotion de la Santé, Ecole de Santé publique,
Université Libre de Bruxelles (ULB), Bruxelles, Belgique
ARE@Santé (Association pour le Renforcement de l'Enseignement et de l'Apprentissage
en Santé), Bruxelles, Belgique. florence.parent@ulb.ac.be*

Mots-clés : *compétences, développement d'outils, évaluation, approche participative*

Résumé.

Dans le cadre de l'implémentation d'un référentiel de compétences intégré (RCi) en kinésithérapie, notre objectif est de développer des outils en termes d'ingénierie des compétences et de mettre en place des processus en termes d'ingénierie de projets. Ceux-ci répondent à la finalité d'évoluer vers un changement des pratiques d'apprentissage pour les étudiants et d'évaluation pour les superviseurs et maitres de stages. Les outils et leurs dispositifs associés visent le développement des compétences réflexives et métacognitives pour le portfolio et de la plupart des compétences du RCi pour la grille d'évaluation en stage. Conscients du bien fondé de l'intégration des logiques pédagogiques et professionnelles dans l'approche socioconstructiviste d'un curriculum de formation professionnalisante, nous avons privilégié une approche participative avec l'organisation formelle, dans notre méthodologie, de feed-back des acteurs concernés

1. Introduction

Dans le cadre d'une réforme générale de l'enseignement dans les Hautes Ecoles en Communauté française de Belgique ayant pour objectif l'adaptation de la formation aux standards internationaux (processus de Bologne, de Copenhague et Cadre européen de certification) en référence à une approche pédagogique basée sur une logique de compétences, nos institutions ont

saisi l'opportunité d'entrer dans cette réforme du curriculum par la construction d'un référentiel de compétences intégré (RCi).

Ce projet est initié en 2008 par le groupe des directeurs des institutions partenaires, en offrant un accompagnement méthodologique (AM) mené par Florence Parent (ESP-ULB) à une équipe de pilotage (EP), constituée à parts égales d'enseignants des institutions partenaires du projet.

Le cadre de référence pédagogique s'appuie sur la méthodologie de Parent (Parent et al, 2011) et s'inscrit dans un paradigme socioconstructiviste (Jonnaert, 2009). Selon une logique de compétences, les concepts de « situation » et « d'apprentissage en contexte » sont prépondérants. S'inscrire dans une perspective socioconstructiviste, c'est accorder une place centrale à la dimension sociale dans les apprentissages, en valorisant la co-construction en contexte, des connaissances et des compétences. L'apprentissage est alors davantage considéré comme le produit d'activités sociocognitives liées aux échanges didactiques enseignant-élèves, élèves-élèves, mais aussi enseignants-enseignants, enseignants-experts professionnels.

Notre démarche repose sur une pédagogie de l'intégration adaptée au secteur de la santé et plus précisément l'approche par compétences intégrée (APCi). Elle est construite et consolidée en constante interaction avec tous les acteurs concernés selon un processus de consensus. La construction d'un curriculum en santé nécessite également d'être en adéquation avec les besoins des professionnels et de la société, raison pour laquelle l'EP s'est entourée, tout au long du processus, de nombreux acteurs issus du monde professionnel.

Ainsi pour cette première étape, qu'a été la construction du référentiel de compétences, la représentativité des acteurs a été élargie à l'ensemble des situations professionnelles rencontrées sur le terrain de la pratique ainsi qu'aux acteurs stratégiques et au secteur de la santé.

Après cette étape de traduction didactique, qui a fait notamment appel au cadre conceptuel que représente la didactique professionnelle (Pastré et al, 2006) et à une taxonomie spécifique (Parent et al. 2008, Parent et al, 2012), ayant permis d'élaborer notre référentiel, nous avons entamé la phase de transposition didactique qu'on pourrait nommer également transposition pédagogique avec l'élaboration des référentiels de formation (RFi) et d'évaluation intégrés (REi). Dans le même temps nous avons implémenté, sous forme expérimentale, les premières réformes que sous-tend l'approche par compétences. Le RFi issu du RCi, permet de questionner le programme de formation, d'entrer dans le développement de nouveaux dispositifs et d'organiser de façon cohérente et transparente la formation (Parent, 2008).

Par ailleurs, les résultats observés dans différentes institutions de formation, montre que le référentiel joue bien un rôle sur le pilotage de la formation. Au départ de son rôle de cadre de référence, il peut produire un effet sur les pratiques individuelles des enseignants (Postiaux, 2009).

En tant qu'EP nous constatons aussi le rôle inducteur du référentiel sur nos motivations au changement.

Nous situant dans une approche programme, c. à d. dans une vision plus systémique du curriculum, nous privilégions la représentativité exhaustive de l'ensemble du corps enseignant dans la mise en place de nouveaux dispositifs. Dans le cadre précis, de l'implémentation des dispositifs de formation et d'évaluation liés aux stages (grille d'évaluation de stage et portfolio), nous y impliquons en partie les maîtres de stages (professionnels encadrant les étudiants dans les lieux de stages).

Vu à cette plus large échelle, le processus innovant dans lequel nous sommes engagés ne va pas sans engendrer auprès des acteurs concernés, certaines peurs et réticences au changement (contenus et dispositifs de formation et d'évaluation). Conscients de leur impact, nous nous

sommes d'emblée inscrits dans une approche participative, impliquant un maximum d'acteurs, de façon à stimuler l'appropriation progressive des concepts et la remise en question des pratiques liées à ce nouveau cadre de référence.

Avant d'entrer plus avant dans la description des outils de formation et d'évaluation mis en place situons plus précisément notre référentiel de compétences intégré (RCi).

Dans l'APC, la compétence se développe dans l'action et en situation. De la notion de compétence, émerge aussitôt celle de situation : « *Les compétences ne peuvent se définir qu'en fonction de situations, elles sont donc tout autant situées que les connaissances dans un contexte social et physique* ». Une situation est un « *tout contextuel* » assimilable à un champ d'activités définies par les relations interactives d'une personne avec son environnement. La situation n'existe pas en dehors de la personne et de ce qu'elle en fait (Jonnaert et al, 2006).

Donc, *"pour amener les élèves à mobiliser leurs savoirs, il faut les placer dans des situations appropriées"* (Scallon, 2004).

Notre référentiel de compétences intégré¹ définit :

- Dix compétences, issues des activités réflexives, cognitives, métacognitives, psychoaffectives, sociales, gestuelles/procédures, déclinées en capacités ;
- Les critères de qualité et des exemples d'indicateurs et/ou de rubriques permettant l'évaluation ;
- Les situations d'intégration en adéquation avec les situations professionnelles rencontrées et regroupées en cinq familles : 'consultations', 'institutions de soins', 'milieux particuliers', 'recherche' 'gestion/management'.
- Un cadre de référence professionnel et un cadre de référence pédagogique.

2. Situation de la problématique et des objectifs

L'adoption de référentiels de compétences intégrés (RCi) par les Hautes Ecoles belges représente une nouvelle orientation qui peut être à l'origine de craintes et résistances par manque de compréhension des tenants et aboutissants qu'implique un changement de paradigme pédagogique. Dans ce contexte, il ne faut pas négliger non plus les raisons sociales et politiques relatives aux appropriations utilitaristes des logiques de compétences, d'où la nécessité de balises méthodologiques favorisant un développement de curriculum par compétences, dans une vision non réductrice, globale et intégratrice.

Ce- inclus le questionnement de la place des savoirs, la remise en question des rapports et rôles des acteurs concernés : étudiants, enseignants, maîtres de stages, mais aussi le souhait d'optimiser le partenariat entre les institutions de formation et avec les terrains de stages.

Dans le processus engagé dans nos deux Hautes Ecoles en kinésithérapie et conscients de tous ces aspects, dont les enjeux managériaux, c'est en priorisant l'appropriation du RCi par les acteurs

¹ Référentiel de compétences intégré en kinésithérapie consultable sur : http://www.helb-prigogine.be/index.php?option=com_content&view=article&id=1280&Itemid=425&limitstart=5

concernés dans les activités d'intégration professionnelle que nous avons développé deux nouveaux outils en adéquation avec un apprentissage par compétences.

Dans le cadre d'une formation professionnalisante, l'implémentation d'un RCi intègre deux logiques d'acteurs différentes : celle de l'équipe pédagogique et celle des professionnels de terrain. La problématique est donc relative à l'implication des différents acteurs conjointement à l'adéquation des outils de formation et d'évaluation en situation d'activité d'intégration professionnelle dans un apprentissage par compétences et à l'évaluation de la mise en place de ces nouveaux outils.

En référence au choix du paradigme socioconstructiviste, notre but est de sensibiliser les acteurs concernés, à l'intérêt des pédagogies actives et donc aux processus d'apprentissage davantage centrés sur l'étudiant (Jouquan, 2003). Nous y voyons une opportunité de transformer les dispositifs d'apprentissages et les outils d'évaluation actuels.

Ceux-ci répondent à la finalité d'évoluer vers un changement des pratiques d'apprentissage pour les étudiants (activités réflexives et métacognitives) et d'évaluation pour les superviseurs et maitres de stages (évaluation formative partagée, qualitative et critériée).

Notre objectif est donc de développer des outils en termes d'ingénierie des compétences et de mettre en place des processus en termes d'ingénierie de projets dans une perspective d'intégration.

3. Description des outils et dispositifs associés

L'élaboration des outils et de leur dispositif associé se base sur une réflexion quant à leur forme et contenu de manière à soutenir les objectifs de développement des compétences réflexives et métacognitives pour le portfolio et de la plupart des compétences du RCi pour la grille d'évaluation en stage.

Considérant que la compétence se développe au cours d'un processus dynamique d'apprentissage en situation complexe, il a semblé évident d'intégrer de nouveaux dispositifs tels que l'auto évaluation et l'évaluation formative afin de permettre à l'étudiant d'adapter ses propres processus en cours d'apprentissage (Rey et al, 2003).

Les outils et dispositifs mis en place devraient idéalement revaloriser la fonction formatrice des stages, vue sous l'angle d'une approche par compétences dans un cadre de référence socioconstructiviste.

En termes d'outils, nous avons refondu la grille d'évaluation afin de la rendre concordante au RCi. Elle permet une évaluation qualitative des principales compétences et capacités par une série d'indicateurs à titre d'exemples (*Figure 1 : Extrait de la grille d'évaluation en stage*).

Optimiser les fonctions physiologiques <i>Activités gestuelles, opératives et procédurales</i>	<i>Appliquer des techniques et des procédures spécifiques</i>						Utiliser les techniques en connaissance de cause (indications, contre-indications,...)
							Réaliser les gestes techniques avec précision, dextérité
							Respecter les protocoles, les procédures et les recommandations existantes
							Autres
	SO	1	2	3	4	5	
	<i>Mesurer des paramètres biométriques, physiologiques en utilisant des tests, des outils</i>						Mesurer avec la plus grande précision
							Se référer à des normes et/ou des standards
							Assurer l'innocuité des mesures
							Autres
	SO	1	2	3	4	5	
	<i>Utiliser la technologie</i>						Respecter les règles d'utilisation, de sécurité, d'ergonomie et d'hygiène
							Diversifier les techniques et les outils utilisés
							S'adapter aux possibilités techniques de l'environnement de travail
							Autres
	SO	1	2	3	4	5	

Figure 1 : Extrait de la grille d'évaluation en stage. L'évaluateur situe le niveau de maîtrise des capacités en référence aux indicateurs. SO = sans objet ; 1 à 5 = niveaux qualitatifs de maîtrise.

Quant au dispositif associé, la grille est proposée en auto évaluation à l'étudiant et en évaluation partagée (co-évaluation) avec le superviseur et le maître de stage, à titre formatif en évaluation intermédiaire et à titre certificatif en fin de stage.

Toujours dans le cadre de la revalorisation de la formation liée au stage, nous avons opté pour la mise en place d'un portfolio comme outil de « *réflexion appartenant à l'étudiant qui lui permet de cumuler des preuves - ou du moins des traces – des apprentissages qu'il a réalisés dans divers contextes (cours, stages et toutes autres expériences professionnelles et personnelles...) et de les organiser de manière à témoigner du développement de ses compétences pendant sa formation* » (Belanger, 2008).

L'objectif de ce portfolio (*Figure 2 : Eléments constitutifs du portfolio*) est de développer la réflexivité et l'autonomie de l'étudiant en lien avec la construction de ses compétences professionnelles.

Présentation personnelle	Traces d'apprentissage	Evaluation	Bilan
Sous forme d'un cv retraçant les éléments pertinents (en rapport avec la kinésithérapie) : - du parcours personnel - de la vision du métier - du projet professionnel	Au départ de l'analyse d'une situation clinique (3 à 5 traces)	Insertion des grilles d'évaluation de stages	Bilan écrit des compétences développées en fin de B3
	Rapport de GEC (groupe d'enseignement clinique)	Analyse des forces et faiblesses en lien avec les compétences au cours de chacun des stages	Bilan écrit des compétences développées en fin de M1
	Rapport au médecin		

Figure 2 : Eléments constitutifs du portfolio.

3.1 Sur le plan méthodologique

Pour les deux outils, nous avons procédé selon les cinq étapes suivantes :

- *Elaboration de l'outil ;*
- *Passation de l'outil aux utilisateurs ;*
- *Phase test ;*
- *Analyse des données ;*
- *Réajustement de l'outil.*

Des dispositifs d'accompagnement des utilisateurs ont été prévus tout au long du processus.

Et nous avons respecté les objectifs suivants :

- *Privilégier l'approche participative :* les outils et dispositifs ont été construits par l'équipe pédagogique et plus particulièrement par l'équipe d'enseignants cliniques étant donné qu'il s'agit du cadre spécifique de la formation en stage.
- *Favoriser l'appropriation des concepts, outils et dispositifs par les professionnels qui encadrent les étudiants dans les lieux de stage :* l'implémentation des outils et dispositifs a fait l'objet d'un accompagnement auprès des professionnels sous forme de réunions d'information, de modes d'emploi écrits explicités par les superviseurs de stages (enseignants de l'école) lors de leur visite en stage.
- *Favoriser l'appropriation des concepts, outils et dispositifs par les étudiants :* présentation du référentiel de compétences, accompagnement à l'utilisation de celui-ci lors des autoévaluations et bilans du portfolio, accompagnement du portfolio par les enseignants cliniques.
- *Intégrer les acteurs concernés dans l'évaluation des outils et dispositifs lors de leur implémentation en phase test :* organisation formelle de feed-back sous forme de questionnaires et d'entretiens.

S'ils s'avèrent efficaces, les choix opérés devraient fédérer les acteurs concernés autour du projet d'apprentissage de l'étudiant, ce qui pourrait remédier au manque d'implication des étudiants dans les activités d'intégration professionnelle mais aussi avoir un impact positif sur la qualité de leur encadrement.

4. Discussion

La grille d'évaluation de stage a été proposée à un échantillon représentatif des lieux de stages. Le portfolio a été proposé aux troisièmes bacheliers.

Les données pertinentes que nous pouvons extraire des questionnaires et entretiens auprès des acteurs concernés sont les suivantes :

4.1 Perception de la nouvelle grille d'évaluation en stage

4.1.1 Par les étudiants

De manière générale, l'accueil est très positif. Les étudiants reconnaissent l'aspect formatif de l'auto évaluation et l'aspect plus objectif qu'apporte l'évaluation partagée. Ils y voient un véritable outil communicationnel qui permet la triangulation entre stagiaire, superviseur et maître de stage.

Pour certains lieux de stage où la communication est faible, cet outil se révèle être d'une importance capitale. Les étudiants apprécient également l'évaluation formative intermédiaire qui offre une réelle opportunité de s'améliorer.

Par ailleurs, la grille étant en concordance avec le RCi, l'appropriation de celui-ci peut se faire par le fait que les étudiants s'y réfèrent, lors de l'autoévaluation, pour identifier leurs points forts et faibles face aux compétences qu'ils développent en stage.

4.1.2 Par les maitres et superviseurs de stage

De manière générale l'accueil est plutôt positif. Cependant, autant les superviseurs que les maitres de stage déplorent la complexité et la longueur de la grille et donc son impact chronophage. Ils reconnaissent éprouver des difficultés à se familiariser avec un vocabulaire spécialisé d'où l'intérêt de disposer du RCi. Ils reconnaissent l'atout communicationnel de l'évaluation partagée et apprécient le choix d'une évaluation qualitative critériée. De ce fait, ils souhaiteraient une échelle plus nuancée, mais aussi de pouvoir disposer de la liberté de formuler des indicateurs spécifiques à leur situation de stage. Ils se demandent quelle position adopter devant le fait que toutes les compétences et capacités ne sont pas nécessairement observables dans une seule situation de stage. Enfin, ils mentionnent le manque de réalisme de la compétence « administrer ».

4.2 Perception du portfolio

4.2.1 Par les étudiants

Un accueil plutôt mitigé. En extrayant du corpus, les idées clés, les étudiants disent de cet outil que « c'est un exercice difficile mais qui aide à prendre conscience et à trouver des idées » et que cela « permet une prise de recul ». Même si ils y voient un intérêt, ils pensent que « cela prend beaucoup de temps », que « c'est un casse-tête administratif », qu'il y a « trop de paperasses ». La structure du portfolio leur semble « morcelée », « pas claire », « peu cohérente ». Enfin, ils soulignent que tous les enseignants n'ont pas la même vision de l'outil, ce qui ajoute au manque de clarté générale.

4.2.2 Par les enseignants

Un bon accueil mais des difficultés liées à la structure et au manque de ligne directrice de l'outil. Les enseignants redécouvrent l'intérêt du travail d'équipe et reconnaissent leur implication dans ce projet. Ils constatent, comme les étudiants qu'il y a une appropriation différente de l'outil par chacun des enseignants, et reconnaissent que cette appropriation prend du temps. Ils ne sont pas vraiment satisfaits de la structure du portfolio. Du fait de la juxtaposition des données plutôt que de leur intégration, il manque d'un lien cohérent témoignant du développement des compétences. Ils relèvent la nécessité de renforcer la guidance de l'étudiant pour rendre cet outil plus opérant. Enfin, ils se questionnent sur la forme et le contenu de l'évaluation du portfolio : pourquoi et comment l'évaluer ?

5. Perspectives et questionnements

Le taux de participation des utilisateurs et la précision de leurs réponses nous permet de proposer une amélioration des outils et dispositifs.

5.1 Concernant la grille d'évaluation en stage

Nous avons simplifié l'utilisation de la grille sans vraiment l'alléger car nous souhaitons laisser la possibilité ouverte pour l'évaluation d'un nombre suffisant de compétences.

Nous avons proposé une échelle d'évaluation plus nuancée.

Nous avons ouvert une possibilité d'ajout d'indicateurs spécifiques.

5.2 Concernant le portfolio

Nous avons proposé une nouvelle structure davantage en cohérence avec les compétences du RCi et notamment la réflexivité.

Nous avons utilisé la plateforme « eduportfolio » pour y déposer un portfolio de référence.

Nous avons élaboré un guide de construction du portfolio et intensifié la guidance des étudiants, notamment en troisième bachelier.

Nous avons élaboré, en équipe, une grille d'évaluation du portfolio sur base des compétences identifiées.

Nous souhaiterions que l'aspect chronophage du portfolio soit considéré en valeur ECTS.

5.3 Nos questions en suspens concernant les outils et dispositifs proposés

L'effort d'appropriation par les utilisateurs et leur ouverture d'esprit, dans la situation particulière de période test, face au changement de paradigme pédagogique s'observera-t-il lors d'une application généralisée ? Comment garantir les aspects dynamique et interactif des outils et dispositifs dans la durée ?

Leur impact positif de triangulation et leur effet communicationnel nous permet par ailleurs de comprendre l'intérêt fondamental de l'intégration des logiques pédagogique et professionnelle dans l'approche socioconstructiviste d'un curriculum de formation professionnalisante.

Enfin, nous sommes confortés dans notre choix de paradigme socioconstructiviste : partagés dans un processus de co-construction, les changements sont plus opérants et agissent comme facteur de cohésion mais aussi de transformation.

6. Références bibliographiques

- Belanger, C. (2008). Rôle du portfolio au supérieur : rendre l'étudiant acteur de sa formation, *Actes du 25ème congrès de l'association internationale de pédagogie universitaire, Montpellier, France*.
- Jonnaert, P., Barette, J., Masciotra, D., Yaya, M. (2006). La compétence comme organisateur des programmes de formation revisitée, ou la nécessité de passer de ce concept à celui « de l'agir compétent », *Bureau International d'Education de l'UNESCO, Genève IBE Working papers issues*, N°4, (14-15).
- Jonnaert, P. (2009). Compétences et socioconstructivisme, un cadre théorique. *Éditions de Boeck*.
- Jouquan, J., Bail, P. (2003). A quoi s'engage-t-on en basculant du paradigme d'enseignement vers le paradigme d'apprentissage ? *Pédagogie Médicale*, 4 (163-175).

- Parent, F. (2008). Formation, compétences et constructivisme: le référentiel d'évaluation, outil de cohérence dans les programmes de formation en santé. In Baillat, G., De Ketele, JM., Paquay, L., Thélot, C., Evaluer pour former. Outils, dispositifs et acteurs. *Ed. De Boeck, Collection Pédagogies en Développement*, 111-139.
- Parent, F., Baulana R., Coppieters Y., Kahombo G., d'Hoop E., Lemenu D., Garant M., De Ketele J-M. (2011). Mieux gérer la cohérence pour renforcer les ressources humaines en santé : paradigmes et méthodes pour une intégration efficace des pratiques professionnelles et communautaires en formation. *Pédagogie Médicale*, 11(2), 111-25.
- Parent, F., Baulana R., Kahombo G., Coppieters Y., Garant M., De Ketele JM. (2010). A comprehensive competence-based approach in curriculum development : experiences from African and European contexts. *Health Education Journal*, 10.1177/0017896910376229.
<http://hej.sagepub.com/content/early/2010/09/16/0017896910376229.abstract>.
- Parent, F., d'Hoop, E., Dury, C., Baulana, R., Kahombo, G., Lejeune, C., De Ketele, JM. (2008). Modèles de navigation dans un cadre d'approche par compétences intégrée (APC¹). Taxonomie dans un cadre d'APC¹. *Bruxelles : ARE@ Santé ASBL*.
- Parent, F., Jougaunt, J., Kerkhove, L., Jaffrelot, M., De Ketele, J.M. (en cours de soumission). Intégration du concept d'intelligence émotionnelle à la logique de l'approche pédagogique par compétences dans les curriculums de formation en santé.
- Pastré, P., Mayen, P., Vergnaud, P. (2006). La didactique professionnelle. *Revue française de pédagogie*, n° 154, 145-198.
- Postiaux, N. (2010) Référentiels de compétences et pilotage de formation à l'Université. Rôles, enjeux et limites. *Thèse de doctorat. Faculté des sciences psychologiques de l'éducation, Université Libre de Bruxelles*.
- Rey, B., Carette, V., Defrance, A. et Kahn, S. (2003). Les compétences à l'école — Apprentissage et évaluation. *De Boeck Université*.
- Scallon, G. (2004). L'évaluation des apprentissages dans une approche par compétences. *De Boeck Université*.
- Tardif, J. (2006). L'évaluation des compétences. Documenter le parcours de développement. *Montréal, Chenelière Education*.

EVALUATION DES ENSEIGNEMENTS AU FORMAT PAPIER ET EN LIGNE : IMPACTS SUR LES PROCEDURES ET RESULTATS

Mallory Schaub*, Marine Chittaro*

* Université de Genève, mallory.schaub@unige.ch, marine.chittaro@unige.ch

Mots-clés : évaluation des enseignements, format papier, format en ligne, taux de réponse, indice de satisfaction

Résumé. Depuis 2009, l'Université de Genève a entrepris une réforme des principes et outils de l'évaluation de ses enseignements. La démarche d'évaluation des enseignements initialement formative, volontaire et individuelle a évolué vers un système indicatif, uniformisé et transparent. Cette réforme s'est accompagnée de l'acquisition du logiciel Evasys©, permettant le traitement automatique, à grande échelle, des évaluations tout en offrant la possibilité d'administrer les questionnaires en ligne ou sur papier. L'introduction des évaluations en ligne a suscité autant de soulagement (économies de temps et de moyens) que de crainte (baisse du taux de réponse, sur-représentativité des étudiants mécontents) de la part des enseignants et des autorités académiques. Si le nombre de répondants a effectivement baissé en ligne, la mesure de l'appréciation globale des étudiants à l'égard des enseignements n'est dans l'ensemble pas modifiée par l'administration en ligne. L'usage concluant du système se heurte néanmoins aux résistances des utilisateurs et bénéficiaires sensibilisés aux problématiques de validité et fiabilité des outils utilisés, en particulier lorsqu'il s'agit d'évaluer leurs propres activités.

1. Introduction

L'université de Genève a instauré des démarches d'évaluation des enseignements depuis un peu plus de 10 ans. Dès le début, les procédures privilégiées étaient l'administration de questionnaires aux étudiants. La démarche adoptée alors était celle de l'autonomie, du volontariat et de la confidentialité. Les enseignants avaient à disposition un secteur (Secteur Formation et Evaluation) leur fournissant sur demande et sur mesure des questionnaires papier à destination des étudiants ou d'autres méthodes d'évaluation (par observation notamment). Le secteur fournissait ensuite un traitement des données réceptionnées ainsi que l'élaboration de rapports de résultats fournis exclusivement aux enseignants. Les enseignants étaient ensuite libres d'interpréter, prendre en compte les données et d'en assurer le suivi. La régulation et le développement des pratiques pédagogiques étaient laissés à la discrétion des enseignants.

1.1 Nécessité d'une réforme des principes de l'évaluation des enseignements

Depuis 2009, l'Université de Genève a entrepris une profonde réforme de ses principes et procédures d'évaluation des enseignements (Schaub 2010). Les raisons qui ont suscité ce changement sont multiples et touchent tant à des principes de transparence, d'uniformisation, d'équité et d'autonomie. Elles ont émergé tant de ses propres constats à l'analyse de ses pratiques que de recommandations issues d'un audit de son système qualité par l'OAQ (Organe d'Accréditation et d'assurance Qualité des hautes écoles suisses) en 2007. Ces éléments ont amené le Rectorat à élaborer une directive institutionnelle en matière d'évaluation des enseignements et des enseignants. L'objectif poursuivi était de fournir un cadre et des principes directeurs pour l'ensemble des membres de la communauté universitaire. Cette réforme de l'évaluation des enseignements s'est accompagnée de l'élaboration d'une directive en matière d'évaluation des

programmes de formation et d'  valuation des entit  s (facult  s notamment), afin de favoriser l'  laboration d'une d  marche qualit   globale et coh  rente.

Dans cette r  forme, les objectifs vis  s par l'  valuation ont pour l'essentiel   t   maintenus par rapport aux pr  c  dentes attentes de l'  valuation. Celle-ci vise en priorit   de soutenir le d  veloppement de la qualit   des enseignements en permettant : 1) la d  tection, sur la base de l'avis des   tudiants, des forces et faiblesses des enseignements, 2) la mesure de l'impact d'innovations p  dagogiques ; 3) de guider le soutien    fournir aux enseignants pour qu'ils puissent d  velopper et valoriser leurs comp  tences et pratique p  dagogiques, et enfin 4) de servir la gestion et le pilotage interne aux facult  s en mati  re d'enseignement.

La directive sur l'  valuation des enseignements et des enseignants propose des recommandations minimales, chaque facult   pouvant adapter certains points    ses propres sp  cificit  s ou contraintes. Elle fournit un cadre pour l'uniformisation de l'  valuation en   non  ant les principes suivants : 1/ la syst  matisation de l'  valuation des enseignements dispens  s dans l'institution, 2/ la standardisation des outils d'  valuation et 3/ l'exigence de p  riodicit  s minimales d'  valuation. Elle vise aussi    favoriser une plus grande transparence pour une gestion collective des r  sultats et suites de l'  valuation. Elle   nonce une participation et responsabilisation accrue du doyen par l'obtention des r  sultats d'  valuation, par des recommandations en mati  re d'utilisation de ces donn  es et par la mise en place de mesures de suivi.

1.2 Adapter les outils au service de l'  valuation des enseignements

L'adoption de ces nouveaux principes a impliqu   de devoir adapter les outils au service de l'  valuation des enseignements. En effet, les principes d'  quit  , d'uniformit   et de standardisation ont n  cessit   l'  laboration de nouveaux questionnaires    destination des   tudiants. L'outil reste identique dans le principe (le questionnaire, soumis aux   tudiants au terme de l'enseignement au format papier, est compos   de questions ouvertes et ferm  es), mais change dans son contenu et son support.

Concernant le *contenu*, dor  navant, les questionnaires comportent un bloc d'  nonc  s standardis  s pour l'ensemble des enseignements de m  me type,   labor  s avec la facult   et touchant diff  rentes dimensions de l'enseignement. Ils comportent aussi un bloc d'items libres laiss  s    l'enseignant pour int  grer ses propres   nonc  s. Il contient aussi un   nonc   portant sur l'appr  ciation globale de l'enseignement consid  r  . Cet item est strictement identique pour tous les enseignements de la facult   permettant ainsi, dans les bilans globaux, d'avoir une mesure commune permettant toute analyse transversale et longitudinale. Ensuite, il inclut la possibilit   de fournir des commentaires libres concernant les forces et faiblesses des enseignements.

En outre, le *support* a chang   avec l'acquisition d'un logiciel sp  cifique pour l'  valuation des enseignements (Evasys  ). Ce syst  me a   t   acquis pour trois raisons principales. La premi  re touche    des aspects *  conomiques* : il permet de g  rer un grand nombre d'  valuations qui peuvent   tre g  n  r  es, diffus  es et trait  es simultan  ment en grande partie de mani  re automatis  e. Il permet ainsi de r  pondre    l'accroissement des   valuations impliqu   par le principe de syst  matisation et de p  riodicit   (depuis 3 ans, le secteur enregistre une hausse de plus de 20% des   valuations effectu  es par an). Auparavant saisis    la main, les questionnaires sont maintenant scann  s, r  duisant consid  rablement le temps consacr   auparavant    la saisie manuelle des r  ponses aux questions ferm  es¹. Ainsi, les forces en pr  sence peuvent   tre r  allou  es    l'analyse des donn  es et au suivi des   valuations. Ce syst  me permet de proposer un support papier ou en ligne pour les   valuations ce qui, engendre potentiellement une grande r  duction des consommations de papier.

¹ Pr  cisons n  anmoins que le secteur continue    retranscrire les commentaires libres des   tudiants afin de garantir le respect de leur anonymat lorsqu'ils remplissent les questionnaires au format papier.

La deuxi  me touche    des aspects *m  tier*. Le secteur en charge des   valuations peut proposer un syst  me soit en ligne soit papier, r  pondant aux caract  ristiques des enseignements    distance et au profil des   tudiants et enseignants actuels qui sollicitent ce type de moyen technologiques. Le stockage et la s  curit   des donn  es sont accrus, l'anonymat des   tudiants est mieux assur   et le risques de pertes de donn  es est r  duit (pour autant que ce soit le format en ligne qui soit utilis  ). Ensuite, il constitue un puissant outil d'analyse qui permet de compiler un nombre consid  rable de donn  es dans des rapports longitudinaux ou transversaux. Si ces analyses   taient d  j   possibles avant, elles n  cessitaient un travail long et fastidieux de recherche et agglom  ration de donn  es.

La troisi  me raison touche    des aspects de *pilotage*. Le logiciel permet de mettre    disposition des autorit  s facultaires des indicateurs qualit  s permettant de comparer des groupes (   l'  chelle de la facult   ou d'un programme) ou individus    une valeur   talon bas  e sur une cohorte pass  e ou pr  sente ou sur une valeur seuil. Ces choix sont effectu  s en concertation avec les autorit  s que nous conseillons sur les alternatives possibles et leurs   ventuelles cons  quences. Ces donn  es peuvent ensuite   tre mises sur la plateforme d'  valuation accessible au doyen ou directeur de d  partement qui assure ensuite le suivi des   valuations aupr  s des enseignants.

2. Informatiser l'  valuation des enseignements

La mise en place de cet outil fait l'objet d'un projet institutionnel de syst  me d'information r  alis   en deux parties. La premi  re phase, dont les r  sultats liminaires sont pr  sent  s dans le cadre de cette communication, a consist      tester l'outil dans une facult   de notre institution en deux   tapes : un premier semestre au format papier et un second au format en ligne, d  crivant ainsi une transition progressive pour les enseignants et pour le service devant alimenter des bases de donn  es enseignants et enseignements dans un premier temps, puis, en plus dans un second temps, une base de donn  es   tudiants.

A son terme, les donn  es obtenues ainsi que le calcul des co  ts de cette phase font l'objet d'un bilan interm  diaire exposant les forces et faiblesses de l'outil, les nouveaux besoins   ventuels et des r  ajustements n  cessaires en vue de son d  ploiement durant la deuxi  me phase. Durant celle-ci, l'outil sera utilis   pour l'  valuation des enseignements de l'ensemble des 6 facult  s, 3 instituts, et de la formation continue. Un bilan final permettra de prendre la mesure de l'impact de la mise en   uvre d'un tel outil en termes de co  ts et ressources, de proc  dures et aussi en terme d'ad  quation par rapport aux attentes et besoins des diff  rents acteurs institutionnels.

Afin de prendre la mesure de l'impact d'un tel changement, le secteur a entrepris de comparer les r  sultats obtenus au terme de chacun des deux semestres de la premi  re phase    ceux obtenus l'ann  e pr  c  dente sur les deux m  mes semestres pour les m  mes cours donn  s par les m  mes enseignants. Une telle analyse est rendue possible en raison des conditions de l'  valuation instaur  es dans la facult   qui sert de pilote pour la premi  re phase. En effet, tous les enseignements ont   t     valu  s durant deux ann  es cons  c  tives, avec les m  mes questionnaires, rendant les comparaisons de donn  es des   valuations possibles.

Durant la premi  re phase, le secteur a   valu   l'ensemble des enseignements d'une facult   qui compte 3544   tudiants et 286 enseignants qui dispensent environ 600 enseignements par ann  e. Pour effectuer les   valuations, 3 questionnaires de base (cours, TP, s  minaire) disponibles en fran  ais et en anglais ont   t   mis    disposition des enseignants et   tudiants. Durant le semestre d'automne, les enseignants ont re  u un de ces questionnaires²    distribuer    leurs   tudiants au terme d'une des derni  res s  ances de cours du semestre. Ces questionnaires, une fois remplis, ont   t   retourn  s au secteur, afin de pouvoir proc  der    la saisie num  rique, la retranscription des

² A noter que les enseignants avaient la possibilit   de faire ajouter des commentaires libres, ce que seuls 5 d'entre eux ont effectu  .

commentaires libres et la mise en forme des rapports de r  sultats. Au semestre de printemps, chaque   tudiant a re  u une adresse URL lui donnant acc  s    un questionnaire en ligne pour chaque cours suivi. En tout, 17'500 URL ont   t   fournies et 6'700 questionnaires remplis ont   t   enregistr  s en retour. Au terme de chacun des semestres, les enseignants ont re  u un rapport de r  sultats pour chaque enseignement   valu  . Ils ont   t   fournis avant le terme exact du semestre afin que les enseignants puissent s'ils le souhaitent, fournir un feedback de ces   valuations aux   tudiants.

3. Premiers r  sultats

3.1 Impacts sur les proc  dures

Le nombre d'enseignements   valu  s a augment   de plus de 35% dans cette facult  , ce qui est imputable    la mise en   uvre de la directive avec r  duction de la p  riodicit      son minimum (chaque enseignement   valu      chaque occurrence). Le secteur a proc  d      l'  valuation de 588 enseignements, dont 307 au format papier (12'000 questionnaires saisis) et 281 au format en ligne (6'700 questionnaires remplis). Parmi les enseignements   valu  s, un peu plus de la moiti   sont des cours (58%), un tiers des cours-s  minaires, 9% des s  minaires et enfin 3% des TP ou ateliers.

Le premier impact du passage en ligne est une variation du taux de r  ponse. Il n'y a pas eu de baisse du nombre de r  pondants entre les semestres d'automne 2009 et 2010, tous deux   valu  s au format papier. Au contraire, entre ces deux semestres, il y a une augmentation du nombre de r  pondants de pr  s de 30%. En revanche, entre les semestres de printemps 2010 et 2011   valu  s respectivement au format papier puis en ligne, il y a un quart de r  pondants en moins. Le taux de r  ponse moyen global pour les   valuations effectu  es en ligne est de 44%.

Au-del   de ces deux mesures, le passage aux   valuations en ligne a aussi permis de r  duire la consommation de feuilles de papier, de r  duire le temps consacr      la retranscription des commentaires libres, de r  allouer ce temps    la r  alisation de bilans globaux et individuels et de pouvoir assurer la prise en charge de l'ensemble des   valuations par un effectif de ressources restreint. Ces avantages d'ordre plut  t   conomiques ne doivent n  anmoins pas occulter les   ventuels impacts d'un passage aux   valuations en ligne sur les r  sultats des   valuations. En effet, tant pour les enseignants que pour le service en charge des   valuations, le point d'int  r  t consiste    savoir si la mesure de la satisfaction des   tudiants pourrait   tre influenc  e par ces variations du taux de r  ponse.

3.2 Impacts sur les r  sultats des   valuations

Au-del   de la baisse du taux de r  ponse, ce qui nous int  resse le plus est son   ventuel impact sur la mesure de la satisfaction des   tudiants. Pour cela nous avons consid  r   les r  ponses fournies    un   nonc   d'appr  ciation globale³ contenu dans l'ensemble des questionnaires. Une valeur comprise entre 0 et 4 est attribu  e    chaque r  ponse donn  e (« oui », « plut  t oui », « plut  t non », « non »). La somme de ces valeurs est ensuite divis  e par le nombre total de r  ponses    cet   nonc   (moins les « sans avis » qui ne sont pas comptabilis  s) afin d'obtenir un indice d'appr  ciation globale.

Nous relevons que la mesure de l'appr  ciation globale des   tudiants    l'  gard des enseignements n'est dans l'ensemble pas influenc  e par l'administration en ligne. En effet, l'appr  ciation globale moyenne varie de 3.49 (sd=0.31) au format papier    3.41 (sd=0.39) au format en ligne, avec des m  dianes tr  s similaires (respectivement 3.50 et 3.48) et une proportion d'enseignements consid  r  s comme insatisfaisants identique (respectivement 7% et 9%). Au niveau des donn  es de l'ensemble de la facult  , les r  sultats entre les deux semestres et donc entre les deux formats, sont tr  s similaires. Les donn  es fournies ici repr  sentent des regroupements et moyennes qui doivent

³ L'  nonc   est le suivant : « globalement, la qualit   de ce enseignement est satisfaisante ».

encore   tre trait  es de mani  re plus d  taill  e pour permettre une analyse et une comparaison plus fines de l'impact de l'outil sur les proc  dures et r  sultats de l'  valuation. En effet, nous devons encore d  terminer si,    un niveau individuel (un cours), les m  mes constats seraient possibles.

4. Conclusion

Au jour d'aujourd'hui, nous dressons un premier bilan satisfaisant de la mise en place de ce logiciel, tant il r  pond aux exigences et principes de l'  valuation dans la nouvelle forme que celle-ci a prise (proc  dures automatis  es,    large   chelle,   laboration d'indicateurs qualit  , etc.) mais aussi aux attentes strat  giques et   conomiques de la direction. Cet usage concluant se heurte n  anmoins aux r  sistances des utilisateurs/b  n  ficiaires. Les enseignants se montrent m  fiant  s    l'  gard de la validit   de cet outil et de la fiabilit   des donn  es qu'il g  n  re quand, notamment, il entra  ne une baisse du taux de r  ponse ce qui a effectivement   t   observ  . Actuellement, avec pour seule mesure incitative un rappel par mail adress   aux   tudiants, le taux de r  ponses des   valuations en ligne atteint un peu plus de 40%, taux habituellement rencontr   pour de telles enqu  tes (Dommeyer, Baum, Hanna, & Chapman, 2004; Nulty, 2008).

Cependant, nous retenons pour l'instant l'absence de diff  rence entre les r  sultats des   valuations obtenus avec un format papier et un format en ligne, corroborant de pr  c  dentes analyses similaires (Liu, 2003; Hardy, 2006). Malgr   la baisse du taux de r  ponse, il ne semble pas y avoir de variation concomitante de la mesure de la satisfaction des   tudiants. Toutefois, nous explorerons davantage dans le d  tail les donn  es issues de ces   valuations pour confirmer cette observation g  n  rale. Cette baisse du taux de r  ponse alimente cependant les craintes des enseignants, sensibles    la validit   et la fiabilit   de l'outil utilis  , d'autant qu'il s'agit d'  valuer leurs prestations d'enseignement. Ce n'est peut-  tre pas tant sur l'outil que sur le fait m  me d'  tre   valu   que se basent ces r  sistances.

La rigueur de notre d  marche d'analyse des r  sultats de l'  valuation fournira - ou pas - des arguments pour mettre en avant la validit   et la fiabilit   de l'outil utilis  . Parall  lement, le secteur travaillera d  s la deuxi  me phase du projet d'informatisation    la mise en   uvre de mesures permettant de relever ou du moins maintenir un taux de r  ponse satisfaisant. La communication, la visibilit   des effets de l'  valuation, des rappels constituent quelques-unes de ces mesures permettant de promouvoir la participation des   tudiants aux enqu  tes effectu  es en ligne (Nulty, 2008). Il s'agira si n  cessaire de rassurer la communaut   sur la pertinence de cet outils et r  orienter les pr  occupations des enseignants non pas sur le taux de r  ponse mais sur les r  sultats des   valuations, leurs utilisations et suivis. En outre, la souplesse et l'adaptabilit   du syst  me permet de r  pondre    l'essentiel des requ  tes des enseignants et autorit  s facultaires. Ainsi, cela devrait favoriser l'adoption du syst  me et de la nouvelle d  marche d'  valuation dans son ensemble par les diff  rents acteurs et de fait, r  duire cette tension entre les besoins et attentes des diff  rents acteurs institutionnels.

5. R  f  rences

- Dommeyer, C. J., Baum, P., Hanna, R. W., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys : their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education*, 29(5), 611-623.
- Hardy, N. (2003). Online ratings: Fact and fiction. *New Directions for Teaching & Learning*, 96, 31-39.
- Liu, Y (2006). A Comparison Study of Online versus Traditional Student Evaluation of Instruction. *International Journal of Instructional Technology and Distance Learning*, 3(4), 15-29.
- Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education*, 33(3), 301-314.

Actes du 24^e colloque de l'Admécé-Europe
L'évaluation des compétences en milieu scolaire et en milieu professionnel

Schaub, M. (2010). Evaluation de l'enseignement à l'Université de Genève : d'une démarche individuelle vers une démarche institutionnelle de la qualité. *Paper presented at the 26ème congrès de l'AIPU*, Rabat, 17-21 mai 2010.

**ANALYSE DE VALIDITE DES TESTS DE CLASSEMENT EN LIGNE POUR LES
LANGUES A L'AIDE DU MODELE DE RASCH**

Cornelia Gick*, Herv   Platteaux **, Sergio Hoein*, Catherine Blons-Pierre****, Patricia Kohler*******

* *Universit   de Fribourg (Suisse), cornelia.gick@unifr.ch*

***Universit   de Fribourg (Suisse), herve.platteaux@unifr.ch*

*** *Universit   de Fribourg (Suisse), sergio.hoein@unifr.ch*

**** *Universit   de Fribourg (Suisse), catherine.blons-pierre@unifr.ch*

***** *Universit   de Fribourg (Suisse), patricia.kohler@unifr.ch*

Mots-cl  s : *test de classement en ligne, IPT/Analyse d'item, Mod  le/Analyse de Rasch, Apprentissage des langues, Plateforme Moodle*

R  sum  . *L'Universit   bilingue de Fribourg (Suisse) a r  alis   des tests de classement en ligne adaptatifs pour quatre langues. Avec la plateforme Moodle, les   tudiants passent un test, qui   value leurs comp  tences langagi  res, afin de pouvoir s'inscrire    un cours de langue   trang  re adapt      leur niveau. Depuis quatre semestres, les tests de classement en ligne ont   t   utilis  s avec succ  s (tr  s peu d'  tudiants ont d   changer de niveau de cours). Pour augmenter encore la qualit   de cette proc  dure, nous analysons la validit   et de la fid  lit   des items des diff  rents sous-tests con  us pour l'allemand (DaF) et le fran  ais (FLE)    l'aide du mod  le de Rasch (Winsteps). Cela pour deux actions : la constitution des tests et de leurs sous-tests (mode adaptatif) et l'am  lioration des items (qui participe    la formation continue des comp  tences des   quipes de d  veloppement des items).*

1. Introduction

Le Centre de Langues et le Centre Nouvelles Technologies et Enseignement, de l'Universit   bilingue de Fribourg (Suisse), ont r  alis   des tests de classement (placement tests) en ligne adaptatifs pour l'allemand, l'anglais, le fran  ais et l'italien langues   trang  res. Avec la plateforme d'apprentissage en ligne Moodle, les   tudiants passent un test, qui   value leurs comp  tences langagi  res, avant de pouvoir s'inscrire, via un syst  me d'inscription en ligne,    un cours de langue   trang  re adapt      leur niveau. Mis ensemble, notre test de classement et le syst  me d'inscription permettent d'automatiser le processus d'inscription et d'admission aux cours de langues.

Notre test de classement en ligne est un dispositif d'  valuation formative puisque son objectif premier est de d  terminer le niveau de comp  tence de chaque apprenant afin de lui proposer des cours correspondant    son niveau. Pour le Centre de Langues qui g  re le test, le but est aussi de cr  er des classes de niveaux relativement homog  nes.

Le choix d'une plateforme d'apprentissage en ligne pour mettre en   uvre nos tests de classement adaptatifs (Linacre, 2000) provient d'une volont   de mettre    profit plusieurs forces de ces technologies. En premier lieu, elles permettent de r  aliser une   conomie organisationnelle dans le cadre d'un enseignement/apprentissage en grands groupes. Avec une plateforme en ligne, chaque

semestre, 1'500   tudiants passent le test o   et quand ils veulent sans qu'ils aient besoin de faire le pied de grue dans les couloirs du Centre de Langues. Le stress des enseignants, et du secr  tariat, en est diminu   d'autant durant la m  me p  riode de la rentr  e. L'  valuation de comp  tences r  ceptives permet un calcul automatique du niveau langagier de l'  tudiant    partir de ses r  ponses au test en ligne.

Quelques principes r  sument le mod  le de ces tests :

- Ils sont compos  s de trois cat  gories de sous-tests   valuant les comp  tences r  ceptives la compr  hension orale et la compr  hension   crite ainsi que les structures de la langue.
- Il int  gre l'auto  valuation selon les descripteurs du Cadre Europ  en Commun de R  f  rence pour les langues du Conseil de l'Europe (CECR).
- Pour chaque cat  gorie, plusieurs sous-tests sont propos  s aux   tudiants selon une modalit   adaptative.
- Les deux tests utilisent un instrument qui permet    l'  tudiant d'  tre dirig   de mani  re efficace dans un sous-test de son niveau. Pour le test FLE (Fran  ais Langue Etrang  re), c'est l'auto  valuation sur la base du CECR ; pour le test DaF (Deutsch als Fremdsprache), ce sont 8-10 items de grammaire dont le r  sultat dirige vers un autre test de grammaire de 18-20 items.
- La cr  ation des items (questions et propositions de r  ponse) des sous-tests est bas  e sur le CECR (Blons-Pierre, 2010).

Les tests ont   t   utilis  s avec succ  s : tr  s peu d'  tudiants ont d   changer de niveau de cours. Depuis le lancement des tests en ligne, chaque semestre d  s f  vrier 2009, une r  colte syst  matique de donn  es a permis de mener diverses analyses que nous r  sumons dans une premi  re partie de cet article. Sur cette base, pour augmenter encore la qualit   des banques d'items, nous nous attachons    analyser plus en d  tails la validit   et la fid  lit   (Tagliante, 2005) des items des diff  rents sous-tests con  us pour les tests DaF et FLE.

Pour ce faire, nous utilisons la th  orie de r  ponse aux items (Lord, 1980 ; Wright & Stone, 2004), en particulier le mod  le de Rasch (1980), et le logiciel Winsteps, qui nous permettent de d  terminer le pouvoir de discrimination d'un item et son degr   de difficult   par rapport aux autres, ainsi que de formuler des indications concernant leur qualit  . Nous disposons maintenant d'une population suffisante pour effectuer un tel traitement. Pour le test FLE, 400   tudiants environ ont pass   le test entre les sessions de septembre 2010 et de f  vrier 2011. Pour le test DaF, 1000   tudiants environ, entre f  vrier 2009 et septembre 2011, dont 650 environ durant les sessions septembre 2010 et 2011 et f  vrier 2011 avec un test inchang  .

Ces r  sultats sont utilis  s avec deux finalit  s qui forment le c  ur du pr  sent article :

- En premier lieu, l'analyse de Rasch intervient dans la conception d'une structure adaptative des tests. Elle permet de positionner les items dans les sous-tests, en fonction de leurs degr  s de difficult  , de v  rifier si les degr  s de difficult   des items (et des sous-tests) correspondent    la population test  e (Gick, 2011) et de v  rifier la qualit   des items utilis  s.
- En second lieu, l'analyse de Rasch sert au d  veloppement continu de la production des items. Elle permet d'identifier des items probl  matiques dans les diff  rents tests cibl  s sur un niveau. Ces items sont alors soumis    l'  quipe de conception qui propose des am  liorations. Les analyses selon le mod  le de Rasch servent ensuite    v  rifier si ces changements portent leurs fruits lors de la passation suivante.

2. R  sultats des analyses pr  c  dentes

La base des deux tests en ligne DaF et FLE sont les tests papier pr  existants et valid  s. Depuis le lancement des tests en f  vrier 2009, diverses analyses ont   t   men  es,    partir d'une r  colte syst  matique de donn  es, dans le but de v  rifier la validit   de notre mod  le de test adaptatif et son op  rationnalisation dans Moodle.

On a regard   le cheminement de l'  tudiant dans le test (nombre sous-tests faits, dur  e totale, etc.) et la perception qu'il en construisait (Kohler, Platteaux, & Blons-Pierre, 2012). Ces analyses ont permis de r  gler les seuils de r  ussite des sous-tests, leurs limites en temps et le nombre d'items qui les composent. En parall  le une analyse quantitative, faite    partir des r  sultats obtenus aux tests en ligne par les premiers   tudiants selon le mod  le de Rasch, a permis de faire un premier tri sur la qualit   des items, de supprimer des items estim  s probl  matiques, ainsi que d'ajouter et de valider de nouveaux items (Gick, 2012).

Les feedbacks des   tudiants sur les tests en ligne ont aussi permis de conna  tre et d'am  liorer l'acceptance des tests en ligne (Kohler, Platteaux, & Blons-Pierre, 2010). Cette analyse se poursuit tant sur l'acceptance que sur le degr   d'accord entre l'auto  valuation effectu  e par les test  s et le r  sultat de l'  valuation propos   par le test de classement (Kohler, 2012 ; Blons-Pierre, 2011 ; Blons-Pierre, Kohler, Gick, Hoein et Platteaux, 2012).

On a   galement v  rifi  , qu'   partir des r  ponses au test en ligne de l'  tudiant, un calcul automatique pouvait   tre mis en place dans Moodle en utilisant les fonctions de tests de cette plateforme et les   l  ments d'  valuation associ  s    son tableau de notes pour   valuer des comp  tences r  ceptives langagi  res (Platteaux & Hoein, 2010).

3. Analyse de Rasch dans le contexte de nos tests

Dans la th  orie de la r  ponse aux items (IRT), il y a une estimation pour les personnes,    partir des items r  pondus vers les items sans r  ponse, en partant d'une probabilit   de r  solution de 50 % (Rost 2004). Ceci permet de formuler des observations sur les items et la totalit   du test m  me si, pour une partie du test, nous avons un nombre limit   de participants. Cette analyse nous donne des informations sur le degr   de difficult  , l'adaptation au mod  le et la qualit   des items. Le LMS Moodle permet, avec un module suppl  mentaire, d'obtenir les r  ponses d  taill  es de chaque   tudiant    chaque item. Avec le logiciel Winsteps, nous obtenons relativement rapidement les r  sultats que nous cherchons. Si ce logiciel rend les analyses tr  s rapides et permet   galement d'obtenir un grand nombre de formats de sortie et de repr  sentation des r  sultats, il est n  cessaire de pr  parer les donn  es en les codant en chiffres.

En raison de la structure adaptative de nos tests, il nous faut   galement r  unir toutes les r  ponses de tous les   tudiants aux diff  rents sous-tests dans un seul tableau. Nous traitons uniquement les questions de type QCM, qui constituent la grande majorit   des items du test FLE et sont le seul type dans le test DaF. Pour que l'analyse soit assez fiable, nous ne prenons en compte que les tests pour lesquels nous avons un nombre suffisant de r  ponses. A cause du mod  le adaptatif, il arrive que nous n'obtenions pas assez de r  ponses en une seule session. Si cela se produit, nous d  cisons de n'analyser que les sous-tests avec un nombre de r  ponses suffisant, ou d'attendre un semestre ult  rieur avant de proc  der    des changements. A partir de 70 r  ponses les r  sultats semblent   tre significatifs. Pendant le d  veloppement du test DaF, on a utilis   sous r  serve les r  sultats    partir de 30 r  ponses.

M  me si les deux tests FLE et DaF sont semblables, il existe des diff  rences qui se manifestent   galement lors de l'utilisation de l'analyse de Rasch : les niveaux du CECR (Cadre europ  en commun de r  f  rence pour les langues) sont au centre du test FLE ; chaque sous-test se positionnant par rapport    un des niveaux. Cela n  cessite une discussion approfondie sur ces niveaux dans l'  quipe de d  veloppement, ainsi qu'une phase d'exp  rimentation (Kohler, 2012). Ces deux actions ont pu   tre mises en   uvre et il a   t   possible de construire sur des travaux pr  alables. L'auto-positionnement, qui sert d'entr  e dans chaque comp  tence, a un r  le important. Pour les   tudiants, le niveau test   dans un sous-test est toujours clair. L'  tudiant est dirig   vers un sous-test d'un niveau inf  rieur, d'un niveau sup  rieur ou vers une autre comp  tence. Le chemin    travers le test passe d'abord par la compr  hension orale (CO) et   crite (CE), pour arriver ensuite    la composante « lexique et structure de la langue ». Dans les tests de CO et CE, plusieurs questions sont pos  es sur la base d'un texte (  crit ou oral). Pour l'utilisation de l'analyse de Rasch, cela

signifie que si un item est d'un niveau inf  rieur ou sup  rieur, il ne pourra pas simplement   tre d  plac   dans un autre sous-test, mais devra   tre adapt  . Cela justifie donc l'utilisation de l'analyse de Rasch dans une   tape ult  rieure, pour v  rifier les r  sultats de cette adaptation.

Le test DaF est bas   sur une progression et un continuum. Il a ainsi   t   possible de commencer avec un nombre r  duit d'items (18 courtes t  ches de compr  hension   crite) et peu de sous-tests (trois tests de compr  hension   crite). L'analyse de Rasch a permis le placement des items sur une   chelle de difficult   et l'identification des items probl  matiques. Le niveau de cours est d  termin   par un calcul, invisible aux   tudiants, fait    partir de leurs r  sultats aux sous-tests. La difficult   majeure   tait de trouver les seuils de c  sure entre les niveaux. Le test ne pouvant   tre mis    l'essai avant son utilisation en situation r  elle, le d  veloppement d'un tel test en si peu de temps impliquait un certain risque. Par cons  quent l'attribution des niveaux de cours a   t   faite, pendant la premi  re semaine de la premi  re passation du test, par le biais d'une confirmation manuelle. Pour cela, les r  sultats calcul  s ont   t   compar  s avec l'auto-  valuation des   tudiants, en tenant compte des renseignements personnels, et le calcul a   t   en partie ajust  . C'est seulement apr  s cette premi  re semaine que tout a   t   automatis  . L'auto-  valuation du niveau de langue bas  e sur le CECR se fait    la fin du test, et continue    servir uniquement comme une aide    la d  cision en cas d'h  sitation entre deux niveaux.

Au cours des sessions suivantes, de nouvelles t  ches ont pu   tre ajout  es. La r  partition sur plusieurs sous-tests et l'int  gration de nouveaux items ont montr   des probl  mes pendant la session de printemps 2010. Depuis le semestre d'automne 2010, nous pouvons utiliser un test fiable, qui peut   tre analys   plus en d  tail et affin  . Cela n'aurait pas   t   possible sans l'utilisation en parall  le de l'analyse de Rasch.

4. Rasch pour la conception des tests

Depuis le d  but, le d  veloppement du test DaF a   t   accompagn   pas    pas par l'analyse de Rasch. Apr  s chaque session, nous nous int  ressons en premier lieu aux valeurs suivantes :

- Le degr   de difficult   (Measure) des items par comp  tence. Les valeurs indiqu  es ici ont   t   utiles pour la distribution des items dans les sous-tests.
- La relation entre la difficult   des items et le niveau de la population test  e, donc la question: est-ce que les items propos  s couvrent suffisamment l'  ventail des niveaux ? Cela est identifiable dans le « Item Map ».
- La relation des trois sous-tests entre eux.
- La qualit   des items et des distracteurs par rapport    l'« Infit » et au « Point-Measure ».

Pour d  velopper le test DaF on s'  st bas   sur les questions    choix multiples de grammaire du test de classement papier qui fournissaient des indicateurs utiles pour diff  rencier les niveaux A2 - B2. Nous avons analys   les donn  es des t  ches de grammaire de la derni  re passation du test papier, au semestre d'automne 2008. Cela a fourni des points de rep  re pour la premi  re version   lectronique des items de grammaire. A partir du semestre de printemps 2009 les donn  es obtenues dans Moodle ont   t   pr  par  es pour   tre ensuite analys  es avec le logiciel Winsteps. Ce qui nous a int  ress  s tout particuli  rement,    part les valeurs de difficult   (Measure), a   t   le spectre de difficult  s que ces t  ches couvraient. Pour cela nous avons utilis   l'« Item map ». L'impression subjective qu'il manquait des items dans le segment des niveaux sup  rieurs a   t   confirm  e. On a montr     galement ainsi que le test   tait globalement trop bas par rapport au niveau de notre population ; il fallait donc le compl  ter avec des items    des niveaux sup  rieurs. Les valeurs de la « Measure » et de l'« Infit » ont permis de s  lectionner huit items de diff  rents niveaux pour le sous-test de distribution qui repr  sente la porte d'entr  e du test. L'analyse a   galement permis d'impl  menter une progression du degr   de difficult   dans le test. Pour le semestre d'automne cela signifie qu'au moins un sous-test a pu   tre construit sur la base d'items valid  s, et qui a   t   utilis   comme point d'ancrage pour le d  veloppement futur.

Les items nouvellement d  velopp  s pour les comp  tences r  ceptives ont   t   utilis  s pour la premi  re fois    la session d'automne 2009. Ils ont   t   regroup  s, selon le degr   de difficult   suppos  , en trois sous-tests de compr  hension   crite (CE) et quatre sous-tests de compr  hension orale (CO). Pour obtenir un nombre suffisant de tentatives, les   tudiants ont   t   envoy  s, selon leur r  sultat    un sous-test, au sous-test de difficult   sup  rieur ou inf  rieur. Dans cette structure, il peut arriver qu'un   tudiant obtienne un r  sultat bas dans le test de grammaire, et qu'il soit donc dirig   vers un test CE de bas niveau, qu'il r  ussisse avec un bon r  sultat et qu'il se trouve devoir passer tout les trois sous-tests CE. En r  ussissant le test CE de plus haut niveau, il est ensuite dirig   vers un test CO de haut niveau. S'il n'a pas alors un bon niveau dans cette comp  tence, et obtient un r  sultat bas, il est redirig   vers le test CO d'un niveau inf  rieur. Si son r  sultat est    nouveau bas, il est encore redirig   vers un niveau inf  rieur. Dans cette configuration, la dur  e du test est rallong  e, mais cela est une contrainte    accepter si l'on veut obtenir des donn  es fiables pour l'analyse de Rasch.

TABLE 13.1 DaF-Einstufungstest-HS09													
ZOU368WS.TXT Mar 11 17:54 2012													
INPUT: 273 Student 114 Item REPORTED: 255 Student 18 Item 2 CATS WINSTEPS 3.71.0.1													
Student: REAL SEP.: 1.04 REL.: .52 ... Item: REAL SEP.: 6.03 REL.: .97													
Item STATISTICS: MEASURE ORDER													
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFI T	INFI T ZSTD	OUTFI T	OUTFI T ZSTD	PT-MEASURE CORR.	PT-MEASURE EXP.	EXACT OBS%	MATCH EXP%	Item
69	14	37	3.23	.41	.80	-1.1	.78	-.8	.68	.57	78.8	74.1	LV3_03
67	17	38	2.78	.39	.91	-.5	.84	-.7	.60	.54	73.5	70.4	LV3_01
72	16	33	2.64	.42	.81	-1.2	.74	-1.3	.66	.54	75.9	69.7	LV3_06
71	24	36	1.66	.40	1.09	.6	1.17	.8	.37	.44	62.5	69.2	LV3_05
66	36	157	1.58	.22	.93	-.6	.84	-.7	.49	.45	83.2	80.3	LV2_06
68	28	36	.98	.44	1.12	.7	1.09	.4	.29	.36	78.1	75.3	LV3_02
70	29	37	.88	.43	1.04	.3	.94	.0	.35	.36	72.7	76.0	LV3_04
62	62	145	.43	.19	1.13	1.6	1.20	1.4	.40	.49	66.2	71.7	LV2_02
63	65	148	.33	.19	1.09	1.1	1.12	.9	.42	.49	71.9	71.7	LV2_03
60	43	90	-.29	.25	.97	-.4	.90	-.7	.52	.49	65.0	67.3	LV1_06
65	89	135	-.72	.20	1.08	.9	1.06	.4	.39	.44	69.9	72.9	LV2_05
61	107	163	-.90	.19	.98	-.1	.88	-.7	.51	.49	72.8	74.1	LV2_01
58	71	117	-1.08	.22	.95	-.6	.91	-.8	.51	.47	69.5	67.7	LV1_04
59	83	121	-1.52	.22	1.01	.1	1.02	.2	.45	.46	70.4	72.2	LV1_05
64	107	140	-1.53	.22	1.12	1.1	1.12	.5	.36	.43	75.4	79.2	LV2_04
55	119	154	-2.36	.22	.84	-1.3	.73	-1.5	.58	.47	81.3	79.4	LV1_01
56	113	139	-2.57	.24	.87	-.9	.72	-1.2	.51	.42	83.3	81.5	LV1_02
57	124	136	-3.55	.33	1.06	.3	1.07	.3	.31	.35	91.1	91.2	LV1_03
MEAN	63.7	103.4	.00	.29	.99	.0	.95	-.2			74.5	74.7	
S.D.	38.1	50.1	1.89	.09	.11	.8	.15	.8			7.0	5.8	

Figure 1 : DaF HS09 – Item Statistics : Measure Order

L'analyse de Rasch pour la compr  hension   crite (cf. Figure 1) donne l'image suivante :

- La relation entre la difficult   attendue (Entry Number: Item 55-72) et celle effectivement mesur  e (Measure)   tait d  j   bonne lors de la premi  re passation du test.
- La distribution des items dans les trois sous-tests (LV1-LV3) montre   galement des bons r  sultats.
- Les items LV2-04, LV2-02 et LV3-02 ont attir   notre attention    cause de leur « Infit MNSQ », qui dans le cas id  al devrait se situer autour de 1.00. Mais nous pouvons continuer    utiliser ces items en constatant : 1) que la valeur z « Infit ZSTD », qui devrait se trouver entre -2.00 et +2.00, est bon pour les trois items ; 2) que les valeurs du « Point Measure » sont sup  rieures    .20 ; 3) et que les valeurs qui   quilibrent les diff  rences de niveau entre les sous-tests (PT-Measure Exp.) sont   galement bonnes.
- Pour l'item n. 63 (LV2-04), la valeur basse du « Measure » (-1.53) est surprenante. Dans ce cas pr  cis, il a paru int  ressant d'examiner de plus pr  s.

Le graphique des « Expected Score ICC » pour la « Entry Number 63 » (LV2-04) montre que les   tudiants avec un niveau bas ont r  ussi    r  soudre la t  che. Cela est surprenant, car le texte de cet item de compr  hension   crite contient beaucoup de mots qui ne sont probablement pas compris    ce niveau (cf. Figure 2). Nonobstant cela, l'item a pu   tre r  solu correctement.

Verwirrung im Vogelreich

Schlecht geschlafen, weil Drossel, Fink oder Star so laut gepfeifen haben? Selber schuld. Die V��gel konnten nichts daf��r. Sie passen sich bloss ihrer Umgebung an und machen es demzufolge wie die Menschen: Je h��her der Ger��uschpegel, desto energischer wird geschrien. Amseln, Rotkehlchen oder Kohlmeisen mutieren also insbesondere dann zu Schreih��lsen, wenn sie aus dem stillen Wald in die l��rmige Stadt ziehen. Und das tun die Piepser, weil sie in den Strassen und Gassen vielf��ltigere Nahrung finden und weniger Feinde haben. F��r das sorglose Leben im ��berfluss m��ssen die V��gel aber einen hohen Preis bezahlen: Die Urbanisierung bedroht ihre Existenz.	Und zwar deshalb, weil der ��berm��ssige Stimmeinsatz die Verst��ndigung mit den Artgenossen im Wald erschwert. Zwitschert zum Beispiel die Stadt-Meise lauter, als sie soll, kann die Land-Meise sie f��lschlicherweise f��r eine Amsel halten und ihren Balzgesang ganz einfach ignorieren. Was macht die Stadt-Meise dann? Na klar: noch lauter pfeifen. Damit lockt sie aber nicht die Meise an, sondern das Rotkehlchen. Totales Chaos im Vogelreich. Und schuld daran ist, wie immer, der Mensch. Liebe V��gel, macht es doch ein bisschen besser als wir! Quelle: NZZ, 12.4.2009
---	---

Welche Aussage steht im Text?

Antwort w  hlen:

- ☐ In der Stadt gibt es weniger Nahrung als auf dem Land.
- ☐ Stadt- und Landv  gel verstehen einander gut.
- ☐ V  gel, die lange in der Stadt leben, singen lauter.
- ☐ Je lauter die Umgebung ist, desto kraftvoller pfeifen die V  gel.

Figure 2 : DaF Item LV

Avec les bonnes strat  gies, cet item a pu   tre r  solu par des   tudiants avec un niveau bas, car il suffit de retrouver les mots phare « Umgebung » et « je ... desto » dans le texte. De plus, les distracteurs 1 et 2 peuvent probablement   tre identifi  s comme faux sur la base de connaissances g  n  rales. Cet item ne peut   tre consid  r   comme valide car il devrait tester la compr  hension, et pas les strat  gies   voqu  es. Ce qui est enseign   en tant que strat  gie de lecture et compr  hension pendant les cours de langue se montre probl  matique dans une situation de test de classement. Ici, des r  gles centrales du d  veloppement d'items n'ont pas   t   respect  es. Une reformulation des distracteurs aurait   t   possible. Il a cependant   t   d  cid   de ne pas r  utiliser cet item pour les tests    venir. L'analyse de Rasch en elle-m  me n'a pas reconnu cet item comme probl  matique. Par contre, uniquement l'  cart entre le niveau   valu   par une didacticienne de langue   trang  re et le degr   de difficult   effectivement calcul   nous a rendus attentifs    cet item.

5. Mod  le de Rasch et discussion pour le d  veloppement des tests

Les tests se d  veloppent en proc  dant    une analyse continue de la validit   des items des sous-tests. Pour ce faire, nous voulons   tablir une analyse en boucle compos  e des   tapes suivantes :

- Une analyse    l'aide du mod  le de Rasch sert    identifier des items probl  matiques.
- L'  quipe de conception des items discute ceux-ci et propose des am  liorations.
- Les am  liorations sont impl  ment  es dans le test pour la session suivante.
- Le mod  le de Rasch sert    v  rifier si ces changements portent leurs fruits.
- L'  quipe de conception discute des r  sultats obtenus.

D'autres analyses, men  es en parall  le (satisfaction, changements de classe, ...), servent    garantir que le test continue    fonctionner suite    ces modifications et    identifier les points avec une marge d'am  lioration (Blons-Pierre et al., 2012 ; Gick, 2012).

Nous voulons mettre en place cette proc  dure en boucle pour am  liorer les items des tests de fa  on continue. Actuellement nous sommes en pleine phase d'impl  mentation d'une premi  re boucle. Nous avons fait une analyse des items, identifi   les items probl  matiques, et analysons ces items dans l'  quipe de conception, pour en faire ressortir les am  liorations qui seront impl  ment  es dans le test d  s le prochain semestre.

5.1 Identifier les items potentiellement probl  matiques

Le but de notre premi  re   tape est d'identifier quelques items probl  matiques qui seront analys  s et   ventuellement modifi  s. Comme nous l'avons vu plus haut, nos tests de classement remplissent d  j   leur fonction, et nous voulons   videmment que cela reste ainsi. La proc  dure d'am  lioration se faisant sur le test de classement en utilisation (pour g  n  rer des r  ponses en situation r  elle), nous limitons donc la proportion d'items modifi  s entre deux passations. Cela permet au test, qui a fait ses preuves sur le terrain, de continuer    atteindre son but (placer les apprenants dans des classes le plus homog  nes possibles), tout en permettant d'am  liorer les items.

L'analyse de Rasch nous aide    identifier les items potentiellement    modifier. Nous utilisons ici le mod  le de Rasch de fa  on pragmatique : il nous sert    identifier les items probl  matiques mais ne nous sert pas pour l'identification de la source du probl  me. Pour identifier ces items probl  matiques, nous utilisons principalement deux repr  sentations des r  sultats propos  s par le logiciel Winsteps : l'« item : map » et le tableau de sortie selon Item. Nous regardons d'abord toutes les donn  es ensemble, puis nous les s  parons selon les comp  tences test  es.

Les graphiques des « Item : Map » donne un premier aper  u de la position des items les uns par rapport aux autres selon leur degr   de difficult  . Nous regardons si des items se placent en dehors de la « zone » de difficult   pour laquelle ils ont   t   cr   s. Id  alement, avec la structure adaptative de nos tests, nous aimerions trouver, du bas vers le haut, d'abord les items des tests des niveaux A, ensuite, au milieu, ceux des niveaux B et, tout en haut les items des niveaux C. Si la majeure partie des items se place bien dans cette structure, on en trouve aussi qui sont plac  s plus haut ou plus bas que leur niveau id  al. Ces items sont retenus pour la discussion dans l'  quipe (  tape 2).

Avant cela, ces items sont regard  s dans les tableaux d'analyse par item pour savoir s'ils sont   galement identifi  s comme probl  matiques selon d'autres facteurs (p.ex. Pt.Measure, Infit, Outfit) et la courbe de score attendu (Expected Score ICC). Cela est souvent le cas. Ces tables de r  sultat permettent   galement de ressortir des Items n'attirant pas l'attention dans l'item map, mais pour lesquels d'autres indications font penser qu'ils sont probl  matiques : p.ex. un Pt.Measure bas, un Infit ou Outfit   lev   (ou bas). « Measure », « Point Measure » (Wright et Stone, 1979). Nous identifions ainsi un certain nombre d'autres items potentiellement probl  matiques, qui seront soumis    l'  quipe de conception des items.

Par exemple, pour les donn  es du test FLE des deux semestres 2011 (nous avons regroup   ces deux semestres pour obtenir un nombre suffisant de r  ponses), nous identifions une quinzaine d'items sur un total de 210.

5.2 Identification des probl  mes, proposition et impl  mentation des modifications

Dans cette phase, l'  quipe de conception des items va :

- Regarder les r  ponses des   tudiants    chaque item potentiellement probl  matique (r  sultats des tentatives dans Moodle)
- Discuter et poser des hypoth  ses sur : Est-ce que il y a effectivement un probl  me avec cet item ? Si oui : quelles pourraient en   tre les causes ? Si non, en quoi l'analyse l'identifie comme probl  matique ?
- Proposer des possibles am  liorations pouvant aller du changement d'une virgule    l'  limination/substitution compl  te de l'item.
- D  cider des items    changer en cherchant le maintien du fonctionnement des tests.

Nous sommes actuellement dans la phase d'identification des probl  mes et de proposition de modifications. Pour le test FLE nous avons par exemple commenc   une discussion sur certains items. Tous les items probl  matiques ont   t   ou seront discut  s de cette fa  on, et l'  quipe d  cidera lesquels modifier. Voici quelques exemples de discussion :

Un exemple d'item de compr  hension   crite :

- Question : Elle s'int  resse    la planche    voile depuis la pr  -adolescence.
- R  ponses : Vrai, Faux, La r  ponse n'est pas dans le texte.
- Analyse de Rasch : Cet item a un pt. mesure tr  s bas, et un infit haut (avec zsdt haut), la courbe ICC montre   galement un probl  me.

Hypoth  se formul  e :

- Probl  me avec formulation/interpr  tation: « LT monte premi  re fois » dans le texte    lire n'est pas   gale    « s'int  resse » dans la question.
- Possible solution: reformuler « int  resse ». Changer avec : « *****    la planche    voile ».

Ce changement va   tre impl  ment  .

Un exemple d'item de lexique et structure :

- Question : Il s'agit d'une phrase dans laquelle on cherche un accord verbal.
- R  ponses : a choisi, choisira, ait choisi, avait choisi.
- Analyse de Rasch : Cette question a un pt. mesure bas, et sa courbe montre   galement un probl  me. La distribution des r  ponses sur les distracteurs donne un indice : la majorit   des personnes choisissent une m  me mauvaise r  ponse (distracteur 3).

Hypoth  se formul  e :

- Item trop difficile, distracteur interpr  t   comme juste.
- Solutions possibles : a) Il s'agit d'un test de niveau   lev  , donc il doit   tre difficile. b) Il est trop difficile, on pourrait changer le distracteur en « choisissait ».

Ce changement demande qu'il soit encore discut   avant d'  tre impl  ment  .

5.3 Analyse avec les items modifi  s

Lorsque les modifications seront impl  ment  es dans les tests des prochaines sessions, nous pourr  s collecter une nouvelle s  rie de donn  es avec lesquelles nous effectuerons une nouvelle analyse sur la base du mod  le de Rasch. Nous pourr  s ensuite regarder avec l'  quipe si la modification a port   ses fruits. Si oui, on aura appris que l'hypoth  se sur le probl  me de l'item   tait fond  e. Autrement nous pourr  s d  cider si une modification additionnelle s'impose ou si l'item doit   tre substitu  . Dans les deux cas on aura un indice de l'efficacit   des modifications propos  es.

6. Conclusions et perspectives

L'analyse de Rasch nous appara  t utile pour la conception et le d  veloppement de tests adaptatifs. Elle contribue    la transformation rapide, pour les tests DaF et FLE, d'un test papier progressif en un test adaptatif en ligne valide en se servant de donn  es r  colt  es en situation r  elle. Dans une phase ult  rieure, cette analyse soutient   galement l'optimisation des tests en visualisant les degr  s de difficult  s des items les uns par rapport aux autres. De plus elle donne des indices sur la validit   de chaque item individuel.

Cette analyse comporte aussi des limites. Notre exp  rience montre notamment le besoin de continuer les analyses quantitatives tout en y associant d'autres m  thodes plus qualitatives pour d  velopper des tests de classement en langues   trang  res adaptatifs. Ainsi, l'analyse de Rasch ne permet pas de d  tecter tous les items « probl  matiques ». Le regard des experts en didactique des langues   trang  res reste incontournable. En effet, les r  flexions de ces experts sur les items

probl  matiques permettent d'expliciter les probl  mes li  s    certains items. Pour tirer avantage des forces des deux m  thodes, nous avons d  cid   de mettre en place une proc  dure en boucle, alliant analyse de Rasch et regard des experts.

Les premi  res exp  riences avec cette proc  dure portent leurs fruits. Nos perspectives de travail sont de la rendre plus syst  matique. Ensuite, il s'agira d'effectuer un bilan de la proc  dure, en termes des choses apprises sur le d  veloppement des items par les   quipes des tests (formation continue au sein du Centre de Langues).

7. Bibliographie

- Blons-Pierre, C. (2010). *Tensions between change and quality assurance : A dynamic necessary for innovation in university language centres/Tensions entre politique du changement et assurance qualit   : une dynamique n  cessaire pour l'  volution des centres de langues universitaires*. Communication pr  sent  e    la 11th International CercleS Conference, Helsinki, 4 September.
- Blons-Pierre, C. (2011). *L'impact du CECR/GeR/CEFR sur le fonctionnement d'un centre de langues dans une universit   bilingue suisse*. Communication pr  sent  e    la 4th International ALTE Conference, 7-9 July.
- Blons-Pierre, C., Kohler, P., Gick, C., Hoein, S. & Platteaux, H. (2012). Analyse qualitative des tests de classement en ligne pour les langues : acceptance et r  le de l'auto  valuation. Dans *Actes du XXIV  me Colloque de l'ADMEE Europe. L'  valuation des comp  tences en milieu scolaire et en milieu professionnel*, Luxembourg 11-13 janvier.
- Conseil de l'Europe. (2001). *Cadre europ  en commun de r  f  rence pour les langues : apprendre, enseigner,   valuer*. Paris : Didier.
- Gick, C. (2011). *Online-Einstufung und Online-Einschreibung in studienbegleitende Deutschkurse mittels adaptivem Einstufungstest: Entwicklungsschritte und Stand der Dinge*. Communication pr  sent  e au Bremer Symposium 2011 (Fremdsprachenzentrum der Hochschulen im Land Bremen), Bremen 4-5 M  rz.
- Gick, C. (2012). Einstufungstests im Spannungsfeld von Bologna-Reform, Referenzrahmen f  r Sprachen, Fachwissenschaft und Nutzern, In C. Blons-Pierre (dir.) : *Apprendre, enseigner et   valuer les langues dans le contexte de Bologne et du CECR / Sprachen lernen, lehren und beurteilen im Kontext von Bologna und dem GER*. Bern : Peter Lang, 189-217.
- Kohler, P. (2012). Gestion de l'auto-  valuation dans un dispositif de test en ligne: limites et perspectives. In C. Blons-Pierre (dir.), *Apprendre, enseigner et   valuer les langues dans le contexte de Bologne et du CECR / Sprachen lernen, lehren und beurteilen im Kontext von Bologna und dem GER* (pp. 219-240). Bern : Peter Lang.
- Kohler, P., Platteaux, H., & Blons-Pierre, C. (2010). *R  actions des   tudiants    l'usage des tests en ligne pour   valuer les comp  tences langagi  res*. Communication pr  sent  e au XXVI  me Congr  s AIPU. R  formes et changements p  dagogiques dans l'enseignement sup  rieur, Rabat 17-21 mai.
- Kohler, P., Platteaux, H., & Blons-Pierre, C. (2012). Un test de classement en ligne pour   valuer les niveaux de comp  tence et constituer des groupes classes. *Revue RIPES - Num  ro Sp  cial « Innover dans l'  valuation des apprentissages : pourquoi et comment ? »*, 27(2), R  cup  r   du site de la revue : <http://ripes.revues.org/532>.
- Linacre, J. M. (2000). *Computer-Adaptive Testing: A methodology whose time has come*. Chicago: MESA Psychometric Laboratory - University of Chicago.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Mahwah, NJ: Erlbaum.
- Platteaux, H., & Hoein, S. (2010). *Tests de positionnement en langues   trang  res sur Moodle*. Communication pr  sent  e    la Conf  rence Moodle Moot 2010, Troyes 28-30 juin.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research Copenhagen 1960, expanded edition with foreword and afterword by B.D. Wright. Chicago: The University of Chicago press.

Actes du 24^e colloque de l'Adm  -Europe
L'  valuation des comp  tences en milieu scolaire et en milieu professionnel
C. Gick, H. Platteaux, S. Hoein, C. Blons-Pierre, P. Kohler
Analyse de validit   des tests de classement en ligne pour les langues    l'aide du mod  le de Rasch

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Zweite,   berarbeitete und erweiterte Auflage, Bern, G  ttingen, Toronto, Seattle: Hans Huber.

Tagliante, C. (2005). *L'  valuation et le Cadre europ  en commun*. Paris: CLE International.

Wright, B. D., & Stone, M. H. (2004). *Making measure*. Chicago: The Phaneron Press.

ANALYSE QUALITATIVE DE TESTS DE CLASSEMENT EN LIGNE POUR LES LANGUES : ACCEPTANCE ET ROLE DE L'AUTOEVALUATION

Catherine Blons-Pierre*, **Patricia Kohler****, **Cornelia Gick*****, **Sergio Hoein******,
Hervé Platteaux*****

* Université de Fribourg (Suisse), catherine.blons-pierre@unifr.ch

** Université de Fribourg (Suisse), patricia.kohler@unifr.ch

*** Université de Fribourg (Suisse), cornelia.gick@unifr.ch

**** Université de Fribourg (Suisse), sergio.hoein@unifr.ch

***** Université de Fribourg (Suisse), herve.platteaux@unifr.ch

Mots-clés : tests de classement en ligne, coefficient Kappa de Cohen, analyse qualitative, acceptance

Résumé. Le Centre de langues de l'université bilingue de Fribourg en Suisse a mis en place avec la collaboration du Centre Nouvelles Technologies et Enseignement des tests de classement (placement tests) en ligne et adaptatifs pour l'allemand, l'anglais, le français et l'italien, langues étrangères. Ces tests font régulièrement l'objet d'analyses quantitatives et qualitatives.

L'étude présente est destinée, d'une part, à réfléchir sur les avantages et les contraintes liés à l'automatisation des tests et de l'inscription et, d'autre part, à présenter une analyse statistique des tests pour le français et l'allemand faite à l'aide du coefficient Kappa de Cohen, qui permet d'évaluer le degré d'accord entre l'autoévaluation effectuée par les testés et le résultat de l'évaluation proposé par le test de classement. Cette analyse statistique est complétée par une analyse qualitative de ces tests (acceptance), effectuée à l'aide des données recueillies auprès d'un échantillon d'étudiants au moyen d'entretiens.

1. Introduction

En 2008, le Centre de langues de l'Université de Fribourg (Suisse) a décidé de développer un dispositif de tests de classement (placement tests) en ligne et adaptatifs, destinés à constituer des groupes classes homogènes. Ces tests en ligne ont été implémentés sur la plateforme d'enseignement Moodle et complétés par un système d'inscription en ligne, qui automatise complètement le processus d'inscription et d'admission aux cours de langues.

Un certain nombre de questions ont été prises en considération par le Centre de langues, le Rectorat et le Centre Nouvelles Technologies et Enseignement avant de se lancer dans l'aventure organisationnelle, pédagogique et technologique que représente le passage du test papier progressif au test en ligne adaptatif, comme par exemple :

- Quels peuvent être les avantages d'une telle (r)-évolution au regard de l'institution académique, de l'administration, du personnel enseignant et également de l'utilisateur ?
- Y a-t-il des limites ou du moins des contraintes liées à la mise en place d'un tel dispositif ?
- Comment peut-on mesurer la qualité des tests et évaluer leur acceptation auprès des étudiants ?
- L'autoévaluation des étudiants pourrait-elle suffire pour constituer des classes de niveaux homogènes rendant ainsi les tests de classement inutiles ?

Ce sont ces questions primordiales, qui touchent à la qualité du dispositif et à la validité de nos tests en ligne, que nous allons traiter dans cette contribution.

2. Avantages et contraintes liés à l'automatisation des tests et des inscriptions

Les avantages du test en ligne par rapport au test papier sont, pour l'institution, la rapidité du processus et la disponibilité des personnels enseignants et administratifs, qui se trouvent déchargés, grâce à l'automatisation du processus, de l'organisation lourde liée à la passation des tests de classement papier : manutention du matériel, présence de surveillants, corrections des tests et traitement des inscriptions en différé, après les corrections manuelles. Pour les enseignants et les étudiants, sur un plan pédagogique, le test en ligne donne l'assurance que les étudiants bénéficient exactement des mêmes modalités d'évaluation. Pour les étudiants, s'ajoute l'avantage du choix du moment et de l'endroit où ils décident de passer ce test.

Néanmoins, pour que l'évaluation soit valide et fidèle, il convient de mettre en place des processus capables de mesurer les différents paramètres en œuvre dans les tests en ligne et permettant de faire évoluer la qualité de cette évaluation. Depuis le début, nos tests sont analysés à l'aide de statistiques quantitatives et d'analyses qualitatives. La validité et la fidélité des items des différents tests sont analysées à l'aide de la théorie de réponse aux items (modèle de Rasch). Ces analyses nous ont permis de faire évoluer les banques d'items et d'assurer ainsi une progression de la qualité à ce niveau¹.

L'aspect de la qualité des tests pose également d'autres questionnements, qui requièrent des types d'analyses différents. Les deux points importants à examiner lorsque l'on passe des tests papier aux tests en ligne paraissent être, d'une part, l'acceptance de ce nouveau dispositif par les usagers : ergonomie du test, conditions de passation, et, d'autre part, la mesure de l'écart entre l'autoévaluation et l'évaluation faite par le test en ligne. Il existe, en effet, depuis le développement du Cadre européen commun de référence pour les langues (CECR), des descripteurs permettant aux individus désireux d'apprendre une langue d'évaluer leur niveau de façon autonome. Dès lors, il est légitime de s'interroger sur l'utilité de faire des tests de classement si l'autoévaluation est suffisamment fiable ; il est donc important de vérifier quelle est l'intensité de l'accord constaté entre l'autoévaluation et le résultat des tests.

L'étude qui suit est destinée à présenter l'analyse de ces tests faite à l'aide du coefficient Kappa de Cohen, qui permet d'évaluer le degré d'accord entre l'autoévaluation faite par les testés et le résultat de l'évaluation proposé par le test de classement pour la compréhension orale, la compréhension écrite ainsi que le lexique et les structures de la langue. Cette analyse statistique sera, dans un deuxième temps, complétée par une évaluation d'acceptance de ces tests de classement, effectuée à l'aide des données recueillies auprès des étudiants au moyen d'entretiens. Ces deux types d'analyse sont présentés pour les tests de classement utilisés pour le français langue étrangère et l'allemand langue étrangère. Ils sont complémentaires.

3. Analyse qualitative des tests en ligne à l'aide de l'indice Kappa de Cohen

3.1 Méthodologie pour mesurer l'écart entre autoévaluation et résultats obtenus aux tests

Nous sommes partis du recueil des données du semestre d'automne 2010 complétés par les données du recueil du semestre de printemps 2011 pour les tests de français (FLE) et allemand

¹ Voir in *L'évaluation des compétences en milieu scolaire et en milieu professionnel*, Analyse de validité des tests de classement en ligne pour les langues à l'aide du modèle de Rasch, Gick, Platteaux, Hoein, Blons-Pierre, Kohler.

(DaF) langues étrangères. Ces données permettent de comparer les résultats de l'autoévaluation avec les résultats des tests. Dans notre cas, on est amené à comparer deux jugements : celui de l'étudiant qui s'autoévalue et celui du test. Pour ce faire nous avons choisi d'utiliser l'indice Kappa de Cohen car nous traitons des variables nominales et non des variables ordinales².

Il convient en outre de préciser que la fonction de l'autoévaluation diffère dans les tests pour le français (FLE) et pour l'allemand (DaF). Dans le test FLE, l'autoévaluation joue le rôle de pré-test proposé au début de chaque sous-test (compréhension orale, compréhension écrite, lexique et structure) alors que, dans le test DaF, le pré-test est assuré par un sous-test de lexique et structure et l'autoévaluation se trouve à la fin du test.

3.2 Les différentes étapes de l'analyse

Il s'agit tout d'abord de préparer les données en repérant les valeurs manquantes ou aberrantes. Il faut ensuite créer la matrice de confusion ou tableau de contingence. Cette matrice s'obtient en utilisant un tableau croisé dynamique. Par exemple, si l'on considère les données permettant de croiser les résultats obtenus par les étudiants pour la compréhension écrite au test de français du semestre d'automne 2011 (SA 2011) avec leur autoévaluation pour cette même compétence, on obtient le tableau suivant :

Niveau CE	A1 autoév.	A2	B1	B2	C1	Total
A1 test	16	15	3	0	0	34
A2 test	2	37	49	20	3	111
B1 test		2	23	24	2	51
B2 test		2	3	11	1	17
C1 test			1	2	1	4
Total	18	56	79	57	7	217

Tableau 1 : Tableau de contingence CE SA 2011 test FLE

D'après le tableau 1, on constate que sur les 217 étudiants qui ont passé le sous-test de compréhension écrite pour le français au semestre d'automne 2011 :

- 88 se sont autoévalués en accord avec le niveau identifié par le test
- 117 se sont surévalués par rapport au test
- 12 se sont sous-évalués.

Les données présentées ci-dessus sont globalement équivalentes à un pourcentage de 40.55% d'autoévaluations en accord avec le sous-test de compréhension écrite et à un pourcentage de 59.45% d'autoévaluations en désaccord avec le sous-test. On peut également ajouter que la majorité des autoévaluations en désaccord avec le sous-test présente un écart égal en valeur absolue à 1 (voir tableau 2).

Sous-évalués	Nombre		Sur-évalués	Nombre
-1	9		+1	89
-2	3		+2	25
-3	0		+3	3
Total	12			117

Tableau 2 : Nombre de sous-évalués et sur-évalués pour CE SA 2011 test FLE

² Pour les variables ordinales, on utilisera, par exemple, l'indice W de Kendall.

Comme le montre le tableau 2, les désaccords entre autoévaluation et résultat du sous-test de CE pour la session du SA 2011 se répartissent principalement en $N+1 = 89$ et $N-1 = 9$, soit un total de 98 pour l'écart 1 en valeur absolue et un pourcentage de 76% par rapport au nombre total des écarts constatés entre autoévaluation et résultat du sous-test.

Cependant, cette première analyse, si elle donne des indications précieuses quant à la surévaluation et à la sous-évaluation, ne tient pas compte de la proportion d'accord qui peut être imputée au seul fait du hasard.

Il est donc important de recourir à un type d'analyse statistique prenant en compte l'aléatoire. Le calcul du coefficient Kappa de Cohen permet cette prise en compte. En effet, ce coefficient est un indice statistique variant entre 1 et -1, qui est utilisé pour évaluer le degré d'accord (concordance) entre deux juges par rapport à une variable nominale et non ordinale. L'accord maximal est obtenu lorsque l'indice est égal à 1, le désaccord total lorsqu'il est égal à -1. Ce coefficient peut s'interpréter comme la proportion d'éléments classés de la même manière par deux juges, ici, l'étudiant et le test ; ce qui le différencie du simple calcul d'une proportion, c'est qu'il introduit une correction pour prendre en compte le fait qu'une certaine proportion d'accord peut être imputée au hasard. Le coefficient κ se calcule en appliquant la formule suivante :

$$K = (P[A] - P[H]) / (1 - P[H])$$

avec :

$P[A]$ = proportion d'accords

$P[H]$ = proportion de cas où selon la théorie des probabilités, on peut s'attendre à un accord simplement dû au hasard³.

3.3 Les résultats de l'analyse avec le coefficient K de Cohen

Pour analyser les résultats obtenus avec le coefficient K de Cohen, on peut se référer à la table de Landis et Koch (voir tableau 3), qui étalonne les résultats de ce coefficient de la manière suivante⁴ :

Intensité de l'accord	Coefficient Kappa
Excellent	>0,80
Bon	0,80-0,61
Modéré	0,60-0,41
Médiocre	0,40-0,21
Mauvais	0,20-0,0
Très mauvais	<0,0

Tableau 3 : table de Landis et Koch

Par exemple, si l'on applique le calcul du coefficient K de Cohen et la grille de lecture de Landis et Koch au test pour le français, on obtient les résultats suivants :

	K semestre automne 2010	K semestre printemps 2011
CO compréhension orale	0.22 médiocre	0.26 médiocre
CE compréhension écrite	0.21 médiocre	0.22 médiocre
LS lexique et structure	0.40 médiocre	0.31 médiocre

Tableau 4 : Coefficient Kappa pour test FLE semestres automne et printemps 2011

³ Pour plus d'informations sur K, voir http://www.irdp.ch/edumetrie/lexique/coef_kappa.htm.

⁴ Pour la table de Landis et Koch, voir http://kappa.chez-alice.fr/Kappa_2juges_Def.htm

D'après les données présentées dans le tableau 4, on peut dire que l'accord entre les variables « autoévaluation » et « résultat du test » pour le test FLE est médiocre pour la compréhension orale, la compréhension écrite et lexicale et structure, de façon constante, pour les deux sessions de test : semestre d'automne 2010 et semestre de printemps 2011.

Etant donné que l'autoévaluation figure dans le test FLE au début des sous-tests CO, CE et LS, on pourrait faire l'hypothèse que, n'étant pas suffisamment sensibilisé à la description des niveaux de langue, l'étudiant ne consacre pas suffisamment de temps à s'autoévaluer car il n'en perçoit pas clairement les enjeux. Cette attitude paraît, du reste, être corroborée par le fait que, au fur et à mesure de l'avancée dans le test, les étudiants ont tendance à s'autoévaluer de plus en plus rapidement, ils s'appuient en fait sur l'autoévaluation qu'ils ont faite pour le sous-test précédent et la reconduisent de sous-test en sous-test, alors même que les descripteurs sont différents pour chacun des sous-tests proposés. En effet, les habiletés et connaissances attendues diffèrent selon les niveaux et les compétences testées. Ainsi, un étudiant peut avoir un degré de compétence élevé pour la compréhension orale et un degré de compétence moindre pour la compréhension écrite et lexicale et structure. Ce que le test cherche à établir c'est un profil moyen du testé, il est donc important de proposer au début de chaque sous-test une nouvelle autoévaluation qui sert d'entrée dans le sous-test adaptatif.

Ceci étant posé, il paraît maintenant intéressant d'examiner si l'on obtient les mêmes écarts entre autoévaluation et résultats du test pour le test DaF, qui, lui, utilise l'autoévaluation en fin de passation du test avec un statut facultatif et, uniquement, pour la compréhension orale et la compréhension écrite. Pour ce faire, il est possible de faire une analyse comparative des coefficients K de Cohen obtenus pour les tests FLE et DaF, à l'aide du tableau suivant.

	Coefficient K Test FLE	Coefficient K Test DaF
CE SA 2010	0.21	0.14
CO SA 2010	0.22	0.21
CE SP 2011	0.22	0.25
CO SP 2011	0.26	0.35

Tableau 5 : coefficient K pour test FLE et test DaF

Si l'on compare, d'après le tableau 5, la qualité de l'accord entre autoévaluation et résultats des tests pour les deux sessions du SA 2010 et SP 2012, on constate que le test FLE montre des résultats stables pour les deux semestres, en revanche, les variations sont plus importantes pour le test DaF : l'accord qui était jugé mauvais d'après l'indice K pour la CE au SA 2010 devient médiocre pour le SP 2011 alors que l'accord pour la CO qui se situait plutôt à la limite mauvais/médiocre au SA 2010 devient clairement médiocre pour le SP 2011. Globalement, cela traduit une légère augmentation de la qualité de l'accord entre autoévaluation et résultats du test DaF entre les deux sessions.

Toutefois, que ce soit pour le test FLE ou le test DaF, l'accord entre l'autoévaluation des étudiants et les tests ne dépasse jamais le degré de qualité médiocre, ce qui signifie qu'en aucun cas on ne pourrait, à l'heure actuelle, se baser uniquement sur l'autoévaluation pour placer les étudiants dans les groupes classes si l'on a l'objectif de constituer des groupes relativement homogènes.

Il semble donc important, afin de pouvoir mieux utiliser l'autoévaluation, par exemple, dans la perspective de raccourcir le temps de passation des tests adaptatifs qui ont comme point d'entrée l'autoévaluation, de réfléchir à la manière de permettre à l'étudiant de mieux s'autoévaluer. L'analyse statistique réalisée à l'aide du coefficient K de Cohen nous a permis de dresser un constat prenant en compte la part de l'aléatoire, il nous faut maintenant recourir à un autre type d'approche, l'analyse qualitative, qui a pour fonction d'analyser finement les mécanismes à l'œuvre dans le processus d'autoévaluation à travers les témoignages de ceux qui la pratiquent.

4. L'analyse qualitative des tests

4.1 Méthodologie utilisée pour évaluer l'acceptance des tests

Nous avons procédé à des entretiens de type semi-directifs en décembre 2011 auprès d'étudiants qui avaient passé le test FLE en septembre 2011. Comme le modèle de test proposé au semestre d'automne 2010 n'a subi que de faibles modifications d'ordre formel par rapport à celui proposé en automne 2011, il nous a été ainsi possible de recueillir des données actualisées. L'objectif reste de mesurer le degré d'acceptance de l'évaluation des trois compétences (compréhension orale, compréhension écrite, lexique et structure) donnée par le test par rapport à l'autoévaluation. Quels sont les ressentis de l'usager face à l'offre de cours qui découle du résultat du test de positionnement et les représentations qu'il a de son niveau à travers l'autoévaluation ?

L'étudiant est au centre de notre dispositif, en tant qu'acteur dont nous observons les réactions dans la construction des stratégies d'apprentissage à travers le discours qu'il porte sur les représentations de son niveau et les écarts entre niveau effectif et niveau ressenti, ou autoévaluation. Nous partons donc du postulat que l'étudiant possède des compétences d'agir et de réagir par rapport à l'environnement du test en ligne et qu'il convient de les rechercher à des fins d'analyse et d'amélioration du dispositif d'évaluation et d'apprentissage. Cette interaction fait de l'étudiant bien plus qu'un agent, mais un médiateur dans la mesure où il participe d'emblée à la co-construction de son pré-contexte d'étude.

Cette analyse, de type qualitatif, s'effectue sur un échantillon restreint de population concernée par le seul test FLE afin de révéler les marqueurs précis des manifestations des ressentis et de rendre transparents la complexité des pratiques d'apprentissage et le rôle de l'implicite. L'échantillon des personnes interviewées est composé d'un groupe de trente-trois étudiants volontaires répartis comme suit : 10 étudiants qui suivent un cours objectif A1.1, 8 étudiants en cours objectif A1.2, 11 au niveau objectif B1 et 4 au niveau objectif B2. Dix à douze semaines après la passation du test en ligne, ces étudiants ont été invités à se présenter à l'entretien d'une durée d'environ 30 minutes, et ce, individuellement ou en groupe. La langue de l'entretien a pu varier en fonction des niveaux car il était plus facile pour l'intervieweur de laisser l'interviewé choisir entre l'anglais, l'allemand et le français afin de récolter des données les plus fiables possibles dans l'expression du ressenti des participants majoritairement anglophones ou germanophones. Les débutants n'auront ainsi jamais fait recours au français tandis que les niveaux avancés auront facilement pratiqué le passage d'une langue à l'autre ou code-switching. Le canevas d'entretien a mis l'accent sur les trois thématiques suivantes : le déroulement du test (lieu, durée, conditions-cadre), l'utilisation et la lecture des grilles d'autoévaluation (lecture, compréhension, utilisation, utilité), et le degré d'acceptance du niveau du cours suivi et déterminé par le test en ligne, par rapport à celui de l'autoévaluation.

Bien que déjà globalement émises lors d'analyses antérieures (Kohler, P., Platteaux, H., & Blons-Pierre, C., 2012), nos hypothèses sont déterminées par la volonté d'éviter tout biais méthodologique que constituerait l'influence des recherches déjà ou parallèlement menées, afin de mettre en lumière le seul point de vue de l'étudiant à un moment donné. Nous avons donc formulé les trois hypothèses suivantes :

- Les étudiants auraient tendance à sous-évaluer ou surévaluer leur niveau de langue par rapport au niveau effectif révélé par le test en ligne.
- Le niveau de césure entre la sous-autoévaluation et la sur-autoévaluation par l'étudiant de son niveau de langue par rapport au niveau effectif révélé par le test en ligne, serait lié à une représentation du savoir langagier.

- Les stratégies de lecture des descripteurs détermineraient l'utilisation, et donc la fiabilité de l'autoévaluation dans un dispositif de test de classement en ligne pour les langues.

4.2 Résultats

Les entretiens ont permis de dégager des résultats portant sur les deux principaux groupes, d'une part les débutants A1.1 et faux-débutants A1.2, et d'autre part les intermédiaires B1 et B2. Ces résultats peuvent être répertoriés en deux grandes catégories que constitue le double degré d'acceptance par l'étudiant, tout d'abord celui de l'autoévaluation et de l'utilisation des descripteurs du CECR, et, d'autre part, celui de la corrélation entre cette autoévaluation et le niveau effectif attribué par le test en ligne.

Comment tout d'abord l'étudiant réagit-il à l'autoévaluation et à l'utilisation des descripteurs du CECRL ? Contre toute attente, c'est l'étudiant débutant qui se livre à toute une série de réflexions sur la durée du test et le seuil de césure :

« The test for beginners is too short. »

« The test for beginners may not be able to distinguish A1.1 and A1.2. »

Si le manque de descripteurs pour les débutants est également relevé par plusieurs étudiants, ces derniers n'hésitent pas non plus à s'exprimer sur la place où devrait se situer une éventuelle autoévaluation dans le test proposé :

« It's important to have a self-assessment test to know in what level we are. »

« I would like a little of self-assessment. »

« Some short assessment is nice. »

« I missed the feed-back directly after the test. »

Ces réflexions indiquent combien il est important de valoriser l'apprenant-agent de son apprentissage dès le test puisqu'il réclame de façon implicite le même traitement que tout autre étudiant. Il prétendrait ainsi posséder des compétences que le test ne lui aurait pas permis de mettre en œuvre ni dans l'autoévaluation, ni dans l'évaluation. Il resterait à identifier la nature de ces compétences nouvelles et à en produire le mode d'évaluation adéquate, au-delà des catégories préexistantes. Le développement de nouvelles pratiques qui se construisent par les réalités sociales de l'apprenant permettraient-elles de trouver la clé de ces nouveaux savoir-faire langagiers ?

D'autre part, les remarques sur la durée du test pour les débutants confirment certains résultats antérieurement obtenus (Kohler, 2011) . Le sérieux du test serait proportionnel à sa longueur et à sa durée qui contribuent à son degré de fiabilité et d'acceptance. Nous pouvons nous interroger sur les motivations réelles d'une telle demande de la part de débutants complets dont le savoir en français reste entièrement à construire et n'a donc pas nécessairement besoin d'être autoévalué. Mais si l'étudiant le revendique, c'est qu'il semble imprégné des valeurs véhiculées par la société du savoir à laquelle il appartient, le mode d'apprentissage de ces savoirs et de leur évaluation dans la course au « capital linguistique » (Bourdieu, 2001).

Le lien entre ces deux derniers résultats chez les débutants, soit la demande d'évaluation de nouveaux savoir-faire langagiers à l'intérieur d'un test plus long, reste néanmoins inattendu. Cela prouve l'importance et le rôle de l'écoute de l'utilisateur dans l'élaboration de nouveaux outils d'évaluation comme un test en ligne afin de prendre en compte ces implicites où se réfugient des croyances socioculturelles impliquant des pratiques d'apprentissage.

L'évaluation des compétences en milieu scolaire et en milieu professionnel

Si nous considérons maintenant le groupe des interviewés faux débutants dont nous présumons une auto-sous-évaluation, leurs témoignages révèlent un autre élément très nouveau sur le processus même par lequel va se déconstruire ou reconstruire le capital linguistique de l'étudiant. Si la représentation de la langue étrangère se trouve réduite au seul bagage lexical, l'école est souvent citée comme le lieu d'acquisition du savoir, mais ce mot-clé est fréquemment associé à l'oubli sous la forme d'une concession rectificative :

« A l'école, j'ai eu du français, mais j'ai oublié tout. »

« I have learned it at school over three years but I have forgotten a lot of vocabulary. »

Pour d'autres étudiants qui ne font pas référence au passé, le test en ligne est perçu comme une ouverture valorisante qui permet la reconnaissance de savoirs jusque-là méconnus de l'apprenant lui-même, surpris de sa réussite :

« Je suis entrée débutant complet, mais je suis A1.2. » (sourire)

« Je suis entré débutant complet mais suis inscrire pour A1.2. » (sourire)

« Although I didn't take any french course, I know some basic words A1.2. »

Ce dispositif déclenche une attitude de confiance et de satisfaction visible (sourires, bonne humeur etc.). Cette base de départ peut favoriser le développement de nouvelles stratégies d'apprentissage telles que des réflexions sur le savoir et le savoir-faire, sur les interactions entre l'apprenant et l'ordinateur, la découverte de nouvelles façons d'apprendre et les nouveaux rôles complémentaires attribués à l'enseignant comme facilitateur et médiateur dans la fabrication des items de l'évaluation.

Concernant notre deuxième groupe d'étudiants pour lesquels les résultats obtenus ont été significatifs, soit ceux obtenus par les intermédiaires B1/B2, plusieurs pistes méritent d'être retenues : tout d'abord, seules deux étudiantes avaient connaissance de ces outils d'autoévaluation que constituent les descripteurs. Les autres étudiants qui les consultaient pour la première fois, n'ont été aucunement gênés dans leur consultation. D'autre part, tous les étudiants de ce groupe se sont surévalués dès le départ, en compréhension orale. L'origine de ce phénomène semble essentiellement liée à leur mode d'apprentissage du français, fondé sur les approches communicatives. Enfin, ils ont été sollicités à trois reprises pour l'autoévaluation, correspondant à la phase préliminaire de chaque palier du test, soit avant la compréhension orale, la compréhension écrite, et le test de lexique et structure de la langue. Or, la durée moyenne que les étudiants interviewés consacrent à la lecture des descripteurs avant chacune de ces phases, est de plus en plus courte : environ 50 secondes pour la compréhension orale pour passer à seulement 11 secondes en lexique et structure. Cette lecture des descripteurs est donc biaisée par le phénomène-tremplin déjà observé depuis 2010 (Kohler, Platteaux, Blons-Pierre, 2012). Ils ne lisent pas réellement les descripteurs proposés pour l'autoévaluation dont ils retiennent les premières et dernières lignes par compétence, sans oublier l'effet-balayage d'une lecture syntagmatique :

« Ich habe überflogen... »

« c'est trop compliqué... »

Certains interviewés, replacés en situation de consultation des descripteurs devant l'écran d'ordinateur, citent les expressions qu'ils lisent vraiment et qui leur apparaissent comme fondamentales pour déterminer leur niveau : « basic words », « radio-télévision », « AlltagsBerufssprache », « Reisen ». Les mots-clés retenus sont donc directement liés au mode de vie et de communication de l'utilisateur du test en ligne.

Tous les étudiants, depuis le niveau A1.2 jusqu'à B2, ont manifesté leur satisfaction d'avoir été consultés avant le début du test par l'autoévaluation dont ils ont bien compris les enjeux et l'utilité dans le dispositif : ils gagnent du temps et ils maîtrisent leur apprentissage. Malgré la surévaluation constatée pour la totalité des étudiants B1/B2 interrogés, ils n'ont pas été insensibles à la réflexion sur leurs apprentissages, déclenchée par l'autoévaluation. Le seul point reste la difficulté de lecture de descripteurs parfois trop longs, et qui mériteraient d'être adaptés pour un meilleur déroulement du test en ligne.

Il convient donc de retenir trois points fondamentaux concernant le degré d'acceptance de l'auto-évaluation en début de test de classement en ligne pour les langues : avant toute lecture, l'étudiant possède déjà un « capital linguistique » qui a construit ses choix et représentations de son niveau de langue. La lecture des descripteurs va donc s'opérer en fonction de la recherche de ses repères préconstruits que la présentation à l'écran participe à reproduire visuellement sous la forme d'un choix préétabli entre trois degrés, débutants, intermédiaires et avancés. Ces comportements induits chez l'étudiant vont être renforcés par le dispositif lui-même, mais aussi par des mécanismes comme la lecture « balayage » où les mots-clés sont également déterminés par des codes étroitement liés à des pratiques culturelles. Certains éléments sur les représentations d'une sous-évaluation des niveaux faux débutants et celles d'une surévaluation des niveaux intermédiaires, dépendraient aussi des pratiques d'apprentissage en relation avec la société du savoir à laquelle appartient l'étudiant.

Enfin, tous les étudiants semblent séduits par la démarche de l'auto-évaluation dans le processus de classement et la convivialité du système proposé : avant le cours, ils feraient partie d'une communauté d'apprentissage par Internet.

5. Conclusion

En conclusion, on peut dire que les analyses faites avec le coefficient Kappa de Cohen ont permis de constater que le degré d'accord entre les autoévaluations et les résultats des tests de classement en ligne ne dépassait jamais la qualité d'un accord médiocre pour les tests français et allemand, et ce, pour deux sessions de tests. Ce constat conduit à considérer que l'autoévaluation faite sur la base des descripteurs du CECR ne peut, dans l'état actuel, se substituer au test de classement.

Les raisons de l'écart entre autoévaluation et résultat des tests sont à chercher du côté des personnes testées, qui sont à même d'explicitier les stratégies mises en œuvre dans l'autoévaluation et dans le test en ligne. Une première analyse qualitative faite sur la base d'entretiens sur une population très réduite a livré quelques pistes. Il convient, cependant, d'en noter les limites puisque cette analyse s'est effectuée sur un faible échantillonnage. Cette analyse qualitative ainsi que les analyses statistiques menées à l'aide du coefficient Kappa de Cohen mériteraient d'être étendues à des populations plus importantes et sur plusieurs sessions de tests de façon à pouvoir vérifier les variations du coefficient Kappa de Cohen et la pertinence des hypothèses émises dans l'analyse qualitative. Il reste également à mettre en place des stratégies et des techniques visant à améliorer l'accord entre autoévaluation et résultats des tests et à vérifier leur efficacité sur le terrain.

Références et bibliographie

- Bétrancourt, M. (2007). L'ergonomie des TICE : quelles recherches pour quels usages sur le terrain ? In Charlier, B. et Peraya, D. (Eds), *Regards croisés sur la recherche en technologie de l'éducation*. Bruxelles : De Boeck, tecfa.unige.ch/perso/mireille/papers/Betrancourt08_REF.pdf.
- Blons-Pierre, C. (2010). Tensions between change and quality assurance : A dynamic necessary for innovation in university language centres/Tensions entre politique du changement et assurance qualité :

L'évaluation des compétences en milieu scolaire et en milieu professionnel

une dynamique nécessaire pour l'évolution des centres de langues universitaires. Paper presented at the 11th International CercleS Conference, Helsinki, 2-4 September.

Blons-Pierre, C. (2011). L'impact du CECR/GeR/CEFR sur le fonctionnement d'un centre de langues dans une université bilingue suisse. Paper presented at the 4th International ALTE Conference, 7-9 July.

Bourdieu P. (2001). *Langage et pouvoir symbolique*, Paris: Seuil.

Conseil de l'Europe. (2001). *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Paris: Didier.

Gick, C. (2011). Online Einstufungstest und Online-Einschreibung in studienbegleitende Deutschkurse: Entwicklungsschritte und Stand der Dinge. Paper presented at the Bremer Symposium 2011 (Fremdsprachenzentrum der Hochschulen im Land Bremen), Bremen 4-5 März.

Gick, C. (2012). Einstufungstests im Spannungsfeld von Bologna-Reform, Referenzrahmen für Sprachen, Fachwissenschaft und Nutzern. In C. Blons-Pierre (éd. / Hrsg.), *Apprendre, enseigner et évaluer les langues dans le contexte de Bologne et du CECR / Sprachen lernen, lehren und beurteilen im Kontext von Bologna und dem GER* (pp. 189-217). Bern : Peter Lang.

Kohler, P., Platteaux, H., & Blons-Pierre, C. (2012). Un test de classement en ligne pour évaluer les niveaux de compétence et constituer des groupes classes. *Revue RIPES - Numéro Spécial "Innover dans l'évaluation des apprentissages : pourquoi et comment ?"*, 27(2), <http://ripes.revues.org/532>.

Kohler, P. (2012). Gestion de l'auto-évaluation dans un dispositif de test en ligne: limites et perspectives. In C. Blons-Pierre (éd. / Hrsg.), *Apprendre, enseigner et évaluer les langues dans le contexte de Bologne et du CECR / Sprachen lernen, lehren und beurteilen im Kontext von Bologna und dem GER* (pp. 219-240). Bern : Peter Lang.

Platteaux, H., & Hoein, S. (2010). Tests de positionnement en langues étrangères sur Moodle. Paper presented at the Moodle Moot 2010, Troyes 28-30 juin.

Les compétences fondamentales et le développement : peut-on évaluer les systèmes éducatifs par le niveau d'acquisition homogène d'un bloc de compétences de base ?

Nadir ALTINOK* et Jean BOURDON**

* Université Paul Verlaine (Metz) / IREDU, nadir.altinok@gmail.com

** CNRS / IREDU, jean.bourdon@gmail.com

Problématique

La logique de l'éducation universelle suggère d'utiliser les concepts de compétences de base et des compétences dans le sens des capacités. Les compétences de base, leur degré de maîtrise plus ou moins également réparti dans la population vont contribuer à améliorer la capacité d'une population et en particulier les plus pauvres à exercer un certain degré de contrôle sur leur propre vie. Par exemple, prendre part avec d'autres dans les aspirations et choix qui influent sur les contextes de leur vie et d'envisager un futur alternatif pour eux et pour leurs familles (Sen, 1999). Cela s'accorde totalement avec le sous-objectif 6 de l'EPT concernant la qualité de l'éducation : "améliorer sous tous ses aspects la qualité de l'éducation dans un souci d'excellence, de façon à obtenir pour tous des résultats d'apprentissage reconnus et quantifiables – notamment en ce qui concerne la lecture, l'écriture et le calcul et les compétences indispensables dans la vie courante" (Unesco, 2000 ; souligné par nos soins).

L'enjeu d'aborder l'acquisition des compétences de base pour le développement durable est la poursuite du défi posé par la conférence de Jomtien, puis celle de Dakar. Y a été défini les besoins éducatifs fondamentaux comme outils d'apprentissage essentiels (tels que l'alphabétisation, expression orale, calcul, et la résolution de problèmes) et les contenus éducatifs fondamentaux (connaissances, compétences, valeurs et attitudes) requis par les jeunes adultes pour vivre et travailler dans la dignité, prendre des décisions éclairées, comme citoyens, et de continuer à se perfectionner par l'apprentissage tout au long de leur vie.

S'appuyant sur les débats au sein de sa zone d'intervention, l'OCDE (OCDE, 2005:5) présente une typologie des compétences de base en trois grandes catégories :

- 1- Les individus ont besoin pour être en mesure d'utiliser un large éventail d'outils pour interagir efficacement avec l'environnement : suivant des axes physiques tels que la technologie de l'information et d'autres socioculturels tels que l'utilisation et la maîtrise de la langue des échanges. Ils ont besoin de bien comprendre ces outils d'expression pour les adapter à leurs propres fins et d'utiliser ces outils de manière interactive.
- 2- Dans un monde toujours plus interdépendant, les individus ont besoin d'être en mesure de dialoguer avec les autres, et depuis qu'ils vont rencontrer des gens de divers horizons, il est important qu'ils soient capables d'interagir dans des groupes hétérogènes.
- 3- Troisièmement, les individus doivent être en mesure d'assumer la responsabilité de la gestion de leur propre vie, de situer leurs choix dans le contexte social le plus large et pouvoir agir de manière autonome.

Ces trois approches des compétences peuvent se résumer par une dotation de savoirs individuels à l'image du capital humain incorporé, mais aussi et surtout par des capacités à interagir dans la société ; cette dernière peut de ce fait atteindre des objectifs de plus en plus ambitieux d'efficacité que ce soit au niveau de la croissance ou des enjeux de bonnes gouvernances. Cela, il faut bien le reconnaître trace une situation optimale ; l'exemple au quotidien du développement montre que l'éducation peut-elle même tomber dans le piège de pauvreté. Cette situation pénalisante, fort bien décrite par Pritchett (2001), souligne que le lien éducation croissance est inopérant. Cela peut venir à l'évidence d'une non-adéquation entre les savoirs

enseignés et ceux qui sont demandés en rapport avec les emplois offerts dans l'économie. Plus redoutable est le piège dans lequel peut tomber le système éducatif. Si l'enseignement de base fonctionne en tant qu'organisation, le service offert peut-être de si médiocre qualité que les savoirs ne sont pas véritablement acquis. Les familles conscientes de cet état n'attachent plus de valeurs au système éducatif. La demande d'éducation plafonne ou décroît, certains se détournent vers un secteur privé de l'éducation, ce qui tend à accroître les inégalités... Le développement de l'éducation peut donc être confronté à la question de la qualité des savoirs dans son expansion.

Dans ce contexte l'objectif de l'éducation universelle serait bien évidemment lié à la scolarisation de tous les enfants comme condition nécessaire, mais de plus l'assurance de cet objectif et son caractère pérenne imposeraient de s'assurer de la qualité effective des compétences apportées par le système. Voici pourquoi il nous semble utile de faire un point comparatif sur la variété de l'inégalité de répartition dans le monde de la qualité des apprentissages fondamentaux.

Qualité de l'éducation

La démarche qualité s'attache à décrire si la mise en œuvre d'un produit ou d'un service se réalise suivant une certaine efficacité et si la réalisation répond aux attentes des utilisateurs. Ramené au service éducatif, ceci conduit à s'intéresser aux savoirs acquis et si les moyens utilisés dans cet objectif l'ont été efficacement. Délicates questions, car tout ceci dépend du contexte, les objectifs sont propres au système éducatif : le programme et pour les moyens l'éducation comparée à toujours montré que toute comparaison était délicate, ceci dépendant trop des contextes de terrain.

On retrouve dans ceci les aléas de l'économie de l'éducation qui a toujours privilégié une vision fonctionnelle des activités éducatives. Ceci implique un lien voulu entre le produit de l'éducation et les moyens qui y sont affectés et revient à tenter d'identifier une « fonction de production » de l'école. Ce lien implique que les systèmes éducatifs sont relativement pilotés par l'observation du rapport entre le produit éducatif et les moyens, suivant le concept d'efficacité. Hélas, l'identification de la fonction de production d'éducation reste à ce jour un exercice inachevé. Les études économiques de comparaison internationale ont montré que de nombreuses variables éducatives étaient un facteur déterminant de la croissance du PIB par tête des pays (Barro, 1991 ; Mankiw et al., 1992). Cependant, les problèmes de données ont apporté de nombreuses limites : les variables éducatives, telles que les taux de scolarisation ou le nombre moyen d'années scolaires sont des indicateurs imprécis de la mesure du capital humain acquis par l'éducation (Benhabib et Spiegel, 1994 ; Gurgand, 2000 ; Pritchett, 2001). En cela, ils sont davantage une mesure quantitative de l'affectation de moyens à l'éducation qu'une mesure de résultats par l'évaluation des compétences acquises dans la formation initiale. De plus, la littérature de recherche reste toujours largement dans l'expectative pour lier les moyens consacrés à l'éducation et les résultats qui en ressortent, Hanushek (2006). L'analyse des différences internationales des taux de croissance du PIB montre que les connaissances en mathématiques et sciences sont des composants essentiels du capital humain incorporé à la force de travail (Hanushek et Kimko, 2000). Pour autant, ces compétences ne sont pas toujours précisément corrélées avec les mesures quantitatives de l'éducation ou encore des indicateurs de financement de l'éducation. D'où l'intérêt de rechercher une mesure plus précise de la qualité de l'éducation, même si chaque intervenant est conscient que les savoirs acquis ne le sont pas uniquement dans le système d'enseignement. L'approche du rendement de l'éducation montre que pour le même nombre d'années d'études, les rendements de l'éducation diffèrent significativement entre les pays et les niveaux de revenus (Psacharopoulos et Patrinos, 2004). Un apport a été de montrer que pour expliquer la diversité salariale, à un niveau donné, certaines compétences non cognitives et non acquises dans les programmes scolaires devaient être prise en compte, Boissiere & Knight (1985) ; par ailleurs la maîtrise ou non de ces éléments non cognitifs influence significativement le rythme des acquisitions cognitives, Borghans, Meijers, et al. (2008). Depuis quelques années, de nouvelles études démontrent une relation positive entre les acquis des élèves et les revenus des travailleurs (voir Unesco, 2004, chapitre 2, pour une revue de littérature). Ces recherches utilisent généralement les scores des élèves aux tests sur les

compétences en mathématiques, en sciences ou en lecture. Trois études menées aux États-unis montrent qu'une augmentation d'un écart-type du score des élèves en mathématiques est associée avec une hausse de 12 % des revenus (Mulligan, 1999 ; Murnane et al., 2000 ; Lazear, 2003). En utilisant les données d'une enquête sur la littératie des adultes (enquête *International Adult Literacy Survey* qui concerne 15 pays, dont le Canada, le Chili, les États-unis ainsi que 12 pays européens), Leuven, Oosterbeek et van Ophen (2004) montrent que les capacités cognitives des travailleurs – mesurées par leurs niveaux en lecture – agissent significativement sur leurs salaires. Ces effets persistent même si les auteurs introduisent les années d'éducation, ce qui renforce l'importance que l'on doit accorder à la qualité de l'éducation. En ce qui concerne les pays en développement, Hanushek et Woessmann (2007) soulignent que les rendements de l'éducation seraient supérieurs pour ces groupes de pays en comparaison aux pays développés. Par ailleurs, en utilisant les données du Chili dans l'enquête IALS sur la littératie des adultes, Sakellariou (2006) montre que l'augmentation d'un écart-type du niveau des travailleurs en lecture induit une hausse de 15 à 20 % sur leurs revenus.

Autre raison, il faut également souligner l'effet de la qualité de l'éducation sur la société prise dans sa totalité. Au-delà des seuls gains économiques de l'éducation, il est reconnu que celle-ci permet des gains extra-économiques, notamment en matière de santé (par exemple sur les mères et leurs enfants), dans la réduction de la mortalité infantile, des migrations, de l'âge des mariages, de la violence civile, de la citoyenneté, etc. Les rendements sociaux de l'éducation – qui incluent l'ensemble de ces aspects – dépassent largement les rendements privés. Plus précisément, la qualité de l'éducation apporte de nombreuses améliorations sociales. De meilleurs scores en lecture et mathématiques sont associés avec une baisse des taux de fertilité au Ghana (Oliver, 1999) et en Afrique du Sud (Thomas, 1999).

Les effets de l'éducation sur le développement économique sont la quatrième raison. La relation entre éducation et croissance économique a souvent été évaluée à partir d'indicateurs quantitatifs de l'éducation. La plupart de ces travaux ont souligné que l'éducation était un facteur clé pour le développement économique. Pour autant, Pritchett (2001) a montré que l'effet de l'éducation sur la croissance n'était pas aussi évident que cela. En soulignant l'importance de la qualité de l'éducation dans le processus de développement économique, Pritchett montre qu'un enseignement de piètre qualité pouvait être associé à un faible développement économique. La piètre qualité envoie un signal qui restreint la demande d'éducation des familles, ce qui cantonne la population dans la pauvreté. En soi, la dimension qualitative de l'éducation peut être un accélérateur fondamental de la croissance économique. Les récents travaux associant la qualité de l'éducation et la croissance économique tendent tous à souligner l'importance de la dimension qualitative de l'éducation (Lee et Lee, 1995 ; Hanushek et Kimko, 2000 ; Barro, 2001 ; Coulombe et Tremblay, 2006 ; Hanushek et Woessmann, 2007). Le travail le plus influent montre notamment que l'augmentation d'un écart-type du score des élèves, au primaire, est associée avec un gain de 1 % du taux de croissance annuel du PIB par habitant (Hanushek et Kimko, 2000).

Enfin cinquième et dernière raison toujours en suivant Vegas et Petrow, il est nécessaire de souligner l'impact du niveau des acquis des élèves sur les inégalités. La relation entre l'éducation et les inégalités est complexe. Si d'un côté, l'objectif de scolarisation peut conduire à une baisse des inégalités, il ne faut pas négliger le pouvoir qu'a le système éducatif de légitimer les inégalités sociales (Bourdieu et Passeron, 1979). De récentes études montrent que la qualité de l'éducation peut être responsable de l'aggravation des inégalités de revenus et ainsi que l'augmentation de la qualité de l'éducation pour les populations pauvres peut réduire ces inégalités. L'accès à l'éducation primaire a augmenté dans la quasi-totalité des pays du monde depuis une vingtaine d'années. Parallèlement à cette augmentation de la scolarisation, la question de la qualité de l'éducation reste fondamentale. La généralisation de la scolarisation permet-elle réellement une baisse des inégalités, ou au contraire, les renforce-t-elle ? Comme le souligne Reimers (2000) pour l'Amérique latine, s'assurer que les élèves apprennent des savoirs et compétences est une condition indispensable pour garantir l'égalité des chances.

Les enquêtes qui vont être analysées retiennent cette vision économique de la "production d'école". L'enquête sur les acquisitions scolaires retient que l'acquisition des compétences s'explique par une offre d'éducation dans la classe et l'école, d'une part, mais aussi à décrire le contexte socio-économique du milieu de l'élève, d'autre part. Les enquêtes se situent dans une logique d'explication dite de la "fonction de production d'école" toutefois cette pierre angulaire de la recherche en éducation, si elle est reconnue comme schéma de réflexion, reste largement discutée au niveau de sa validation, Hanushek (2003). La comparaison internationale de ce type d'analyses, en particulier si l'on se base sur le travail comparatif mené, par Hanushek & Luque (2003), sur les enquêtes TIMSS n'a pas validé l'idée dominante d'un plus fort impact dans les PMA des ressources des écoles sur la qualité des apprentissages. Ce constat initial venait des conclusions, dans le début des années 1980, de Heyneman et Loxley (1983) qui après avoir examiné les effets de la situation socio-économique et des facteurs scolaires sur les acquisitions des élèves à l'école primaire dans 16 pays à faibles revenus, et 13 pays à revenus élevés, avaient observé que l'influence des antécédents familiaux, sur les acquisitions, variait considérablement suivant le développement économique entre les pays, et que le pourcentage de la variance expliquée par la dotation des écoles était négativement corrélé avec le niveau de développement d'un pays. Cela a donc été infirmé par Hanushek & Luque (2003), mais aussi par Chudgar & Luschei (2009) ; en suivant une approche d'héritage social Gameron & Long (2007) soulignent, pour les pays en développement, une rapide décroissance des effets-école dans l'acquisition à mesure que l'éducation croissante des parents va permettre d'amplifier les effets d'héritage entre générations et donc relativement diminuer les effets purement scolaires.

Pourquoi alors ne pas s'orienter de manière plus directe sur une mesure comparée des savoirs acquis. Nous avons déjà souligné que l'objectif de l'éducation universelle était intimement lié à l'acquisition effective de compétences essentielles comme la compréhension par la lecture soit la littératie et les outils de calcul essentiels soit la numéracie.

Problèmes du comparatif et information prises en compte

Les principes qui précèdent sont loin d'être validés, surtout dans le domaine du comparatif des moyens utilisés : les systèmes éducatifs étant régulés par un ensemble de règles et de conventions propres à chaque pays. Le meilleur exemple est celui du salaire enseignant (de 75 à 85 % des coûts totaux du primaire), niveau de salaire influencé par des normes et le « poids de l'histoire » propres à chaque société. Ce dernier point est d'autant plus important pour une analyse sur des systèmes nationaux, alors que l'initiative pour l'éducation de base est largement déconcentrée.

Il existe aussi de très fortes difficultés sur le produit éducatif : à quel niveau d'apprentissage celui-ci doit être mesuré :

- Doit-il l'être uniquement par rapport à la comparaison moyens/ résultats « école efficace » ?
- Ou doit-il l'être par rapport à un standard homogène de savoirs acquis dans un contexte d'école « école juste » ?

Au niveau du primaire, on peut néanmoins adopter l'idée d'une homogénéité de l'objectif éducatif liée à la maîtrise de compétences de bases soit la maîtrise de la langue d'enseignement et du calcul. Cette notion d'accomplissement des savoirs de base correspond à la période qui va de la 4e à la 6e année du primaire, là où les contenus des programmes sont les plus homogènes entre les pays.

Si l'objectif d'universalité paritaire est précis : taux d'achèvement du primaire pour les deux sexes, le sous-objectif de qualité est intangible.

Par hypothèse, ici l'éducation est définie de qualité, si les objectifs du programme sont atteints ; tels que la mesure de l'évaluation des acquis des élèves le permet

L'enquête TIMSS

La première mesure des acquis au niveau individuel et permettant une comparabilité internationale a été initiée il y a plus de cinquante ans par l'Association Internationale pour l'Evaluation du Rendement scolaire (IEA).

La première enquête évaluant les mathématiques s'est déroulée entre 1963 et 1967 (*The First International Mathematics Study*) qui a regroupé 12 pays développés (Australie, Belgique, Angleterre, Finlande, France, République Fédérale d'Allemagne, Israël, Japon, Pays-Bas, Écosse, Suède et les États-Unis). Les élèves évalués avaient 13 ans ou étaient en dernière année du secondaire (pour plus d'informations, voir notamment Husen, 1967). La deuxième enquête sur les mathématiques (*The Second International Mathematics Study*) s'est déroulée entre 1977 et 1981 et a concerné les élèves de 13 ans ainsi que ceux étudiant en dernière année du secondaire (Burstein, 1992). Dix-neuf systèmes éducatifs ont été évalués, parmi lesquels figurent 2 pays africains (Nigeria et Swaziland). Cependant, c'est surtout le cycle d'enquête TIMSS (*Third International Mathematics and Science Study*) qui va se révéler le plus prometteur dans l'évaluation des mathématiques et des sciences. L'objectif central de l'enquête TIMSS est d'évaluer le niveau des élèves en mathématiques et en sciences ainsi que de décrire le contexte dans lequel les élèves apprennent. Par ce second objectif, les fondateurs de l'enquête TIMSS ont résolument adopté une approche en termes de finalité politique puisque les résultats des élèves sont associés aux différents facteurs utilisés dans le cadre de l'enseignement dans une logique de fonction de production d'école. Quatre vagues d'enquêtes TIMSS ont été effectuées jusqu'à ce jour : la première concerne 45 systèmes éducatifs en 1994-1995 pour trois populations¹ (grades 3 et 4 ; grades 7 et 8 ; dernière année du secondaire) ; la seconde vague concerne 38 systèmes éducatifs en 1999 pour le grade 8 seulement ; enfin la dernière vague finalisée s'est déroulée dans 50 systèmes éducatifs et pour les grades 4 et 8. En février 2009, l'enquête TIMSS 2007 a été également finalisée. Celle-ci concerne les grades 4 et 8 et plus de 59 systèmes éducatifs. L'évaluation des élèves se base essentiellement sur un référentiel commun de connaissances entre les pays. Plusieurs centaines d'items ont été évalués pour déterminer s'ils sont enseignés dans la plupart des pays participants avant d'être insérés dans les questionnaires. Une volonté de maximiser le nombre d'items standardisés pour tous les pays a été recherchée, ce qui n'exclut toutefois pas la possibilité que dans certains systèmes éducatifs, une partie de ces items ne soient pas réellement au programme. Les questionnaires ne concernent pas seulement le niveau des élèves en mathématiques et en sciences. Mis à part ce questionnaire d'évaluation des acquisitions, d'autres sont complétés pour connaître les contextes sociofamiliaux et de scolarisation.

Trois populations différentes participèrent à l'enquête TIMSS en 1994-1995 :

- Population 1 : Elèves de grades adjacents qui rassemblent la plupart des élèves de 9 ans (en général grades 3 et 4) ;
- Population 2 : Elèves de grades adjacents qui rassemblent la plupart des élèves de 13 ans (en général grades 7 et 8) ;
- Population 3 : Elèves en dernière année du secondaire avec une distinction entre deux sous populations (a) les élèves qui ont pris un test en mathématiques et en lecture (b) les élèves qui étaient spécialisés soit en mathématiques soit en physique qui ont pris un test spécialisé.

Quatre différentes vagues de TIMSS ont eu lieu jusqu'à aujourd'hui. Celles-ci se sont déroulées en 1994-95, 1999, 2003 et 2007. Par ailleurs, la prochaine vague est prévue pour 2011 (les résultats seront publiés en février 2013). L'enquête TIMSS 2007 est l'évaluation internationale qui a réuni le plus grand nombre de pays en développement, en comparaison des vagues précédentes et de l'enquête PISA. Les scores TIMSS ont été standardisés de manière à obtenir une moyenne internationale de 500 et un écart type international de 100. Ainsi, il faut interpréter les scores non comme des valeurs absolues, mais plutôt comme des valeurs relatives en rapport à la moyenne.

Encadré 1. La participation des pays en développement aux enquêtes TIMSS

¹ Parfois, certaines provinces canadiennes ou certains États fédéraux des États-Unis ont participé aux enquêtes de l'IEA. Par souci de simplicité, nous n'incluons pas ces régions dans le calcul du nombre de pays participant aux enquêtes.

En 1994/95, 45 systèmes éducatifs ont participé à l'enquête TIMSS et ont été évalués dans les populations 1, 2 et 3 (c'est-à-dire aux grades 3-4, 7-8 et en dernière année du secondaire). Certains pays n'ont été évalués que pour une partie des populations. Parmi les pays participants, on peut notamment recenser un pays africain (Afrique du Sud) et 9 autres pays en développement ou transition. En 1999, 38 systèmes éducatifs ont été évalués et seule la population 2 a été concernée. Parmi les pays participants, trois étaient africains (Maroc, Afrique du Sud et Tunisie) et 16 autres concernaient des pays en développement ou transition (Bulgarie, Chili, Indonésie, Iran, Jordanie, Lettonie, Lituanie, Macédoine, Malaisie, Moldavie, Philippines, Roumanie, Russie, République Slovaque, Thaïlande, Turquie).

Durant l'année 2003, 50 systèmes éducatifs ont été évalués pour les populations 1 et 2. Parmi ces pays, 6 étaient africains (Botswana, Égypte, Ghana, Maroc, Afrique du Sud et Tunisie). Par ailleurs, 25 autres pays en développement ont été évalués (Arménie, Brésil, Bulgarie, Chili, Estonie, Indonésie, Iran, Jordanie, Lettonie, Liban, Lituanie, Macédoine, Malaisie, Mexique, Moldavie, Palestine, Philippines, Pologne, Roumanie, Russie, Serbie, République Slovaque, Thaïlande, Turquie et Uruguay).

En 2007, la quatrième vague TIMSS a inclus 59 pays ou régions du monde. Cette enquête concerne également les populations 1 et 2. Parmi les pays inclus, cinq sont en Afrique (Botswana, Égypte, Ghana, Maroc et Tunisie). Par ailleurs, certains pays ont participé pour la première fois à une évaluation internationale (Égypte, Kazakhstan, Mongolie, Syrie, Yémen).

La dernière série TIMSS a eu lieu en 2011 avec l'enquête TIMSS 2011 qui a réuni 63 pays et régions du monde². Parmi ceux-ci, sept pays proviennent d'Afrique (Botswana, Ghana, Lybie, Maroc, Afrique du Sud, Tunisie et Yémen). Certains pays participent pour la première fois à cette enquête (Azerbaïdjan, Lybie), tandis que d'autres sont présents depuis la première vague (exemples du Danemark, des États-Unis, de l'Iran, etc.). Les résultats sont attendus pour 2013.

L'enquête PIRLS

Dans le domaine de la lecture, les enquêtes de l'IEA ont débuté en 1968-72 avec l'enquête *The Study of Reading Comprehension* qui regroupe 15 systèmes éducatifs, dont 3 pays en développement (Chili, Inde et Iran). Basée sur les élèves de 10 ans, 14 ans et en dernière année du secondaire, cette première enquête sur la lecture a permis d'apprécier la difficulté de mesurer d'une façon standardisée les compétences des élèves dans un domaine fortement lié à la langue du pays considéré. Par la suite, entre 1985 et 1994, est lancée l'enquête *The Reading Literacy Study* qui va servir d'exemple pour les enquêtes futures. L'objectif principal de cette enquête était de produire des tests internationaux valides et des questionnaires qui pouvaient être généralisés à l'ensemble des pays participants (voir notamment Elley, 1992 ; Postlethwaite et Ross, 1992), tests permettant de juger le degré de compréhension par la lecture. Les données ont été collectées entre 1990 et 1991. Deux types de populations étaient visés : les élèves de neuf ans ainsi que les élèves de 14 ans. Au total 32 systèmes éducatifs ont participé à l'enquête, parmi lesquels figuraient 3 pays africains (Botswana, Nigeria et Zimbabwe) ainsi que 5 autres pays en développement (Indonésie, Philippines, Thaïlande, Trinidad & Tobago et Venezuela). Mais c'est surtout l'enquête PIRLS qui va constituer le cycle majeur de l'évaluation de la lecture au niveau primaire.

L'enquête PIRLS évalue le niveau des élèves en compréhension de l'écrit. Basée sur l'enquête *Reading Literacy Study* (RLS) de 1990-91, PIRLS a été effectuée à deux reprises jusqu'à aujourd'hui (2001 et 2006). Seuls les élèves du grade 4 ont été évalués, soit un âge moyen de 9 ans. Deux objectifs de lecture sont concernés par l'évaluation des élèves :

² Par ailleurs, 14 zones, le plus souvent des États fédéraux, participent à l'enquête TIMSS 2011. Il s'agit par exemple de l'État d'Alabama, de Floride aux États-Unis, ou encore d'Abou Dhabi aux Émirats Arabes Unis.

- L'objectif de littératie (en anglais *literacy*) qui renvoie à la lecture qui implique d'imaginer des événements et caractères et de les mettre en mouvement dans un texte ;
- L'objectif d'information (en anglais *informational*) où la lecture va servir à acquérir et à utiliser de l'information qui est organisée chronologiquement (par exemple dans des biographies) et/ou de façon logique (par exemple dans des textes de réflexion).

Au total, quatre processus de compréhension de la lecture ont été évalués. Ceux-ci concernent la capacité à effectuer un repérage et une explicitation d'informations spécifiques ; effectuer des liens à partir de suites logiques ou chronologiques et à relier des événements entre eux ; interpréter et intégrer des idées et des informations et enfin examiner et évaluer le contenu, le langage ainsi que les éléments textuels.

Comme pour l'enquête TIMSS, afin de ne pas faire passer trop de temps à l'évaluation des élèves, l'examen-type a été découpé en dix livrets disposant chacun de questions d'ancrage. Chaque élève n'a reçu qu'un seul livret – soit un dixième de l'examen complet. En ayant recours à des méthodes psychométriques, les spécialistes ont ensuite recomposé des scores comparables pour chaque élève. Deux types de questions ont été utilisées (questions à choix multiples et questions ouvertes).

Par ailleurs, un questionnaire spécifique a été distribué à chaque élève, chaque enseignant ainsi que chaque directeur d'école. Comme les élèves évalués sont assez jeunes – l'enquête concerne essentiellement des élèves du grade 4, soit un âge proche de 9 ans – le questionnaire relatif aux caractéristiques individuelles et familiales de l'élève a été distribué non pas à l'élève directement, mais à l'un de ses parents.

L'enquête PIRLS 2006 est la seule évaluation internationale qui teste un grand nombre de pays en lecture au niveau primaire. En comparaison, PISA n'évalue que les élèves de 15 ans. Les scores PIRLS, à l'image de TIMSS, sont présentés standardisés de manière à obtenir une moyenne internationale de 500 et un écart type international de 100. Parmi les pays les moins performants, on constate la présence de pays arabes tels que Qatar, Koweït ou encore le Maroc. Cependant, le score le plus faible est obtenu par l'Afrique du Sud avec un score moyen inférieur à 300 points, ce qui représente près de 200 points en dessous de la moyenne internationale. Le score de l'Afrique du Sud est ainsi presque égal à la moitié de celui de la Fédération de Russie ou de Hong-Kong. Pour l'Afrique du Sud, on constate même que plus de 98 % des élèves ont une performance inférieure à 500 points. Au contraire, en Fédération de Russie, cette proportion n'est que d'environ 20 %. Aucun élève d'Afrique du Sud n'a pu obtenir un score supérieur à 600 points, or en Bulgarie près d'un quart des élèves dépasse ce niveau. Les scores très bas obtenus par l'Afrique du Sud expliquent en partie pourquoi l'équipe de l'IEA a lancé l'initiative prePIRLS afin de mieux expliquer l'origine de performances faibles à l'âge de 9 ans une fois que la lecture devrait être acquise.

Encadré 2. La participation des pays en développement à l'enquête PIRLS

Dans l'enquête de 2001, 35 systèmes éducatifs ont été évalués. Parmi les pays participants, un seul est issu du continent africain (Maroc), tandis que l'on recense 13 pays à revenu moyen (Argentine, Bulgarie, Colombie, Iran, Lettonie, Lituanie, Macédoine, Moldavie, Maroc, Roumanie, Russie, République Slovaque et Turquie). 41 pays ou systèmes éducatifs ont pris part à l'enquête PIRLS 2006. Parmi eux, seuls deux sont des pays africains (Maroc et Afrique du Sud). Au total, 15 des pays en développement ou en transition ont participé à PIRLS 2006 (Bulgarie, Géorgie, Indonésie, Iran, Lettonie, Lituanie, Macédoine, Moldavie, Maroc, Pologne, Roumanie, Fédération de Russie, République Slovaque, Afrique du Sud et Trinidad et Tobago).

De façon conjointe à l'enquête TIMSS 2011, la prochaine évaluation PIRLS a eu lieu la même année. Suivant les pays trois niveaux d'intégration des deux enquêtes ont lieu. Le premier utilise les mêmes élèves pour les deux enquêtes. Le second teste des classes des mêmes écoles ou en lecture ou en mathématiques. Enfin dans un troisième cas, les échantillons des deux enquêtes sont indépendants entre les établissements. Regroupant 49 pays, dont trois pays africains (Afrique du Sud, Botswana et Maroc), l'enquête PIRLS reste la seule évaluation internationale de la lecture à l'école primaire. Une innovation majeure de PIRLS 2011 est de s'adapter aux besoins des pays participants. En effet, dans certains pays, les élèves du grade 4 sont encore dans un processus d'apprentissage de la langue, ce qui peut remettre en cause la légitimité d'une telle évaluation. C'est pourquoi il apparaît possible d'évaluer les grades supérieurs (grades 5 et 6).

L'enquête SACMEQ

Le consortium connu sous le nom de SACMEQ (*The Southern and Eastern Africa Consortium for Monitoring Educational Quality*) trouve ses origines dans une enquête nationale de grande envergure menée au Zimbabwe sur la qualité de son éducation primaire en 1991, avec le support de l'Institut International pour la Planification de l'Éducation (IIEP : *International Institute for Educational Planning*) (Ross et Postlethwaite, 1991). Désireux de poursuivre cette réussite inspirée sur le modèle de l'IEA, un certain nombre de ministres de l'Éducation des pays d'Afrique de l'Est et du Sud ont exprimé leur intérêt pour cette étude et leur volonté de participer à une telle évaluation. Les 15 ministères de l'éducation qui sont membres du SACMEQ sont : Botswana, Kenya, Lesotho, Malawi, Maurice, Mozambique, Namibie, Seychelles, Afrique du Sud, Swaziland, Tanzanie, Tanzanie (Zanzibar), Ouganda, Zambie et Zimbabwe.

La première vague de SACMEQ s'est déroulée entre 1995 et 1999. SACMEQ I regroupait ainsi sept différents pays et évaluait le niveau en lecture des élèves du grade 6. Les pays participants sont : Kenya, Malawi, Maurice, Namibie, Tanzanie (Zanzibar), Zambie et Zimbabwe. Bien qu'il s'agissait principalement d'études nationales, celles-ci avaient une dimension internationale, car elles avaient en commun de nombreux éléments (questions de recherche, instruments, populations cibles, procédures d'échantillonnage et analyses). Un rapport séparé a été préparé pour chaque pays. Le second round (SACMEQ II) s'est effectué entre 2000 et 2002 et a regroupé 14 pays ainsi qu'un territoire (Zanzibar). L'évaluation concernait cette fois-ci les mathématiques et la lecture. La population cible restait la même que SACMEQ I, à savoir les élèves du grade 6. Les pays participants sont : Botswana, Kenya, Lesotho, Malawi, Maurice, Mozambique, Namibie, Seychelles, Afrique du Sud, Swaziland, Tanzanie, Tanzanie (Zanzibar), Ouganda, Zambie.

Il est important de noter qu'un certain nombre d'items de SACMEQ II ont été repris de l'enquête TIMSS dans le but de produire des résultats comparables. Les questionnaires ont été utilisés pour obtenir des informations relatives aux entrants de l'éducation, aux conditions scolaires, aux questions d'équité dans l'allocation des ressources humaines et matérielles. Les informations relatives au contexte socio-

économique ont été obtenues par le biais de questionnaires élèves. Plus généralement, l'enquête SACMEQ II inclut des items sélectionnés de 4 enquêtes antérieures : l'enquête sur les Indicateurs de la Qualité de l'Education du Zimbabwe (*Indicators of the Quality of Education Study*), SACMEQ I, TIMSS et l'enquête *Reading Literacy* de l'IEA conduite en 1990/9 (*Reading Literacy Study*).

Le troisième volet de l'enquête SACMEQ concerne les mêmes pays que l'enquête de 2002. SACMEQ III s'est déroulé en septembre 2007. La principale originalité de ce troisième volet consiste à évaluer l'éducation vis-à-vis du virus du VIH-Sida. En effet, une partie des questionnaires élèves, enseignants ainsi que directeurs d'école intègre des questions relatives à ce virus. Cette originalité est signe d'une réelle adaptation de l'enquête SACMEQ aux problèmes importants que connaît cette zone géographique. Notons que si le protocole SACMEQ est assez exigeant sur l'échantillonnage, pour des raisons de simplification de ce dernier, les écoles de petite taille sont exclues, écoles comportant en général 15 élèves et moins dans le grade, ce qui peut dans certains pays introduire des biais.

L'enquête PASEC

Les enquêtes issues du "Programme d'Analyse des Systèmes Educatifs" (PASEC) de la Conférence des ministres de l'Éducation ayant le français en partage (CONFEMEN) concernent les pays francophones d'Afrique subsaharienne. En 1990, la 42^{ème} Conférence ministérielle de la CONFEMEN, tenue à Bamako, a constitué une réponse concrète de l'Afrique francophone au défi de l'Education pour tous lancé à Jomtien la même année. Les ministres ont alors décidé d'entreprendre en commun un Programme d'évaluation pour aider à la réflexion et à leur action : le PASEC a vu le jour lors de la 43^e Conférence ministérielle de la CONFEMEN à Djibouti en 1991.

La CONFEMEN a fixé au PASEC quatre objectifs (CIEP, 2007) :

- Identifier des modèles d'écoles efficaces et peu coûteux en comparant les performances des élèves, les méthodes d'enseignement et les moyens mis en œuvre ;
- Développer une capacité interne et permanente d'évaluation du système éducatif dans chacun des pays participants ;
- Diffuser librement les résultats obtenus, de même que les méthodes et les instruments d'évaluation préconisés ;
- Renforcer le rôle d'observatoire permanent des systèmes éducatifs du Secrétariat technique permanent de la CONFEMEN.

S'il existait dans les méthodes présentées jusqu'ici une certaine unité, ce qui est normal, car la méthode a comme origine celle de l'IEA, la rupture est assez nette avec le PASEC

Encadré 3. Principaux résultats de l'évaluation SACMEQ (1995-2007)

De façon générale, on peut regrouper les performances en trois groupes :

- Les pays qui ont connu une évolution positive dans les deux domaines : ils sont au nombre de 6 (Lesotho, Maurice, Namibie, Swaziland, Tanzanie et Zanzibar). La hausse la plus prononcée concerne la Namibie avec environ 45 points de plus dans les deux domaines entre 2000 et 2007. Cependant, le niveau relatif de ce pays était plutôt bas au début des années 2000.
- Les pays qui ont connu une évolution stable ou contrastée : quatre pays sont inclus dans cette catégorie (Afrique du Sud, Botswana, Kenya, Seychelles). Dans les deux premiers pays, la performance augmente peu tandis que dans les deux derniers, elle baisse légèrement. Cependant, ces évolutions ne sont pas significatives.
- Les pays qui ont connu une baisse de leur performance dans les deux domaines : quatre pays sont concernés (Malawi, Mozambique, Ouganda et Zambie). La baisse la plus prononcée se situe au Mozambique avec près de 40 points en moins en lecture et plus de 46 points de moins en mathématiques.

Source : UNESCO International Institute for Educational Planning (2010) et site internet du Sacmeq : <http://www.sacmeq.org/indicators.htm>

Il est possible de recenser trois types d'évaluations menées par le PASEC :

- **L'évaluation diagnostique** qui consiste à mesurer les acquisitions des élèves au cours d'une année scolaire, puis à identifier les facteurs qui influent positivement ou négativement sur les apprentissages (près d'une quinzaine d'évaluations de ce type ont été réalisées ou sont en cours de réalisation jusqu'à aujourd'hui).
- **L'évaluation thématique** qui se base sur les principes de l'évaluation diagnostique, sauf qu'elle s'intéresse à un thème précis tel que la formation professionnelle des enseignants, ou encore le recrutement d'enseignants contractuels (quatre évaluations thématiques ont été réalisées jusqu'à aujourd'hui). L'échantillonnage n'est plus basé sur l'élève, mais sur les catégories d'enseignants que l'on cherche à analyser.
- **Le suivi de cohorte** qui sert à suivre l'évolution d'un même groupe d'élèves pendant cinq années consécutives, en évaluant chaque année leurs acquisitions scolaires (deux suivis de cohorte ont pu être réalisés de façon complète).

L'enquête PASEC vise à évaluer en début et en fin d'année les élèves du grade 2 et du grade 5. Par exemple, le test de mathématiques au grade 5 inclut des items qui évaluent les connaissances dans les propriétés des nombres et leur habilité à effectuer des calculs simples (addition et soustraction). Les tests incluent également des items qui demandent aux élèves d'utiliser l'addition, la soustraction, la multiplication et la division dans la résolution de problème. D'autres items évaluent les connaissances des élèves dans les décimales, les fractions et les concepts basiques en géométrie.

Cette base de données comprend les performances scolaires au primaire en mathématiques et en français. Au niveau du CP2 (deuxième grade du primaire) comme du CM1 (cinquième grade du primaire), entre 2.000 et 2.500 élèves dans une centaine d'écoles, ainsi que leurs professeurs et les directeurs, ont été interrogés dans chacun des onze pays.

Le protocole de l'enquête présente la particularité de deux évaluations des acquis scolaires, l'une en début d'année, la seconde en fin d'année ; il s'agit à ce jour de la seule enquête internationale menée en termes de "valeur ajoutée" (voir CONFEMEN, 2004).

Certains pays ont participé plusieurs fois à l'enquête PASEC. On recense, par ordre chronologique, les participations suivantes :

- Pasec I : Djibouti (1993/1994), Congo (1993/1994), Mali (1994/1995)³,
- Phase II : République Centrafricaine (1994/1995), Sénégal (1995),
- Pasec III : Burkina Faso (1995/1996), Cameroun (1995/1996), Côte d'Ivoire (1995/1996),
- Pasec IV : Burkina Faso (1996/1998), Côte d'Ivoire (1996/1998), Sénégal (1996/2000), Madagascar (1997/1998), Tchad (2000/2001),
- Pasec V : Togo (2000/2001),
- Pasec VI : Guinée (2003/2004), Mali (2001/2002), Niger (2001/2002),
- Pasec VII : Bénin (2004/2005), Cameroun (2004/2005), Mauritanie (2003/2004), Madagascar (2005/2006), Maurice (2006), Tchad (2003/2004),
- Pasec VIII : Congo (2006/2007), Sénégal (2006/2007) et Burkina Faso (2006/2007),
- Pasec IX : Burundi (2009/2010), Côte d'Ivoire (2008/2009), Comores (2009-2010), Liban (2009-2010),
- Pasec X 1^{ère} phase : Tchad, Togo, République Démocratique du Congo (2009/2010),
- Pasec X 2^{ème} phase : Vietnam (2011/2012), Cambodge (2011/2012), Laos (2011/2012),
- Pasec XI : trois nouveaux pays pour l'année scolaire 2011/2012 et quatre pays pour l'année scolaire 2012/2013.

Les enquêtes des phases PASEC V et VI étaient des enquêtes thématiques. La première enquête de Guinée du PASEC VI, a été suivie d'une réplique en 2005/2006 afin d'évaluer un changement dans le cursus de formation des maîtres contractuels. A noter que les enquêtes du Sénégal (1995/2000) et de Côte d'Ivoire (1995/1998), plus partiellement du Burkina Faso (1995/1998) étaient des suivis de cohorte, tandis que les autres concernent des enquêtes diagnostiques. Par ailleurs, les résultats de quatre premières évaluations sont difficilement accessibles (Pasec I et RCA en PASEC II, car les enquêtes n'ont pas été réalisées dans des conditions permettant la comparaison. En 2005, pour alimenter une étude sectorielle en RCA, le Pôle de Dakar (BRED-UNESCO) a appliqué les outils PASEC en RCA, limitant l'analyse à un test unique de fin d'année, ceci en 5^e année uniquement. Enfin, même si la diffusion des résultats est plus rapide que pour le SACMEQ, les résultats d'enquêtes récentes ne sont pas encore publiés par la CONFEMEN comme la très récente enquête sur la Côte-d'Ivoire ou des enquêtes sur des pays non francophones (Guinée-Bissau, Vietnam, Liban). Les données constituent une base excellente pour une analyse des déterminants de la qualité des acquis scolaires au primaire.

3 Dans cette première vague, réalisée par des équipes indépendantes dans chaque pays, des hétérogénéités de traitement ont rendu difficile la comparaison ; aussi les vagues suivantes ont été coordonnées par une équipe technique centrale.

Encadré 4. Quelques résultats de l'évaluation PASEC

L'étude du PASEC au Sénégal a été effectuée au cours de l'année scolaire 2006-2007. Les élèves de deuxième année et cinquième année ont été soumis à des tests de français et de mathématiques, en début et en fin d'année. L'échantillon est élaboré sur la base de données du Ministère de l'Éducation Nationale et fait référence au statut des écoles (privé, public et arabophone) et aux zones géographiques.

Un nombre de 1 979 élèves de deuxième année et 1 910 de cinquième année ont participé à l'enquête. Le taux de déperdition est d'environ 14 %, ce qui reste assez élevé pour une école primaire. Dans le tableau 5, nous présentons les résultats de l'évaluation PASEC pour le Sénégal, mais également huit autres pays. La performance du Sénégal est faible en deuxième année, en comparaison de celle d'autres pays tels que le Cameroun ou Madagascar. Cependant, les scores augmentent de 19 % entre le début et la fin de l'année, ce qui marque une hausse conséquente, surtout lorsqu'on la compare aux évolutions des autres pays. Il est en effet assez alarmant d'observer une baisse de la performance des élèves dans des pays tels que la Mauritanie. Cependant, les questionnaires de fin d'année sont plus compliqués que ceux de début d'année, ce qui peut expliquer ces baisses quasiment généralisées.. Tout comme l'ensemble des pays cités dans le tableau 5, la performance du Sénégal diminue entre le début et la fin de la cinquième année. Les pays les plus performants parmi les pays du PASEC sont le Cameroun et Madagascar. Il faudrait aussi y inclure Maurice qui obtient des résultats plutôt élevés.

Tableau 4. Résultats pour 9 pays au test du PASEC (2004-2007)

Pays	Deuxième année			Cinquième année		
	Début	Fin	Variation	Début	Fin	Variation
Burkina Faso	33,5	33,5	0	40,1	38,2	-5
Cameroun	53,2	54,7	3	54,6	47,2	-14
Congo	47,1	44,8	-5	44,9	36	-20
Madagascar	59,3	53,4	-10	64,0	52	-19
Sénégal	37,9	45,2	19	45,8	40,9	-11
Tchad	48,0	42,5	-11	39,5	34	-14
Benin	39,9	33,1	-17	44,5	31,9	-28
Gambie	45,6	48,3	6	53,1	42,4	-20
Mauritanie	42,1	31,7	-25	24,9	22,2	-11

L'analyse de quelques résultats du PASEC Sénégal montre tout l'intérêt de la stratification utilisée. Ainsi, il se dégage de fortes différences de performance entre le secteur privé, le secteur public de la capitale et le public hors Dakar. Le score final moyen en deuxième année est d'environ 75 points, contre moins de 50 points pour Dakar et environ 41 points pour le public hors Dakar. Ainsi, le score moyen du privé est environ le double de celui du public hors Dakar. Une étude de l'évolution des scores entre 1996 et 2007 est également possible. On constate globalement que la performance est restée plutôt stable en deuxième année et qu'elle a augmenté d'environ 3 points de pourcentage en cinquième année.

Source : CONFEMEN (2007b)

L'enquête LLECE

Le réseau des systèmes éducatifs nationaux des pays d'Amérique Latine et des Caraïbes, appelé *Latin American Laboratory for Assessment of the Quality of Education* (LLECE) a été créé en 1994 et est coordonné par l'Office Régional de l'UNESCO en Amérique Latine et aux Caraïbes.

L'objectif principal de cette enquête est de fournir des informations sur les niveaux des élèves et les facteurs associés à ces performances susceptibles de guider les politiciens dans leur choix en termes de politique éducative pour les pays d'Amérique Latine. Pour ce faire, le LLECE consiste à répondre aux questions suivantes : Qu'apprennent les élèves ? A quel niveau l'apprentissage se réalise-t-il ? Quelles sont les compétences que les élèves développent ? Quand l'apprentissage se réalise-t-il ? Sous quelles conditions se réalise-t-il ? Quel est l'écart entre les zones rurales et urbaines ? (Casassus et al., 1998).

Dans chaque pays, des échantillons d'environ 4 000 élèves du grade 3 (âgés de 8 et 9 ans) et du grade 4 (âgés de 9 et 10 ans) ont été construits. Ces enquêtes ont concerné plus de 50.000 enfants, soit au moins 100 classes par pays.

Enfin, l'enquête LLECE II (ou encore appelée SERCE pour la seconde vague) évalue le niveau des élèves en mathématiques en découpant ce domaine en 5 parties (Unesco-LLECE, 2008, p.14) :

- Domaine numérique
- Domaine de la géométrie
- Domaine de la mesure
- Domaine des compétences basées sur les informations
- Domaine de la variation

Il est possible de constater des similitudes avec l'enquête TIMSS. Par ailleurs, une partie des items des tests LLECE est inspirée de ceux de TIMSS au grade 4⁴. Les items sont basés sur la concordance entre les programmes scolaires et les acquis des élèves. Il est ainsi question de mesurer le niveau des acquis scolaires, plus que les compétences des élèves qui leurs sont utiles dans la vie active (comme le fait PISA). Un questionnaire élève afin de cerner son environnement a été accompagné au test sur les acquis. Par ailleurs, un questionnaire enseignant et un questionnaire remis au directeur de l'école ont été utiles pour avoir des informations sur l'environnement de l'école. De manière générale ces enquêtes privilégient dans l'analyse l'opposition entre les zones urbaines et les zones rurales. L'enquête LLECE I évalue deux grades adjacents (grade 3 et grade 4), mais à la même période, ce qui avait également été le cas de l'enquête TIMSS 1995. Il est pour autant impossible de suivre les élèves sur une période définie, car ce sont des élèves différents qui sont évalués sur ces deux grades. Les élèves étaient âgés entre 8 et 9 ans selon les pays. Pour autant, la deuxième enquête LLECE, appelée SERCE, a consisté à évaluer deux grades différents : le grade 3 et le grade 6 (UNESCO-OREALC, 2008). La première enquête s'intéressait à dans 13 pays du sous-continent (Casassus et al., 1998, 2002). Ces 13 pays sont : Argentine, Bolivie, Brésil, Chili, Colombie, Costa Rica, Cuba, République Dominicaine, Honduras, Mexico, Paraguay, Pérou et Venezuela. Des données pour seulement 11 pays ont été incluses dans le rapport officiel (Casassus et al., 1998). En 2006, le deuxième volet de l'enquête LLECE a été lancé. Il regroupe les mêmes pays que LLECE I. Cependant, à la différence de LLECE I, le second volet de l'enquête a intégré le domaine des sciences en plus des mathématiques et de la lecture pour un nombre de six pays volontaires. Par ailleurs, les grades testés ont été le grade 3 et le grade 6 (UNESCO-OREALC, 2008).

4 A la différence de TIMSS, l'enquête LLECE 2 – SERCE évalue les élèves du grade 3 et du grade 6. En effet, dans TIMSS, seuls les élèves du grade 4 au niveau primaire sont évalués (les élèves du grade 8 sont au niveau secondaire).

L'enquête MLA

Dans le cadre d'un projet conjoint de l'UNESCO et de l'UNICEF, le programme MLA (*Monitoring Learning Achievement*) mène des études sur les acquis de l'apprentissage sur une vaste échelle géographique : ces études sont effectuées dans plus de quarante pays avec la volonté de transférer la capacité d'analyse au niveau national (Chinapah, 2003). Sur la base de données recueillies une fois les apprentissages fondamentaux réalisés, soit suivant les pays entre la 3^e et 6^e année, les pays doivent être en mesure d'identifier les facteurs qui favorisent ou freinent les apprentissages à l'école primaire, d'analyser les problèmes, de formuler des adaptations des politiques éducatives et de suggérer de nouvelles pratiques afin d'améliorer la qualité de l'enseignement. Plus récemment, le projet MLA II élargit les enquêtes au début du secondaire (grade 8). Par rapport aux enquêtes du PASEC et du SACMEQ, où les élèves sont testés uniquement sur des connaissances scolaires, le MLA porte également sur des questions de connaissances pratiques et de prévention ; de plus, il rajoute la possibilité d'ajouter des items à la demande du pays étudié. Au total, 72 pays ont pris part à l'évaluation du niveau des élèves par le biais de l'enquête MLA. Toutefois, les données n'ont pas toutes été publiées ; souvent manque des informations essentielles comme les caractéristiques essentielles de la dispersion entre observations. En complément aux rapports nationaux, un rapport séparé sur MLA I a été rédigé pour 11 pays d'Afrique (Botswana, Madagascar, Malawi, Mali, Maroc, Maurice, Niger, Sénégal, Tunisie, Ouganda, Zambie ; voir UNESCO, 2000). Il est important de noter que l'enquête MLA a été sujette à de nombreuses critiques, notamment sur la qualité de l'administration des questionnaires, sur les procédures d'échantillonnage effectuées dans certains pays et sur la difficulté d'obtenir des items comparables entre les différents pays. A l'inverse, il faut reconnaître que MLA, avec un certain recul dans l'observation, a été pour certains pays, particulièrement en Asie du Sud, le déclencheur d'action d'évaluation des acquisitions et donc de l'introduction d'une pratique de l'évaluation régulière des acquisitions.

Encadré 4. Principaux résultats de l'enquête SERCE

Au grade 3 et en mathématiques, les performances des pays dans l'enquête SERCE peuvent être classées en trois groupes de pays :

- Les pays qui obtiennent une performance significativement supérieure à la moyenne régionale (Chili, Cuba, Costa Rica, Mexique, Uruguay et l'Etat du Nuevo Leon au Mexique) ;
- Les pays qui ont un niveau proche de la moyenne régionale (Argentine, Brésil et Colombie).
- Les pays qui ont un score significativement inférieur à la moyenne régionale (Guatemala, Équateur, El Salvador, Nicaragua, Panama, Paraguay, Pérou et République Dominicaine).

D'autres résultats intéressants peuvent être soulignés. La Seconde étude régionale comparative et explicative (SERCE, *Segundo Estudio Regional Comparativo y Explicativo*), réalisée en Amérique latine, a observé que les filles obtenaient des résultats sensiblement meilleurs que les garçons en lecture en 3^e et en 6^e années dans la moitié des 16 pays participants (Argentine, le Brésil, Cuba, le Mexique, le Panama, le Paraguay, la République dominicaine et l'Uruguay). Dans l'évaluation SERCE, les scores obtenus en lecture par les élèves de 3^e année étaient très variables, quel qu'ait été le niveau de performance des pays. À Cuba, l'écart entre les 10 % d'élèves les meilleurs et les 10 % d'élèves les moins performants était de 295 points sur l'échelle de compétences (779 moins 484). Dans la plupart des autres pays de la région, les écarts étaient plus faibles, comme en Argentine (236 points), au Costa Rica (231), en El Salvador (219) et au Paraguay (241).

Source : UNESCO-OREALC (2008).

Mesure de la qualité des apprentissages

Les enquêtes internationales disponibles permettent une mesure des compétences de base au primaire si du moins on se limite aux savoirs de base en appréciés par la maîtrise de la littératie et de la numéracie,

Les divers résultats peuvent être retraités et quantifier sur une échelle de 0 à 100 traduisant le pourcentage de validation des connaissances donc ici la mesure de la qualité de l'éducation, plus précisément les taux de réussites d'échantillons représentatifs d'élèves à des enquêtes sur les acquis des élèves.

L'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 dans un intervalle court, ici doublon, et par ce biais procéder à un ancrage des enquêtes les unes avec les autres. Par ailleurs, même si la place manque pour tout détailler certaines enquêtes nationales ont pu s'adjoindre aux enquêtes nationales tant sont similaires les méthodes suivies. L'homogénéisation des résultats suit un ancrage dont le détail est présenté dans Altinok et Murseli (2006).

Tableau 5 Les groupes d'enquêtes pris en compte

Série d'enquêtes	Période couverte	Compétences évalués	Niveau	Champ
Timss/Pirls	1995-2007	Mathématiques/ Lecture	Année 4	Pays OCDE puis extension
LLECE- Unesco	1999 et 2006	Mathématiques/ Lecture (Sciences)	Années 4 ou 5	Amérique latine
SACMEQ	1997-2007	Mathématiques/ Lecture	Année 6	Afrique anglophone
CONFEMEN/ PASEC	1995 à 2008	Mathématiques/ Lecture	Année 5	Afrique sub saharienne
MLA-Unesco	1997 à 2000	Mathématiques/ Lecture (Sciences), vie courante)	Années 4 à 6	Afrique et Asie
Enquêtes nationales	1998-....	Mathématiques/ Lecture (autres champs)	Années 4 à 6	

Abréviations : IEA (International Association of the Evaluation of Educational Achievement), TIMSS (Third International Mathematics and Science Study), PIRLS (Progress in International Reading Literacy Study), UNESCO (United Nations Educational, Scientific and Cultural Organization), MLA Monitoring Learning assessments, CONFEMEN (Conférence Francophone des Ministres de l'Education, PASEC (Programme d'Analyse des systèmes éducatifs de la CONFEMEN), SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality), LLECE Laboratorio, bureau régional UNESCO de Santiago, SERCE ou LLECE II Segundo Estudio Regional Comparativo y Explicativo

Dans l'absolu la correspondance du savoir à la compétence dépend de la théorie du TRI/IRT, ceci imposerait au sens strict que toutes les mesures, pour chaque item, dans un pays soient cohérente sur l'ensemble des pays retenus pour définir une macrocompétence étalonnant correctement la difficulté de l'épreuve.

Une méthode d'ancrage revient, comme l'ont proposé, Hanushek et Kimko (2000) 'Schooling, Labor-Force Quality, and the Growth of Nations', American Economic Review, 90 (5), December, 1184-1208 donne un étalon reconnaissance à une catégorie d'enquête et conduit à recadrer les scores en proportion. Ceci suppose que la fonction de répartition des compétences entre individus serait identique entre toutes les enquêtes, limite donc à la comparabilité internationale.

$$(\bar{x}_{ij}^k)_{adjusted} = \frac{\bar{x}_{ij}^k}{\max(\bar{x}_{ij}^k)} \times 100$$

Une troisième voie revient à sauvegarder cette variété où $(\bar{x}_{ij}^k)_{adjusted}$ représente l'ajustement du résultat de l'enquête i pour le pays j dans la compétence k (littératie ou numéracie). L'indice i inclut les groupes d'enquêtes (SACMEQ, MLA, PASEC, LLCE et SERCE, TIMSS et PIRLS). L'indice j représente un pays de ce groupe d'enquête. Le niveau ajusté s'obtient en rapportant chaque résultat initial $\bar{x}_{ij}^k$ à la valeur maximum de l'enquête $\max(\bar{x}_{ij}^k)$, multipliée par 100. Ici nous ne retenons pas un traitement contraignant les distributions, car le nombre et la nature des pays qui ont participé aux tests très différents et souhaitons conserver cette variété transversale. Malgré cette absence de standardisation par la variance des tests, nous avons néanmoins des résultats comparables par harmonisation sur les moyennes des n pays qui ont participé à au moins deux enquêtes (appelés «doublons»); chaque enquête est réajustée ainsi pour obtenir la moyenne des doublons pour chaque enquête :

$$\bar{x}_{i,doublons_n}^k = \frac{\bar{x}_{i,doublons_1}^k + \bar{x}_{i,doublons_2}^k + \dots + \bar{x}_{i,doublons_n}^k}{n} = \frac{\sum_{i=1}^n \bar{x}_{i,doublons_n}^k}{n}$$

Ceci est donc obtenu du calcul de la moyenne des résultats des «pays doublon», à savoir ceux qui ont participé aux enquêtes TIMSS/PIRLS (T/P) et à un autre groupe d'enquêtes dans un intervalle de temps limité, d'où la moyenne :

$$\bar{x}_{T/P_n}^k = \frac{\bar{x}_{T/P_1}^k + \bar{x}_{T/P_2}^k + \dots + \bar{x}_{T/P_n}^k}{n} = \frac{\sum_{i=1}^n \bar{x}_{T/P_n}^k}{n}$$

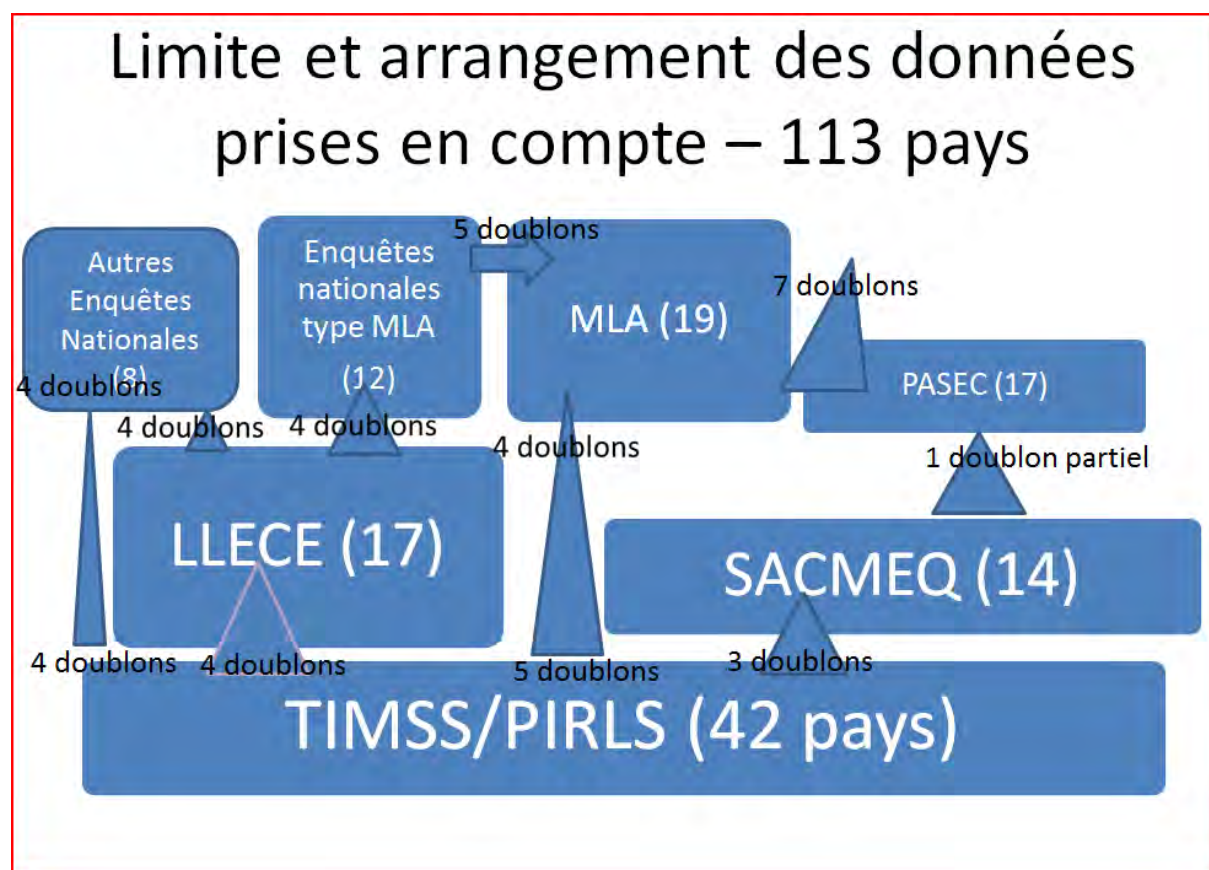
Puis est calculé le poids du transfert de l'enquête i vers l'enquête ajustée en rapport à T/P :

$$poids_i^k = \frac{\bar{x}_{T/P_n}^k}{\bar{x}_{i,doublons_n}^k} = \frac{\bar{x}_{T/P_1}^k + \bar{x}_{T/P_2}^k + \dots + \bar{x}_{T/P_n}^k}{\bar{x}_{i,doublons_1}^k + \bar{x}_{i,doublons_2}^k + \dots + \bar{x}_{i,doublons_n}^k} = \frac{\sum_{i=1}^n \bar{x}_{T/P_n}^k}{\sum_{i=1}^n \bar{x}_{i,doublons_n}^k}$$

puis est réalisé le calcul des valeurs de l'enquête initiale, désormais réajustées en fonction de l'enquête de référence. Cela a été fait en multipliant le résultat initial du pays j dans l'enquête par le poids du transfert :

$$\bar{x}_{ij}^k = (\bar{x}_{ij}^k)_{adjusted} \times poids_i^k$$

C conduit donc à prendre en charge 113 pays pour lesquels l'information est disponible, si 11 » pays sont disponibles pour la numéracie, seulement 111 le sont pour la littéracie. Rappelons en fonction du tableau 5 que cette mesure s'échelonne de 1995 à 2007 pour les cas extrêmes. Dans le cas où plusieurs mesures sont disponibles pour un pays est retenue la plus récente ou sinon celle d'une échelle implicite de confiance (TIMSS/PIRLS,..., MLA et enquêtes nationales).



Une mesure qui laisse visibles les « strates » de développement

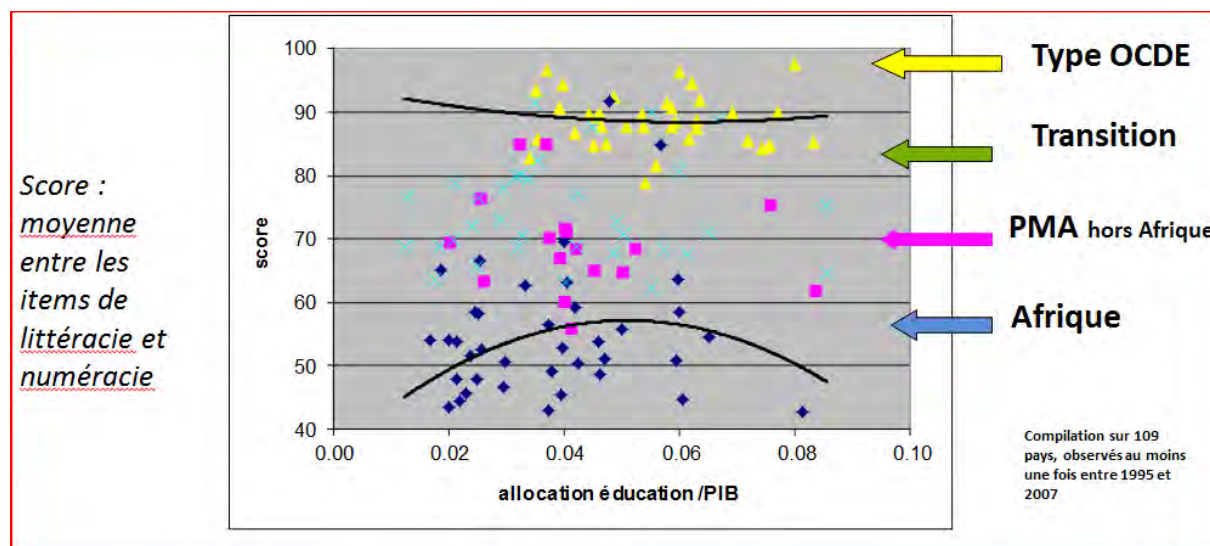
Description et équité

Les scores aux évaluations du primaire et le lien avec l'allocation à l'éducation paraissent largement conditionnés par la position sur l'échelle de développement. Pour mieux représenter ce constat, établissons un lien entre qualités de l'éducation, perçue à travers les scores ainsi harmonisés, et l'effort financier pour l'éducation au primaire. Le graphique 2 montre clairement si en ordonnées l'on représente ainsi la moyenne des 2 scores de littératie et de numération mis à l'échelle et en le croisant, en abscisses, par le coût moyen par élève de la scolarisation primaire en rapport au PIB par tête (notion de richesse) qu'existent à la fois de grandes variétés. Si cela est évident il existe une absence de lien clair entre la dépense et la « qualité » puisque nombre de pays obtiennent un résultat proche pour des coûts relatifs de l'éducation qui varient dans une échelle de 1 à 5. Ce qui paraît d'autant plus cliver les résultats est le classement de la mesure de qualité suivant les grandes régions géographiques indicatrices du niveau de développement. Le résultat comparatif de l'école apparaît donc, indépendamment des moyens consacrés à l'école primaire ; largement influencé par le développement socio-économique du pays concerné.

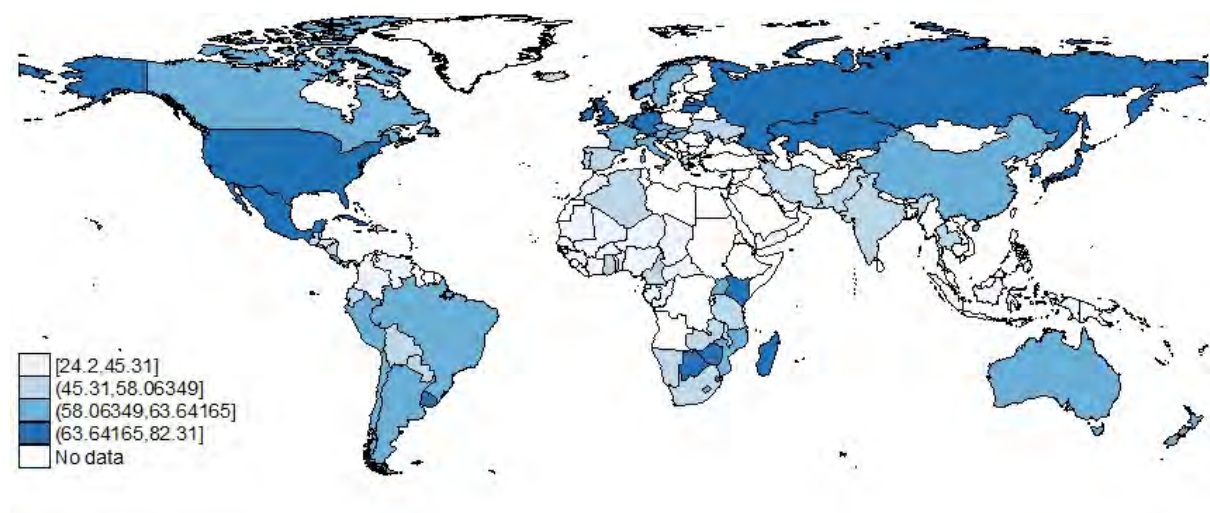
En fonction de cet aspect de dispersion géographique, nous avons voulu représenter ceci e la manière la plus visuelle au travers d'un planisphère. Ainsi, le graphique 3 reprend la visualisation du scores de la numéracie entre les 113 pays. Suivant les spécialistes ce score reste plus important, car il peut préjuger de bonnes bases pour la diffusion des sciences et l'adaptation des pays à l'innovation technologique. L'étendue de la position des pays montre qu'en moyenne et après ancrage les élèves valident de 24% à 82%, suivant les pays, les items auxquels ils ont été soumis. On y remarque certes une différence Nord-Sud indéniable, mais une dispersion interrégionale conséquente qui fait que certains pays dans les zones les

plus défavorisées ont des scores peu différents de ceux enregistrés dans les zones socio économiquement favorisées. À l'inverse des pays des zones les plus favorisées enregistrent des déficits qui les rapprochent des situations moyennes des défavorisées.

Graphique 2 : croisement performance et allocation à l'enseignement primaire.



Graphique 3 Test de numéracie

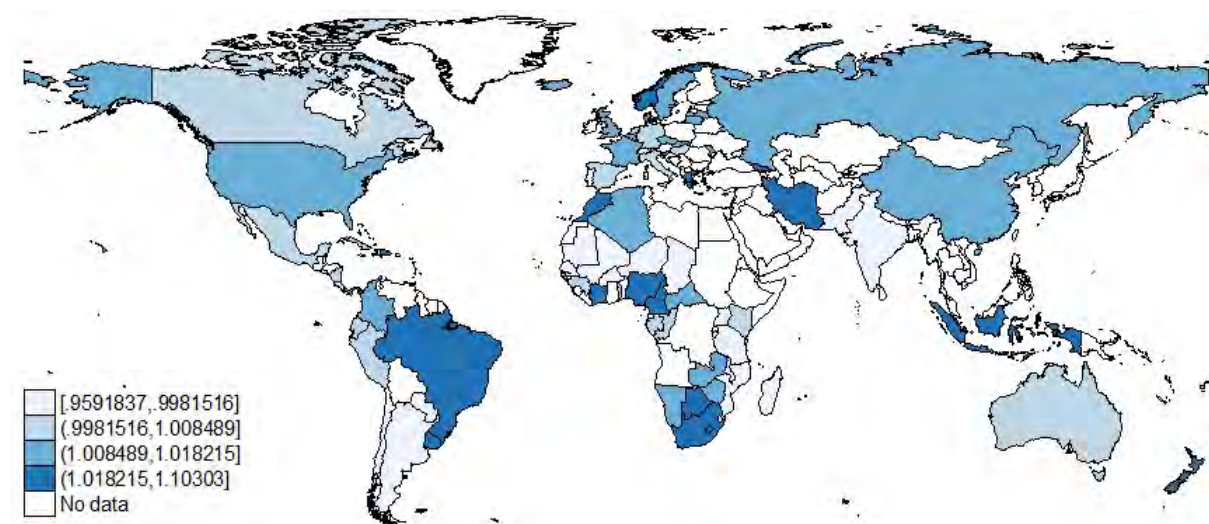


Les ODM ayant clairement affichés l'objectif d'égalité de la qualité de l'éducation suivant les sexes qu'en est-il de cet indicateur ? Pour représenter ceci toujours de manière sommaire, le graphique 4 donne la performance relative en littératie des filles en rapport de celle des garçons. Ici l'étendue de l'indicateur montre une performance relative qui s'étale de 110% à 96% comparée à celle des garçons. Cet indicateur paraît plus pertinent que celui de la performance en numéracie, car les écarts d'accès à la lecture peuvent mieux prendre en compte certaines pesanteurs sociales auxquelles les filles seraient soumises. Le résultat est somme toute encourageant, car il précise qu'à la différence certainement de l'Afrique sahélienne et de l'Asie du Sud, il n'existe pas de biais régionaux importants marquant une sensible discrimination.

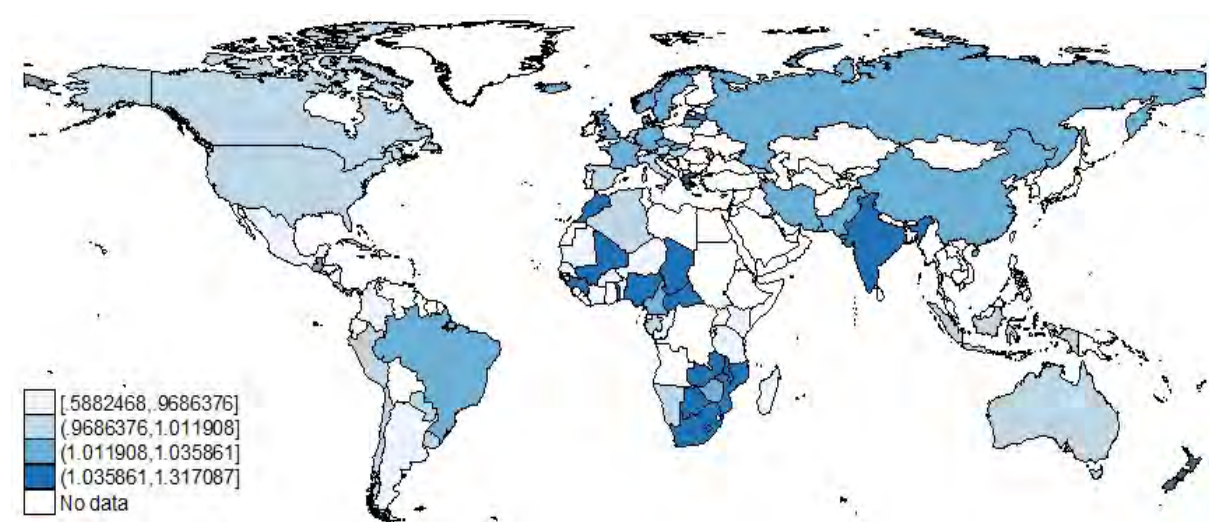
Évidemment ce simple indicateur de spécificité sexuée de la qualité de l'éducation peut être contesté, puisque dans les pays les plus pauvres la principale différence est que les filles accèdent moins à l'école ou

qu'elle ne fréquentent celle-ci que pendant trop peu d'années pour leur permettre l'apprentissage efficace des connaissances de base. Aussi le planisphère du graphique 5 concerne ce m^eme indicateur que le précédent, mais le pondérant par le taux de survie relative scolaire des filles en 5^e année de l'enseignement primaire.⁵ Les différences sont éclairantes, ceci permet d'homogénéiser sensiblement la question pour les pays de l'Amérique latine. Les enjeux pour l'Afrique sahélienne et l'Asie du Sud montrent bien que la question n'est pas celle de la performance des filles, mais pour l'essentiel, celle de leur « survie » dans le système scolaire.

Graphique 4 : La performance relative des filles en littératie



Graphique 5 : La performance relative des filles en littératie corrigée du taux de survie scolaire en 5^e année



Note : ici le rapport entre le score de littératie des filles et celui des garçons est pondéré par le taux de survie « scolaire » des filles en 5^e année

L'école juste

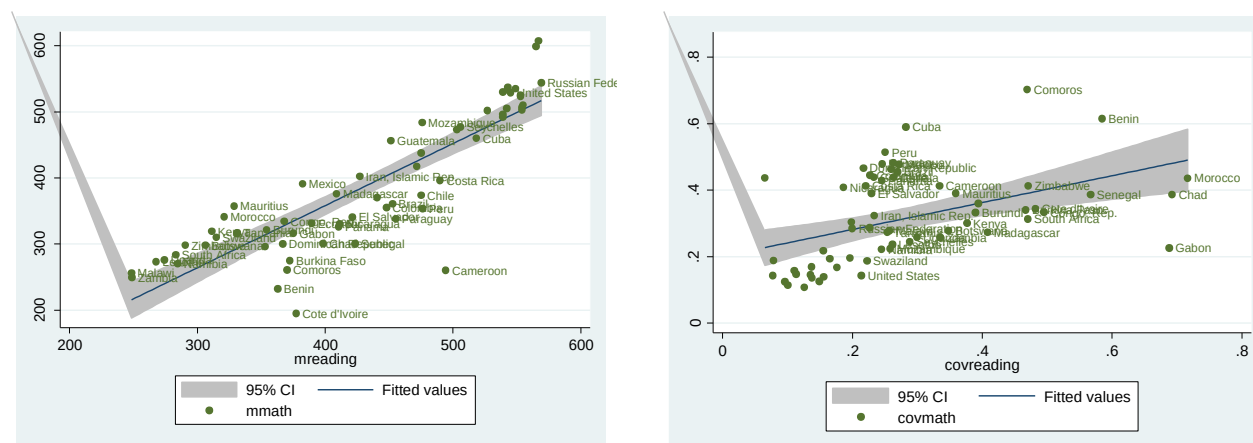
La question de l'égalité garçons/filles étant précisée, il serait bon de mesurer la justice dans l'accès à l'école. Un système en développement est un système qui souvent développe des inégalités, ainsi peut-on s'interroger si l'accès à l'école est identiquement assuré du point de vue de la qualité. On peut en effet

⁵ Données obtenues de la base de données ISU/UNESCO.

penser que dans certains pays une rapide expansion scolaire va de pair avec le creusement d'inégalités dans les apprentissages. Donc, notre question va être d'apprécier si ce niveau de qualité sur lequel nous avons présenté nos résultats, peut-il être mis en cause par le creusement des différences d'apprentissages entre les élèves. Pour arriver à ceci, nous considérons un coefficient de variation σ/μ , c'est-à-dire le rapport, à la fois pour la numéracie et la littératie, entre la dispersion de la qualité d'acquisition entre les élèves d'un pays et le niveau moyen de cette qualité.

Comme nous l'avons mentionné au fil du texte les sources statistiques peuvent être d'une qualité relative et ici cet indicateur est plus sensible qu'une simple moyenne. Aussi avons-nous réduit nos références en ne prenant plus en comptes les ajouts d'enquêtes nationales, la série de résultats du MLA et toutes les enquêtes antérieures à l'an 2000. De plus de cent points de référence, le panel de pays se réduit ici à 78 pour la mesure de littératie et à 74 pour la numéracie ; avec une sous représentation de l'Asie compte tenu des enquêtes MLA dans cette région plus fréquentes. Ces indicateurs de dispersion restent assez largement corrélés pour les deux compétences ($r=0,87$), mais dans l'ensemble n'ont que très peu de liens tous deux avec la dépense vers les enseignements primaires, ce qui ne fait qu'ajouter un doute sur la fonction de production d'école. Un résultat attendu est celui du lien entre ces indicateurs et le niveau de richesse nationale (exprimée par le PIB par tête), mais le lien est moins accentué ici pour l'indicateur de numéracie ($r=0,57$) qu'il ne l'est pour la littératie ($r=0,61$), constat commun à l'ensemble des travaux similaire. Toujours en accord avec les conclusions usuelles de ce type d'analyse, on retrouve le point qu'un système paraît moins inégalitaire à mesure où sa performance de qualité s'accroît. En effet, la corrélation entre le coefficient de variation et le niveau moyen est respectivement de 0,56 pour la numéracie et de 0,60 pour la littératie. Le graphique 6 est double et oppose deux situations, à gauche sont représentés les niveaux croisant la numéracie (axe vertical) et la littératie (axe horizontal). Les labels ne sont pas mentionnés pour les pays les plus riches (sauf pour les États unis – haut droite). La droite de tendance et son intervalle à 95% de confiance (zone grisée) sont aussi repris. Le point le plus significatif vient d'une sous-performance relative en numéracie pour une grande part de l'Amérique latine et plus encore pour les pays francophones de l'Afrique subsaharienne. Le graphique de droite reprend la même organisation du graphique, mais non plus avec les niveaux, mais les coefficients de variation. Ici les résultats sont assez nets avec de très faibles valeurs de deux coefficients pour les pays les plus riches. Si le fait notable paraît ici le relativement important coefficient de variation pour l'Amérique latine en mathématiques, l'importance des coefficients de variation pour des pays francophones de l'Afrique subsaharienne, pourrait conduire à tenter de recherche des causalités en termes d'organisation du système.

Graphique 6 : croisement avec les performances de numéracie (hor.) et de littératie (vert.), niveaux à gauche et coefficients de variation à droite.



Conclusion : Retour sur l'école efficace

Ces résultats ne sont que partiels et sont à compléter par des éléments à approfondir, on peut noter :

- Une recherche à ré exploiter ces scores, pour expliquer les causes de l'efficacité des systèmes éducatifs primaires, pour tenter à lever cette difficulté de lier la « performance » aux paramètres de coût et de financement.
- A l'inverse des éléments comme le taux d'encadrement moyen ou la formation des enseignants paraissent plus explicatifs à ce niveau global que menés sur un groupe d'enquêtes régionales.
- Question à approfondir : dans quelle mesure la marche vers l'EPT s'est réalisée avec une sauvegarde de l'acquisition des compétences de base ?

Références

Bibliographie du texte

- Abadzi, H. (2006). *Efficient learning for the poor*. Washington, DC: The World Bank.
- Altinok et Murseli (2006) "International database on human capital quality," *Economics Letters*, vol. 96(2), pages 237-244,
- Barro R.J., 1991. Economic Growth in a cross section of countries, *Quarterly Journal of Economics*, 151, 407-443.
- Barro, R.J. (2001), "Education and Economic Growth". In Helliwell, J.F. (Ed.), *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being* (pp.14-41), Paris : OECD Press.
- Benhabib J. and Spiegel M., 1994. The role of human capital in economic development : evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34, 143-179.
- Boissiere, M., J. B. Knight, et al. (1985). "Earnings, schooling, ability, and cognitive skills." *The American Economic Review* 75(5): 1016-1030.
- Borghans, L. E. X., H. Meijers, et al. (2008). "The role of noncognitive skills in explaining cognitive test scores." *Economic Inquiry* 46(1): 2-12.
- Burnstein, L. (1992), *The IEA Study of Mathematics III: Student Growth and Classroom Processes*. Oxford: Pergamon Press.
- Casassus, J., Froemel, J.E., Palafox, J.C., Cusato, S. (1998), *First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades. First Report*, Santiago, Chile: Latin American Laboratory of Evaluation of the Quality of Education.
- Center for Global Development. (2006). *When will we ever learn? Improving lives through impact evaluation*. Téléchargé de www.cgdev.org/files/7973_file_WillWeEverLearn.pdf
- Chabbott, C. (2006). "Accelerating early grades reading in high priority EFA Countries: A desk review". Téléchargé de <http://www.equip123.net/docs/E1->
- CONFEMEN (2004). "Les enseignants contractuels et la qualité de l'enseignement de base I au Niger : Quel bilan ?", Document de travail PASEC, Dakar, en ligne (accédé le 15/2/2008) : http://www.confemen.org/IMG/pdf/Rapport_Niger_2004.pdf.

CONFEMEN (2007a). *Quelques pistes de réflexion pour une éducation primaire de qualité pour tous. Évaluation diagnostique du Madagascar*, Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC), Dakar.

CONFEMEN (2007b). *Évaluation PASEC Sénégal*. Programme d'analyse des systèmes éducatifs de la CONFEMEN (PASEC), Dakar.

Elley, W.B. (Ed.). (1992), *How in the World do Student Read?*, Hamburg: Grindeldruck GMBH.

Gameron, A., Long, D.A., (2007), Equality of educational opportunity: a 40-year retrospective. In: Teese, R., Lamb, S., Duru-Bellat, M. (Eds.), *International Studies in Educational Inequality, Theory and Policy*, Vol. 1, Educational Inequality, Persistence and Change. Springer, Dordrecht.

Gurgand, M. (2000). "Capital humain et croissance : la littérature empirique à un tournant ?", *Economie Publique*, vol. 6, p.71-93.

Hanushek, E.A., Kimko, D.D. (2000), "Schooling, Labor-Force Quality, and the Growth of Nations", *American Economic Review*, 90(5), 1184-1208.

Hanushek, E.A., Woessmann, L. (2007), "The Role of Education Quality in Economic Growth", *World Bank Policy Research Working Paper*, 4122, Washington, D.C.

Hanushek, E.A., Woessmann, L. (2010), "The Economics of International Differences in Educational Achievement", *NBER Working Paper*, 15949, NBER, inc.

Hanushek, Eric A. & Luque, J. A., (2003), Efficiency and equity in schools around the world, *Economics of Education Review*, vol. 22(5), pages 481-502

Heyneman, S., Loxley, W (1983). The distribution of primary school quality within high and low income countries. *Comparative Education Review* 27 (1), 108–118.

Husén, T. (Ed.) (1967). *A Comparison of Twelve Countries: International Study of Achievement in Mathematics (Vols. 1-2)*. Stockholm: Almqvist & Wiksell.

Mulligan, C.B. (1999), "Galton versus the Human Capital Approach to Inheritance", *Journal of Political Economy*, 107(6), S184-S224.

Murnane, R.J., Willet, J.B., Duhaldeborde, Y., Tyler, J.H. (2000), "How Important Are the Cognitive Skills of Teenagers in Predictiong Subsequent Earnings?", *Journal of Policy Analysis and Management*, 19(4), 547-568.

OCDE (2005), *Promouvoir la formation des adultes*, Paris.

Oliver, R. (1999), "Fertility and Women's Schooling in Ghana", in *The Economics of School Quality Investments in Developing Countries*, ed. P. Glewwe, 327-344, New York: St. Martin's.

Postlethwaite, T.N. (2004). *Monitoring Educational Achievement*. Paris. UNESCO International Institute for Educational Planning (Fundamentals of Educational Planning, 81).

Postlethwaite, T.N., Ross, K.N. (1992), *Effective Schools in Reading: Implications for Educational Planners: An Explanatory Study. The IEA Study of Reading Literacy II*, Hamburg: IEA.

Pritchett L., 2001. Where has all the education gone?. *World Bank Economic Review*, 15, 367-391.

Psacharopoulos, G., Patrinos, H. (2004), "Returns to Investment in Education: A Further Update", *Education Economics*, 12(2), 111-134.

Ross, K.N., Postlethwaite, T.N. (1991). *Indicators of the Quality of Education: A Study of Zimbabwean Primary Schools*. Harare: Ministry of Education and Culture; Paris: International Institute for Educational Planning.

Sakellariou, C. (2006), "Cognitive Ability and Returns to Schooling in Chile", Background Paper for Vegas, E., Petrow, J. (2008), *Raising Student Learning in Latin America. The Challenge for the 21st Century*, the World Bank, Latin American development forum series.

Thomas, D. (1999), "Fertility, Education and Resources in South Africa", in C.H. Bledsoe, J.B. Casterline, J.A. Johnson-Kuhn, J.G. Haaga (Eds); *Critical Perspectives on Schooling and Fertility in the Developing World*, Washington, D.C: National Academic Press.

UNESCO (2004), *EFA Global Monitoring Report 2005: The Quality Imperative*, Paris.

UNESCO-OREALC (2008). *Student Achievement in Latin America and the Caribbean. Results of the Second Regional Comparative and Explanatory Study (SERCE)*. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, Santiago, Chile.

UNESCO-OREALC (2008). *Student Achievement in Latin America and the Caribbean. Results of the Second Regional Comparative and Explanatory Study (SERCE)*. Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, Santiago, Chile.

Vegas, E., Petrow, J. (2008), *Raising Student Learning in Latin America. The Challenge for the 21st Century*, the World Bank, Latin American development forum series.

World Bank (1995), *Priorities and Strategies for Education*, Washington D.C.: The World Bank.

World Bank: Independent Evaluation Group. (2006). *From schooling access to learning outcomes—An unfinished agenda: An evaluation of World Bank support to primary education*. Washington, DC.

Sites internet d'intérêt

IEA : <http://www.iea.nl/>

LLECE-SERCE : <http://www.llece.org/public/content/view/8/3/lang,en>

PASEC-CONFEMEN : <http://www.confemen.org/spip.php?rubrique3>

SACMEQ : <http://www.sacmeq.org/>

TIMSS-PIRLS : <http://timss.bc.edu/>

Index

Problématique	1
Qualité de l'éducation	2
Problèmes du comparatif et information prises en compte.....	4
L'enquête TIMSS	4
L'enquête PIRLS	6
L'enquête SACMEQ.....	8
L'enquête PASEC.....	9
L'enquête LLECE.....	13
L'enquête MLA.....	14
Mesure de la qualité des apprentissages	15
Une mesure qui laisse visible les « strates » de développement	17
Description et équité	17
L'école juste	19
Conclusion : Retour sur l'école efficace	21
Références	21
Bibliographie du texte.....	21
Sites internet d'intérêt.....	23

LA CLASSE COMME « ZONE DE RESPONSABILITÉ » :

**VERS LA FLEXIBILITÉ SOCIOCOGNITIVE EN TANT QUE COMPÉTENCE
TRANSVERSALE DE L'ENSEIGNANT(E)**

Aneta Mechi

Université de Genève Aneta.Mechi@unige.ch

Mots clés : *attente catégorielle, contexte éducatif hétérogène, flexibilité sociocognitive, situation éducative transversale*

Résumé. *Face à des problèmes éducatifs transversaux (susceptibles de surgir dans chaque cours indépendamment de matière enseignée), nous proposons de développer une compétence transversale qui consiste à voir le comportement humain comme faiblement prédictible et susceptible d'être interprété de plusieurs manières. Les résultats montrent que cette compétence, appelée la flexibilité sociocognitive (FSC) est plus élevée chez les enseignant(e)s qui se sentent davantage concerné(e)s par les problèmes des élèves que ceux qui se sentent moins concernés. En outre, ceux qui se sentent plus concernés ont moins d'attentes catégorielles (se fie moins aux informations déduites de catégorie d'appartenance de l'élève) comme son milieu socioculturel, ses origines, le niveau de classe ou d'école duquel vient l'élève ou laquelle il/elle fréquente.*

1.

*« Essayez de connaître vos élèves, car vous ne les connaissez point »
J.-J. Rousseau*

La coprésence de diverses générations issues de la migration a contribué à la formation d'un public scolaire hautement hétérogène qui appelle des réponses de la part des différents acteurs de l'école. A cela s'ajoutent les phénomènes de la mondialisation et de la mobilité professionnelle qui ont eu un impact sur la diversification de la migration en Europe. Les problèmes liés au contact entre les cultures se juxtaposent à ceux qui existaient depuis longtemps comme l'indiscipline, basse performance ou décrochage scolaire et constituent les défis de l'approche par compétence. Dans ce climat changeant il devient nécessaire de développer de nouveaux outils capables d'équiper les enseignant(e)s face aux sources de tension et aux problématiques pédagogiques dans les contextes éducatifs caractérisés par une diversité sociale et culturelle.

Etant donné une grande variabilité des situations éducatives transversales (SET), il faudrait pouvoir les aborder avec une approche transversale qui permettrait aux enseignant(e)s faire face à des situations éducatives problématiques à travers tous les domaines. Tout l'enseignant peut être confronté dans son cours à des problèmes de discipline ou de performance et en tant qu'individu sujet à des biais de perception et d'interprétation. Ainsi il aurait tendance à les expliquer à l'aide des informations telles que milieu socioculturel (Darley & Gross, 1983) ou statut d'immigrant (Chrysoschoou, Picard, & Pronine (1998). En effet, dans la première étude, les individus étant invités à évaluer une performance ambiguë se basent sur des informations concernant de milieu socioéconomique de l'élève et dans le deuxième exemple les explications de l'échec ou de la réussite de l'élève données par des participants varient selon le prénom (connotation étrangère ou

autochtone) et la profession du père. Il s'agit ce que Jones et McGillis (1976) ont appelé des « attentes catégorielles » : les attentes de l'individu varient selon la catégorie à laquelle appartient la personne qu'il a en face de lui de sorte à pouvoir anticiper ce à quoi il pourrait s'attendre de la part de cette personne.

1.1 Entraves à la réflexion

Les recherches ont mis en évidence deux principales erreurs d'explication du comportement humain qui sont les suivantes : la tendance à expliquer le comportement humain en termes de trait de personnalité (erreur fondamentale), (Ross, 1977) et la tendance à expliquer ce dernier en termes de caractéristiques de groupe d'appartenance (erreur ultime) (Pettigrew, 1979). Ces biais ne permettent pas aux enseignant(e)s de « se sentir concerné(e)s » par des problèmes des élèves et donc prendre en main une situation en essayant de la résoudre ou pour le moins de l'améliorer. Du moment où ils/elles les interprètent comme étant dus au milieu socioculturel ou à la personnalité de l'élève il y a moins de probabilité qu'ils/elles s'engagent dans le processus de réflexion sur des possibles options d'amélioration de situation. Par conséquent, ils/elles peuvent se sentir démunis face à de nombreuses situations éducatives. Par exemple, une situation problématique liée à un problème de discipline ne va pas être abordée de la même manière selon l'explication que l'enseignant(e) donne à l'indiscipline de l'élève en question : l'élève est considéré comme indiscipliné et « agité » à cause de sa personnalité, de son statut d'immigré ou il s'agite et perturbe le cours pour une raison liée au cours, aux relations avec les camarades dans la classe. Dans le premier cas l'enseignant(e) n'a quasiment pas d'emprise sur la situation donc il ou elle ne se sent pas responsable ou concerné(e). En revanche dans le deuxième cas l'enseignant(e) a une grande marge de manœuvre pour que l'élève en question s'agite moins ou perturbe moins le cours. C'est la raison pour laquelle l'attribution de responsabilité pourrait jouer ici un rôle important.

1.2 Sentiment d'être concerné

La théorie de l'apprentissage social (Rotter, 1966) met l'accent sur les attentes de l'individu par rapport à « la probabilité qu'un renforcement particulier se produira à la suite d'un comportement spécifique de sa part dans une des situations spécifiques » (Rotter, 1954, p.107). Dans le cas de contrôle interne l'individu anticipe le lien entre son comportement et le renforcement, alors que dans le cas de contrôle externe il ne le voit pas. Cet aspect d'anticipation est crucial pour notre recherche, car il pointe la distinction entre l'attribution des causes qui s'opère a posteriori (une fois que l'événement a eu lieu) et le « placement du contrôle » qui s'opère avant que la situation advienne. L'individu qui perçoit le lien (contrôle interne) entre son comportement et le renforcement, se voit jouer un rôle dans la causalité de ses comportements, avoir un impact sur la situation face à laquelle il se trouve. Wong et Sproule (1984) l'interprètent en termes d'attribution de responsabilité tout en le distinguant de contrôle perçue qui serait plutôt liée au sentiment de compétence ou d'efficacité. Certains chercheurs ont mis en évidence le lien de l'orientation interne de contrôle avec la conduite de commandement ou avec la recherche d'information.

En effet, Goostadt et Hjelle (1973) ont constaté que face à un ouvrier récalcitrant, les individus avec une orientation interne de contrôle emploient plutôt des techniques de persuasion et d'aide, alors que ceux, avec une orientation externe de contrôle utilisent des pratiques coercitives, moyens punitifs et des menaces de renvoi. L'orientation interne du contrôle a reçu également un grand soutien empirique quant à son lien avec l'acquisition de l'information, c'est qui est, comme nous l'avons mentionné plus haut, un aspect important face à une SET. Plus particulièrement, Seeman et Evans (1962) rapportent que les individus avec l'orientation interne s'informaient plus et se montraient plus insatisfaits des réponses succinctes. De manière générale, ces derniers étaient susceptibles de recueillir plus efficacement une information (même celle de connotation négative) et plus désireux d'acquérir de nouvelles informations. Davis et Phares (1967) dans leur recherche sur la persuasion trouvent que les individus avec l'orientation interne s'informaient davantage,

étaient plus actifs dans la recherche d'informations, plus attentifs et plus sensibles aux informations pertinentes fournies par la situation ou par l'environnement.

Ces recherches apportent un éclairage particulièrement utile pour comprendre quel élément pourrait jouer un rôle dans la recherche d'informations concernant des SET problématiques, le test des hypothèses et le changement dans le temps de ligne interprétative chez les enseignant(e)s. Certes, lors de leur formation les enseignant(e)s pourraient très bien acquérir les connaissances liées aux mécanismes psychosociaux qui pourraient entrer en jeu dans une salle de classe et probablement se sentir plus concernés par la façon dont les élèves vivent leur appartenance à leur classe, leur établissement et leur relation avec l'enseignant(e). Cependant, ces connaissances sont loin d'être suffisantes pour que les enseignant(e)s confronté(e)s aux SET agissent réellement. Il serait également nécessaire d'inclure dans la formation des exercices susceptibles de développer la flexibilité sociocognitive afin que les concepts et les théories vues en cours deviennent de réels outils. Il serait indispensable que les enseignant(e)s, conscient que le comportement humain est tributaire de multiple facteurs contextuels et de ce fait difficile à prédire de façon linéaire, acquièrent également une habitude d'émettre plusieurs hypothèses à l'aide de différents concepts (différents outils) au lieu de s'arrêter sur la première conclusion venant à l'esprit immédiatement et souvent automatiquement. C'est pourquoi, il serait probablement utile de développer ce type de compétence qui serait transversale à tous les cours.

2. Flexibilité sociocognitive

Afin d'être capable d'émettre plusieurs hypothèses et non pas une seule (dimension de plusieurs explications) il semble indispensable d'avoir beaucoup d'informations. La recherche d'informations supplémentaires mise en exergue par des recherches sur le locus of control mentionnées ci-dessus devient un des aspects centraux qui sous-tend la flexibilité. Cette procédure rend possible la déconstruction des associations immédiatement activées par des catégories et diminue la force de première impression.

2.1 Comment s'opère la relativisation ?

La recherche des lectures alternatives d'un même comportement, d'une même attitude ou l'élaboration de plusieurs hypothèses variées sur une même situation fait que le premier jugement « hâtif » perd sa force et ainsi son statut de « seule et unique explication ». Prenons une situation où un élève perturbe le cours en posant des questions « sans lien » ce qui fait rire toute la classe. Si l'hypothèse de milieu sociale ou de personnalité est mise de côté, il est possible d'ouvrir une multitude d'options différentes lesquelles, avec l'intervention de l'enseignant(e) et modification de son feedback, pourraient tout à fait être testées et résolues en classe. Il se peut que le cours ne soit pas donné de manière intéressante, il se peut que l'élève en question veuille impressionner ces pairs ou ne pas paraître « chouchou » de l'enseignant(e). Analogiquement, si un élève a des mauvais résultats et l'hypothèse de milieu socioculturel et de manque de capacités est mise de côté, il devient possible d'envisager sa basse performance sous un angle différent : il se peut qu'il ne veuille pas être considéré par ses pairs comme un « intello », il se peut que la réception de plusieurs feedback perçus comme négatifs l'a fait croire qu'il n'a aucun potentiel et un effort quelconque est inutile.

Néanmoins, ce processus d'élaboration des hypothèses qui pourraient potentiellement jouer un rôle dans la situation problématique est souvent obstrué.

2.2 Blocage de processus exploratoire

De nombreux auteurs mettent en évidence les obstacles à la recherche de lectures alternatives d'une situation observée. Piaget (1956) analyse la tendance de l'individu à se centrer sur un aspect au détriment des autres (centration perceptive) ce qui guide l'interprétation et conduit à la

considérer comme unique (centration représentative). Fiske et Neuberg (1990) étudient la focalisation de l'individu sur une information catégorielle qui empêche une lecture plus individualisante concernant la personne observée. Une autre perspective est présentée par Kruglanski (1983), dans sa théorie de l'épistémologie profane, où il met en évidence la tendance de l'individu à valider son hypothèse sur la base des informations consonantes avec cette dernière (gel de la séquence épistémique), alors que l'auteur souligne l'importance de la prise en compte de celles qui l'infirmeraient pour approfondir le raisonnement (dégeler et poursuite de la séquence épistémique). Ce type de biais aurait aussi des conséquences au niveau comportemental. Dans leur recherche sur la confirmation comportementale, Snyder et Swann (1978) ont démontré par exemple que l'individu ayant reçu une information concernant une personne à interroger aurait tendance de mener l'entretien de façon à confirmer cette information.

2.3 Déjouer les pièges

Certaines approches, comme la théorie de la catégorisation multiple (Crisp et Hewstone, 2007), permettent d'atténuer la force des stéréotypes mais de façon ponctuelle : la capacité de l'individu à voir une personne comme appartenant à plusieurs catégories à la fois permet, certes, de se rendre compte des intersections entre les catégories et ainsi de nuancer la force d'une catégorie perçue comme négative, mais de manière synchronique, sans l'engagement dans le processus d'exploration de lectures différentes de la situation concernant cette même personne. Or, face à des SET problématiques il serait utile de développer une « habitude » de passer de façon diachronique et fluide entre les différents regards ponctuels en explorant différentes interprétations des situations impliquant les élèves. C'est la raison pour laquelle nous proposons une nouvelle notion qui rendrait mieux compte d'un processus consistant à utiliser de manière flexible les catégories dans le temps.

2.4 Approche de la flexibilité sociocognitive (FSC)

La notion de flexibilité est présente en neurosciences comme capacité de cerveau à changer de stratégie en s'adaptant à des situations nouvelles ou complexes, notamment lorsque les routines ou les automatismes s'avèrent insuffisants (Shallice, 1982). Elle est également largement utilisée en psychologie cognitive dans la résolution de problèmes et se définit comme capacité à changer de point de vue ou de mode d'approche quand ils se révèlent inadaptés à la résolution d'un problème (Clément, 2009). Nous proposons ici une définition de la flexibilité spécifiquement en rapport avec des objets sociaux (individus, situations, comportements, opinions) que nous appelons la flexibilité sociocognitive (FSC).

Etant finalisée par le processus de poursuite de la séquence épistémique (au sens de Kruglanski, 1983, cf. ci-dessus), la FSC se réfère à la fluidité du passage entre différentes interprétations possibles face à une situation ou un objet social, supposée empêcher la centration sur une seule lecture et la cristallisation d'une catégorie. Cette définition met l'accent sur l'aspect temporel et la réversibilité d'interprétation en présupposant une construction et déconstruction permanente de ce que nous nommerons les « constellations catégorielles ». Etant à la base une capacité de cerveau, la FSC serait susceptible d'être « exercée » pour devenir opérationnelle et pouvoir être mobilisée dans l'action face à des SET problématiques.

Dans la présente étude nous testons le lien entre une approche flexible chez les enseignant(e)s et leur sentiment de responsabilité pour ce qui se passe en classe et à la façon dont ces derniers utilisent les informations sur l'origine sociale et culturelle de l'élève.

3. Etude

3.1 Population

Cent cinquante-neuf enseignants en formation (femmes=99, hommes=60, moyenne d'âge=31.8).

3.2 Procédure

Les enseignant(e)s en formation ont été invités à remplir un questionnaire anonyme au début de leur formation. Il a été présenté comme un sondage qui permettait améliorer le contenu de la formation pour les années à l'avenir.

Tout d'abord, les participants ont été demandés d'indiquer le degré de leur accord ou désaccord à des affirmations liées à la façon de traiter l'information catégorielle concernant l'élève (son statut socioculturel, le niveau de classe ou d'école, les parents ou l'opinion des collègues enseignant(e)s au sujet de l'élève). Ensuite, les participants ont été demandés d'indiquer le degré de leur accord ou désaccord avec des affirmations liées à l'approche de flexibilité sociocognitive.

3.3 Mesures indépendantes

Sentiment de responsabilité – mesuré par des questions relatives au sentiment d'être concerné par des problèmes des élèves liés tant aux performances qu'au comportement en classe (Exemple : « *En ce qui concerne les problèmes liés à l'indiscipline, ils doivent être réglés uniquement par des spécialistes, tels que le/la psychologue, le/la conseiller(e) scolaire, etc.* », « *Le rôle de l'enseignant(e) ne devrait pas aller au-delà de la transmission des connaissances.*»)

Approche de flexibilité sociocognitive (FSC) – mesurée par des questions relatives à la capacité de mettre en question ou suspendre des jugements hâtifs lié au comportement humain (Exemple : « *Je suis convaincu(e) qu'il y a une seule manière d'interpréter un comportement inapproprié* ») sur une échelle de 1 à 6 (1 = *Pas du tout d'accord* et 6 = *Tout à fait d'accord*). Cette dimension de plusieurs alternatives est inspirée d'une des dimensions présentes dans la mesure de Martin et Rubin (1995) et dans celle de Dennis et Vander Wal, (2010). Une autre dimension mesurée par des questions relatives au degré de prédictibilité du comportement humain (Exemple : « *Les gens peuvent se comporter de manière contradictoire dans des situations différentes* ») est inspirée d'une des dimensions de la mesure des axiomes sociaux de (Singelis, Hubbard, Her & An, 2003).

3.4 Mesures dépendantes

Attentes catégorielles - Le degré selon lequel l'enseignant(e) traite de l'information catégorielle au sujet de l'élèves (niveau de l'école ou de classe, statut socioculturel, opinion des collègues au sujet de l'élève, impression après avoir vu les parents de l'élève) comme un indicateur potentiel de son comportement ou de sa performance scolaire. Les réponses à des questions relatives à ces aspects sont données selon une échelle de 1 à 6 (1 = *Pas du tout d'accord*, 6 = *Tout à fait d'accord*).

3.5 Hypothèses

Hypothèse générale 1 : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait moins d'attentes catégorielles que ceux et celles avec un niveau moins élevé.

H1a : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait moins d'attentes catégorielles concernant le type de classe et d'école que ceux et celles avec un niveau moins élevé.

H1b : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait moins d'attentes catégorielles concernant le statut socioculturel d'élève que ceux et celles avec un niveau moins élevé.

H1c : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait moins d'attentes catégorielles concernant l'opinion de collègues au sujet d'élève que ceux et celles avec un niveau moins élevé.

H1d : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait moins d'attentes catégorielles concernant l'impression que leur donneraient les parents d'élève que ceux et celles avec un niveau moins élevé.

Hypothèse générale 2 : Les enseignant(e)s avec un sentiment de responsabilité élevé aurait un degré de flexibilité sociocognitive plus élevé que ceux et celles avec un sentiment de responsabilité plus bas.

4. Méthode

L'analyse factorielle des questions liées à la flexibilité sociocognitive a indiqué deux facteurs expliquant ensemble 49.9% de variance (KMO=.86). Premier facteur appelé prédictibilité linéaire du comportement humain regroupe les questions telles que « Les gens peuvent se comporter de manière contradictoire dans des situations différentes. » (α de Cronbach=.80). Deuxième facteur (multiples explications de comportement humain) inclus des questions telles que « Je suis convaincu(e) qu'il y a une seule façon d'interpréter un comportement inapproprié. » (α de Cronbach=.78) (voir Tableau 1).

Variables	KMO=.86	α de Cronbach
Faible prédictibilité linéaire		.80
Multiples explications		.78

Tableau 1. Variables liées à l'approche de flexibilité sociocognitive.

Des questions relatives aux attentes catégorielles ont été groupées selon l'indice de α de Cronbach ce qui a permis de calculer plusieurs variables telles que : opinion des collègues (α de Cronbach=.70, voir Tableau 2A), niveau de la classe/école (α de Cronbach=.77, voir Tableau 2B), statut socioculturel (α de Cronbach=.70, voir Tableau 2C) et parents d'élèves (α de Cronbach=.63, voir Tableau 2D).

	Variables liées aux attentes catégorielles	Cronbach- α
A.	Opinion des collègues	.70
	Exemple: « En général, lorsque l'enseignant(e) se renseigne auprès d'un(e)/des collègue(s), il/elle sait à quoi s'attendre concernant la performance d'un(e) élève. », « En général, lorsque l'enseignant(e) se renseigne auprès d'un(e)/des collègue(s), il/elle sait à quoi s'attendre concernant le comportement d'un(e) élève. »	

B.	Niveau de la classe/de l'école	.77
	Exemple: « Quand l'enseignant(e) voit de quelle école vient un(e) élève, en général il/elle sait à quelle performance s'attendre de la part de ce(tte) dernier(e). », « Quand l'enseignant(e) voit de quelle école vient un(e) élève, en général il/elle sait à quel comportement s'attendre. »	
C.	Statut socioculturel	.70
	Exemple: « Globalement, les élèves issus de l'immigration ont des handicaps socioculturels qui rendent compte de leurs problèmes de discipline. », « Globalement, les élèves issus de l'immigration ont des handicaps socioculturels qui rendent compte de leurs faibles performances. Si un(e) élève est indiscipliné(e) c'est qu'en général il/elle vient d'un milieu défavorisé. »	
D.	Parents d'élèves	.63
	Exemple: « Quand l'enseignant(e) voit les parents d'un(e) élève, en général il/elle sait à quoi s'attendre concernant le comportement de ce(tte) dernier(e). », « Quand l'enseignant(e) voit les parents d'un(e) élève, en général il/elle sait à quoi s'attendre concernant la performance de ce(tte) dernier(e). »	

Tableau 2. Variables liées aux attentes catégorielles.

Les variables liées aux attentes catégorielles ont été analysées en tant que mesures dépendantes avec Analyse de Variance (ANOVA) et testées en termes de corrélations.

5. Résultats

5.1 Sentiment de responsabilité et attentes catégorielles

Du point de vue global, les résultats donnent un support empirique aux hypothèses concernant l'effet de sentiment de responsabilité sur les attentes catégorielles. Bien que cet effet ne soit pas significatif pour les attentes liées à l'impression établie sur la base de rencontre avec des parents ($p < .21$), pour celles liées au niveau de la classe ($F(8,150)=6.23$, $p < .00$, $\eta^2=.25$), au statut socioculturel ($F(8,150)=3.55$, $p < .001$, $\eta^2=.16$) ainsi qu'à l'opinion de collègues ($F(8,150)=2.59$, $p < .01$, $\eta^2=.12$), il l'est.

5.2 Sentiment de responsabilité et flexibilité sociocognitive

En ce qui concerne l'effet de sentiment de responsabilité sur le degré de flexibilité en termes de conscience d'existence plusieurs explication du comportement humain, il s'avère positif ($F(8,144)=5.42$, $p < .00$, $\eta^2=.23$). C'est également le cas pour la conscience de faible prédictibilité linéaire du comportement humain. En effet, le niveau de sentiment de responsabilité semble avoir un effet significatif sur cette dernière ($F(8,144)=11.2$, $p < .00$, $\eta^2=.38$)

Les effets mentionnés ci-dessus sont étayés par des corrélations significatives entre différentes variables (voir Tableau 3). Les deux dimensions de la flexibilité sociocognitive corrént positivement avec le sentiment de responsabilité ($r = .41$ pour la prédictibilité et $r = .37$ pour les explications multiples) et négativement avec plusieurs variables liées aux attentes catégorielles : 1) niveau de classe/école ($r = -.27$ pour la prédictibilité et $r = -.32$ pour les explications multiples), 2) milieu socioculturel (*ns.* pour la prédictibilité et $r = -.19$ pour les explications multiples), 3) opinion des collègues ($r = -.22$ pour la prédictibilité et $r = -.20$ pour les explications multiples).

	Niveau de classe/école	Milieu socio- culturel	Opinion de collègues	Parents d'élèves	Sentiment de responsabilité	FSC faible prédictibilité linéaire de comportement humain	FSC Explications multiples de comportement humain
Niveau d'origine (classe/école)							
Milieu socioculturel	.36** (.00) 159						
Opinion de collègues	.44** (.00) 159	.24** (.00) 159					
Parents d'élève	.28** (.00) 152	.18* (.03) 159	.46** (.00) 159				
Sentiment de responsabilité	-.39** (.00) 159	-.24** (.003) 159	-.21** (.01) 159				
FSC faible prédictibilité linéaire de comportement humain	-.27** (.00) 153		-.22** (.01) 153	-.25** (.00) 153	.41** (.00) 153		
FSC Explications multiples de comportement humain	-.32** (.00) 153	-.19* (.02) 153	-.20* (.01) 153		.37** (.00) 153	.53** (.00) 153	

Tableau 3. Corrélations entre FSC, le sentiment de responsabilité et des attentes catégorielles des enseignant(e)s.

Bien que ni le sentiment de responsabilité ni les deux dimensions de la FSC semblent de pas avoir d'effet sur l'impression créée chez les enseignant(e)s par des parents d'élèves, la corrélation entre la conscience de faible prédictibilité du comportement humain et cette impression s'avère significative et négative.

6. Conclusion

L'effet le plus important de cette étude est lié à l'influence de sentiment d'être concerné par des problèmes des élèves sur la façon de traiter et d'interpréter les informations qui parviennent de différentes sources : ici déduites de catégories des élèves. Plus particulièrement, les résultats semblent montrer que plus les enseignant(e)s se sentent concerné(e)s par des problèmes des élèves liés aux performances ou au comportement, moins ils ou elles se basent sur des informations inférées de catégories. Cet effet signifie que cette « zone de responsabilité » de l'enseignant(e) mérite une recherche approfondie pour qu'elle devienne une source de force pour l'enseignant(e) et motivation à résoudre des situations problématiques ayant pour objectif « éveiller » ou maintenir des potentiels des élèves.

Par ailleurs, le fait que les résultats exploratoires de l'approche de la FSC soient significatifs permet d'ouvrir une ligne de recherche liée à son développement. Les études en cours visent la réplique de la structure factorielle et le développement de dimensions supplémentaires de la flexibilité sociocognitive.

7. Références

- Chrysoschoou, X., Picard, M., & Pronine, M. (1998). Explications de l'échec scolaire. Les théories implicites des enseignant(e)s selon l'origine sociale et culturelle de l'élève. *Psychologie et Education*, 32, 43-60.
- Clément, E. (2009). *La résolution de problème. A la découverte de la flexibilité cognitive*. Paris : Armand Colin
- Crisp, R.J. & Hewstone, M. (2007). Multiple social categorization. In M.P. Zanna (Ed.), *Advances in experimental social psychology*, (Vol.39), (pp. 163-254). San Diego, California: Academic Press.
- Darley, J.M. & Gross, P.H. (1983). A Hypothesis-Confirming Bias in Labeling Effects. *Journal of Personality and Social Psychology*, 44, (1), 20-33.
- Dennis, J.P. & Vander Wal, J.S. (2010). The cognitive flexibility Inventory: Instrument Development and Estimates of Reliability and Validity. *Cognitive therapy and research*, 34(3), 241-253.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: influences of information and motivation on attention and interpretation. In M. P., Zanna (Ed), *Advances in experimental social psychology*, (Vol. 23), (1-74). Academic Press.
- Jones, E.E., McGillis, D. (1976). Correspondence Inferences and the Attribution Cube: A Comparative Reappraisal In John H. Harvey, William J. Ickes, and Robert F. Kidd, Ed., *New Directions in Attribution Research*, Vol. 1, Hillsdale, N.J.: Erlbaum, pp. 389-420.
- Kruglanski, A.W. & Freund, T. (1983) The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19(5), 448-468.
- Martin, M.M. & Rubin, R.B. (1995). A new measure of cognitive flexibility. *Psychological Reports*, 76, 623-626.
- Ross, L. (1977). The intuitive psychologist and his shortcomings: Distorsions in the attribution process. In L. Bercovitz (Ed.), *Advances in experimental social psychology* . Vol. 10, 174-221. New York, Academic Press.
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General & Applied*, 80, 1-28.
- Pettigrew, T. F. (1979). The ultimate attribution error: extending Allport's cognitive analysis of prejudice, *Personality and Social Psychology Bulletin*, 5, 461-476.
- Piaget, J. (1956). Centration et décentration perceptives et représentatives. *Rivista di psicologia*, 50(4), 205-223.
- Shallice, T. (1982). Specific impairment in planning. *Philosophical Transactions of the Royal Society of London (Biology)*, 298, 199-209.
- Singelis, Th., Hubbard, C., Her, p. & An S. (2003). Convergent validation of the Social Axioms Survey. *Personality and Individual Differences*, 34, 269-282.
- Snyder, M. & Swann, W. B. (1978). Hypothesis-Testing Processes in Social Interaction. *Journal of Personality and Social Psychology*, 36(11), 1202-1212.

LES COMPETENCES DES ETUDIANTS A L'ENTREE A L'UNIVERSITE : QUELLE INCIDENCE SUR LA REUSSITE?

SOPHIE MORLAIX*, Bruno SUCHAUT**,

* IREDU (Institut de Recherche en Education)/CNRS - sophie.morlaix@u-bourgogne.fr

** IREDU (Institut de Recherche en Education)/CNRS - bruno.suchaut@u-bourgogne.fr

Mots-clés : réussite universitaire, compétences académiques, capacités cognitives

Résumé. Ce travail porte sur les facteurs explicatifs de la réussite en première année universitaire dans le contexte français. Si les variables contextuelles (conditions d'études et modalités d'organisation pédagogique) peuvent expliquer les écarts de qualité de parcours, les caractéristiques personnelles des étudiants (facteurs scolaires et sociaux) jouent un rôle majeur dans le processus de réussite. Peu de recherches dans le contexte national ont tenté de mettre en évidence l'impact des compétences et des capacités initiales des étudiants et d'analyser leurs relations avec les résultats aux examens. Une originalité de notre démarche est donc d'intégrer aux modèles explicatifs de la réussite, des indicateurs spécifiques des compétences des étudiants à leur entrée à l'Université.

1. Des facteurs variés explicatifs de l'échec en première année universitaire

1.1. Généralités

La question de l'échec dans les premières années de l'enseignement universitaire reste cruciale en France, comme dans d'autres pays. L'objectif d'accès de 50% d'une classe d'âge à un diplôme de l'enseignement supérieur est encore loin d'être atteint et un étudiant sur trois quitte la filière dans laquelle il s'est inscrit à l'issue de la première ou de la deuxième année (MESR, DEPP, 2009). Au fil des années, les publics de l'université sont devenus moins homogènes en termes de capital social, culturel et économique, mais aussi sur le plan de l'âge et de la nationalité. Cette hétérogénéité questionne les chercheurs et plusieurs facteurs explicatifs de l'échec universitaire en première année sont mis en lumière.

Ainsi les caractéristiques individuelles des étudiants (passé scolaire, origine sociale notamment) sont souvent mises en lumière comme étant explicatives des différences de réussite ou d'affiliation des étudiants à l'université. En dehors de ces caractéristiques, des facteurs plus contextuels (site universitaire, pratiques pédagogiques, curricula, dispositifs d'accompagnement) se rapportant au contexte universitaire sont également mis en évidence.

Notre recherche se situe dans le prolongement des travaux cités mais en intégrant une dimension complémentaire comme facteur explicatif de la réussite. L'hypothèse générale associée à notre questionnement est que les étudiants ne disposent pas tous des mêmes chances de réussite et que celles-ci sont affectées par l'influence de variables très personnelles comme leurs capacités cognitives et leur niveau académique. Ces performances, liées, par ailleurs, à d'autres caractéristiques des étudiants (sociales et scolaires) seraient alors partiellement explicatives des difficultés rencontrées par certains au tout début de leur parcours universitaire. Le rôle des capacités cognitives dans la réussite scolaire, mis à jour dans le contexte français au niveau de l'école primaire (Barrouillet, Camos, Morlaix, Suchaut, 2008), n'a pas encore fait l'objet de recherches en ce qui concerne l'enseignement supérieur.

Au-delà des capacités cognitives (mémoire de travail, vitesse de traitement de l'information, raisonnement logique), les performances académiques (test de compréhension de l'écrit) des étudiants ont été évaluées. L'identification du poids de chaque groupe de facteurs dans la réussite peut livrer des informations utiles sur les chances de succès (ou les risques d'échec) des étudiants en fonction de leurs caractéristiques personnelles.

1.2. Présentation de la recherche

L'échantillon sur lequel s'appuient nos analyses est constitué d'étudiants inscrits en première année à l'Université de Bourgogne (année universitaire 2010-2011). Si, dans l'absolu, il eut été souhaitable de prendre en compte l'ensemble des filières de l'Université, des contraintes matérielles évidentes dans ce type de travail empirique (pour la passation des épreuves notamment) ont conduit à effectuer des choix dans ce domaine. Nous avons souhaité représenter des filières avec une population d'étudiants relativement hétérogène en fonction du passé scolaire (série du bac en particulier) et avons choisi de nous intéresser aux étudiants de première année inscrit en A.E.S., psychologie et droit. Ces trois filières recensaient 1493 inscrits en septembre 2010, soit 35% de l'ensemble des étudiants inscrits en première année de l'Université de Bourgogne toutes filières confondues. Le tableau en annexe permet de situer la population de référence de notre échantillon par rapport à l'ensemble des inscrits de première année en fonction de plusieurs caractéristiques des étudiants. Toujours pour des raisons de faisabilité de la recherche, tous les étudiants des trois filières n'ont pu être testés dans les différentes dimensions, 616 étudiants ont été volontaires pour passer les tests de raisonnement logique et de compréhension de l'écrit, ces deux épreuves étant administrées sous une forme « papier-crayon ». Les mesures fines des capacités cognitives ont, quant à elles, porté sur un échantillon plus modeste puisqu'elles nécessitaient des passations en petit groupe (un étudiant par ordinateur) et d'une durée plus longue ; au total, on dispose pour 421 étudiants de l'ensemble des indicateurs de performances académiques et cognitives.

Une mesure des connaissances académiques initiales et notamment celle de la compréhension écrite a été élaborée en utilisant le DALF (Diplôme Approfondi en Langue Française). Des mesures sur trois facettes des capacités cognitives des étudiants ont été effectuées : la mémoire de travail, la vitesse de traitement et enfin le raisonnement. La mémoire de travail est mesurée par trois indicateurs : le reading span, l'operation span et l'empan de mémoire de travail issu des épreuves du modèle TBRS (Time-Based Resource Sharing model). La vitesse de traitement est aussi mesurée par trois indicateurs : le simple reaction task (tâche de détection simple), et deux épreuves adaptées des travaux de Posner (toutes deux des tâches de détection complexe). Ces tests cognitifs ont été adaptés de tests américains comme le reading span task (Daneman, Carpenter, 1981) ou l'operation span task (Turner, Engle, 1989), ainsi que les épreuves de vitesse de traitement qui répliquent les tâches de Posner. Les épreuves de mémoire de travail ont été construites selon la structure des tâches informatisées issues du model TBRS. Enfin, les capacités de raisonnement sont mesurées par les matrices de Raven. Des mesures individuelles ont été effectuées dans les trois dimensions cognitives pour chacun des étudiants de l'échantillon. D'autres informations ont ensuite été collectées en février suite aux premiers examens universitaires, de façon à recueillir les résultats détaillés (pour chacune des UE) des étudiants. De façon conjointe, un questionnaire leur a été proposé se rapportant à leurs conditions d'études (activité rémunérée, méthodes et quantité de travail, motivation,...). En fin d'année universitaire, les résultats finaux ont également été collectés.

2. Les déterminants des performances académiques et des capacités cognitives

2.1. Les déterminants des performances académiques

L'ensemble des étudiants de l'échantillon a été soumis au niveau dit "C1" du DALF, et plus particulièrement à l'épreuve de compréhension écrite. L'épreuve est basée sur un texte de Maupassant d'environ deux mille mots, à partir duquel les étudiants doivent répondre à une dizaine de questions, en commentant ou analysant certains passages. Le DALF est composé de 13 items dans sa partie "compréhension de l'écrit". Ces items participent à la construction d'un score dont la valeur la plus élevée est de 22 points. Sur l'ensemble de l'échantillon, le score moyen est de 12,9 avec un écart-type de 3,6 ; la distribution des scores présente une allure gaussienne. Aucune différence de réussite n'est observée entre les filles et les garçons ; il n'y a pas non plus d'écarts en fonction du mode d'hébergement de l'étudiant et entre les étudiants boursiers et les autres. En revanche, les enfants de cadre affichent des résultats supérieurs aux autres professions (2 points de plus que les enfants d'ouvrier). Les variables liées au passé scolaire sont discriminantes puisque des écarts significatifs sont relevés selon la série du bac (résultats plus faibles pour les bacs technologiques et professionnels), la mention du bac (avantage pour les étudiants ayant obtenu au moins une mention assez bien) et l'âge lors de l'obtention du même diplôme (les élèves affichant deux années de retard ayant des scores significativement inférieurs aux autres). Afin d'obtenir une image globale des déterminants de la réussite au DALF, un modèle de régression a été estimé en ne faisant apparaître que les seules variables significatives (tableau 1). Ce sont le type et la mention du bac qui présentent les coefficients les plus élevés, les élèves ayant obtenu un bac professionnel sont les plus défavorisés dans cette épreuve de compréhension écrite (1,6 points de moins que les étudiants ayant obtenu un bac général). On relèvera enfin que les variables prises en compte expliquent moins de 10% de la variance du score au DALF ($R^2 = 0,09$).

Tableau 1 : Modèle expliquant la variance du score au DALF (N = 612)

		Modèle 1		
Modalité de référence	Modalité active	Coefficient	t de Student	Prob.
Constante		+13,0	+63,6	0,00
Age lors de l'obtention du bac				
19 ans ou moins	Plus de 19 ans	-0,9	-1,92	0,05
Série du bac				
Bac général	Bac technologique	-1,0	-2,7	0,00
	Bac professionnel	-1,6	-2,6	0,00
Mention au bac				
Pas de mention ou passable	Mention AB, B ou TB	+1,2	+3,8	0,00
Profession du père				
Autres professions	Cadre	+0,6	1,7	0,08
R^2		0,09		

2.2. Les déterminants des capacités cognitives

Les capacités de raisonnement sont mesurées par les « progressive matrices » de Raven, la version abrégée administrée aux étudiants comporte 23 "problèmes" à résoudre. Pour chaque situation, 8 vignettes représentant des figures sont présentées au sujet qui doit en trouver les règles de succession et choisir parmi 8 autres figures celle qui correspond à la suite de la série. Cette situation de test fait intervenir des aptitudes visuo-spatiales, mais aussi la mémoire de travail et la vitesse de traitement de l'information, l'intelligence fluide et l'induction (Huteau, Lautrey, 1999 ; Jensen, 1997) ainsi que des facteurs spatiaux. Les scores obtenus au Raven sur notre échantillon se distribuent de manière gaussienne et varient de 1 à 20 avec une moyenne de 9,4 et un écart-type de 3,4.

Le tableau 2 expose les estimations d'un modèle expliquant la variance du score au Raven en fonction des caractéristiques des étudiants. Ce modèle explique 13% des différences de résultats au test et certaines variables comme l'âge à l'obtention du bac, la série et la mention du bac ou encore la profession des parents exercent une influence significative sur les scores. La filière fréquentée, même à caractéristiques des étudiants données, fait des différences puisque les psychologues

obtiennent près d'un demi-point de plus que les étudiants d'AES et de droit. Plus surprenant est l'impact du mode d'hébergement au désavantage des étudiants logés en résidence universitaire (la probabilité d'erreur associée au coefficient est toutefois élevée : 8%).

Tableau 2 : Modèle expliquant la variance du score au Raven (N = 612)

		Modèle 2		
Modalité de référence	Modalité active	Coefficient	t de Student	Prob.
Constante		+8,1	+31,2	0,00
Age lors de l'obtention du bac				
19 ans ou moins	Plus de 19 ans	-1,3	-3,3	0,00
Série du bac				
Autres bacs	Bac S	+1,7	+4,9	0,00
Mention au bac				
Pas de mention ou passable	Mention AB, B ou TB	+0,5	+1,1	0,09
Profession du père				
Autres professions	Ouvrier	+0,8	+2,4	0,02
	Agriculteur, artisan, commerçant	+0,9	+2,3	0,02
	Cadre	+1,0	+2,8	0,00
Mode d'hébergement				
Autres modes	En résidence universitaire	-0,6	-1,7	0,08
Filière fréquentée				
AES et droit	Psychologie	+1,4	+5,0	0,00
R ²		0,13		

Trois mesures de mémoire de travail ont été reconnues comme de meilleurs prédicteurs que celles existant précédemment (Barrouillet, Camos, Morlaix, Suchaut, 2008 ; Lépine, Barrouillet, Camos, 2005). Nous utiliserons ici des épreuves impliquant, soit du matériel verbal (Barrouillet, Bernardin, Portrat, Vergauwe, Camos, 2007), soit du matériel visuo-spatial (Vergauwe, Barrouillet, Camos, soumis). La mémoire de travail est mesurée par trois indicateurs : le « reading SPAN », « l'opération SPAN » et le « TBRS » (Time based resource sharing model). Dans l'épreuve du Reading span, les sujets doivent mémoriser des chiffres. Entre chaque chiffre, ils lisent des phrases et doivent juger de la plausibilité de ces phrases. Le score obtenu est le pourcentage de chiffres correctement rappelés. Sur l'échantillon, la moyenne du score est de 81 avec un écart-type de 16, il est expliqué à la hauteur de 8% ($R^2 = 0,08$) par les variables prises en compte (tableau 3).

Tableau 3 : Modèle expliquant la variance des performances en reading span (N = 418)

		Modèle 3		
Modalité de référence	Modalité active	Coefficient	t de Student	Prob.
Constante		+84,7	+40,5	0,00
Garçon	Fille	-6,2	-3,3	0,00
Non boursier	Boursier	-4,9	-3,2	0,00
Age lors de l'obtention du bac				
19 ans ou moins	Plus de 19 ans	-6,3	-3,8	0,00
Filière fréquentée				
Psychologie et droit	AES	-6,8	-2,4	0,00
R ²		0,08		

Les filles obtiennent des performances inférieures de 6 points à celles des garçons. Le retard scolaire et le fait d'être boursier sont aussi des caractéristiques qui influent négativement sur les scores des étudiants.

Pour les épreuves relatives à l'opération span (modèle non présenté), les sujets doivent mémoriser des lettres. Entre chaque lettre des équations sont présentées et les sujets doivent se prononcer sur la justesse du résultat (vrai ou faux). Les modèles de régression indiquent que certaines tendances relevées dans le modèle précédent (reading span) se confirment : moindre performances des filles et des étudiants âgés. D'autres caractéristiques, comme la série du bac (avantage aux scientifiques) ou le mode de résidence (avantage aux étudiants non hébergés en résidence universitaire) exercent également une influence sur les performances dans ce domaine. Enfin dans le TBRS (modèle non présenté), les sujets doivent mémoriser des lettres et juger de la parité de chiffres. Le pourcentage de variance expliquée par les caractéristiques des étudiants est de l'ordre de 10%, par rapport aux épreuves précédentes, la profession du père contribue elle aussi à l'explication de la variance : avantage aux professions intermédiaires. Il faut rappeler qu'à ce niveau d'étude, une sélection sociale a déjà été établie et qu'il est difficile de comparer ces résultats avec ceux habituellement relevés pour les niveaux scolaires inférieurs.

La vitesse de traitement est mesurée par trois indicateurs : le SRT (tâche de détection simple), le Posner 1 et le Posner 2 (tous deux des tâches de détection complexe). Dans la tâche SRT, les sujets doivent appuyer le plus rapidement possible sur la barre espace à l'apparition d'un carré à l'écran. Les carrés apparaissent soit 1000, 2000 ou 3000 ms après le signal. Dans le Posner 1, deux lettres apparaissent côte à côte à l'écran et les sujets doivent repérer si elles sont physiquement identiques. Pour le Posner 2, il s'agit d'une tâche de détection complexe, deux lettres apparaissent côte à côte à l'écran et les sujets doivent repérer si elles sont identiques qu'il s'agisse de majuscules ou de minuscules. Les données recueillies dans les trois épreuves correspondent à des temps moyens millisecondes.

Les trois modèles montrent des similitudes dans les pourcentages de variance expliquée qui restent faibles (de 3 à 4%), peu de variables semblent donc être discriminantes en ce qui concerne les vitesses de traitement. En ce qui concerne les tâches de traitement simples, on retrouve une tendance négative pour les filles ainsi que pour les étudiants âgés, en revanche les possesseurs d'un bac scientifique apparaissent légèrement plus performants que les autres. Pour le Posner 1, les enfants d'ouvrier réalisent des scores un peu plus élevés que les autres mais c'est encore une fois le genre (à l'avantage des garçons) et la série du bac (au détriment des possesseurs d'un bac professionnel) qui sont les caractéristiques les plus discriminantes. Pour les épreuves du Posner 2, on notera que les étudiants de psychologie présentent des temps de réaction inférieurs aux autres, c'est aussi le cas des enfants d'ouvrier. Les étudiants boursiers et ceux titulaires d'un bac professionnel affichent des performances plus faibles que celles des non boursiers et des étudiants titulaires d'un bac général ou technologique.

Afin de disposer de mesures plus générales, il peut-être pertinent de construire des indicateurs globaux rendant compte des deux dimensions ciblées (mémoire de travail et vitesse de traitement) ; pour s'assurer que ce regroupement soit valide, nous avons calculé des corrélations entre les différentes mesures. Les corrélations sont positives et significatives (au seuil de 1%), leur intensité varie toutefois selon les dimensions considérées, la plus forte relation est celle entre les épreuves posner 1 et posner 2 (+0,63). Nous avons donc construit un score moyen en mémoire de travail ainsi qu'un score moyen de vitesse de traitement ; chacun de ces scores prenant en compte chacune des performances¹. Nous disposons ainsi de trois mesures des capacités cognitives : Raven, mémoire de travail et vitesse de traitement. Le tableau suivant indique les corrélations obtenues entre ces trois mesures.

Tableau 4 : Corrélations aux tâches de vitesse de traitement (N = 418)

	RAVEN	Mémoire de travail
Mémoire de travail	+0,23 ***	
Vitesse de traitement	-0,14 ***	-0,24 ***

*** : significatif au seuil de 1%

On constate globalement de faibles corrélationsⁱⁱ, la plus faible étant celle associant le RAVEN et la vitesse de traitement (-0,14), cela suggère que les trois facettes des capacités cognitives appréhendées dans cette recherche ont une composante autonome à prendre en compte dans l'analyse des déterminants de la réussite.

3. L'analyse des facteurs expliquant la variété des notes du premier semestre, et de fin d'année

3.1. Analyse des notes de fin de premier semestre

Les étudiants de droit et d'AES ont composé dans trois UE alors que les étudiants inscrits en psychologie avaient quant à eux quatre UE à valider à la fin du premier semestre. Dans chacune des filières les corrélations entre les notes des UE sont élevées (le plus souvent supérieure à +0,70) et permettent de considérer la note moyenne du premier semestre comme un indicateur global des performances des étudiants. Les moyennes des notes globales du semestre 1 sont proches de 9 sur 20 dans les trois filières. Ces notes des UE des trois filières ne sont pas comparables dans la mesure où les enseignements qui les composent ne sont pas les mêmes, on ne peut donc pas comparer les degrés de réussite des étudiants selon la filière. Cela a pour conséquence que l'utilisation de cette note moyenne sur l'ensemble de l'échantillon devra tenir compte de ce biais ; c'est pourquoi nous introduirons systématiquement la filière fréquentée comme variable de contrôle dans les modèles de régression. Ceci permettra de mesurer l'influence des caractéristiques des étudiants en raisonnant à « filière fréquentée identique ».

Les premiers facteurs examinés sont les capacités cognitives, plusieurs modèles successifs ont été estimés en faisant intervenir les différentes mesures de cette dimension (modèles non présentés). Les performances en mémoire de travail et en vitesse de traitement n'exercent aucun impact sur les résultats universitaires ; en revanche, on relève un effet positif du score obtenu au Raven (le coefficient est positif et significatif). A titre d'illustration, une augmentation de 25% du taux de réussite au Raven entraîne un gain de 0,5 points sur la note moyenne, ce qui n'est pas négligeable. Des modèles globaux (tableau 5) ont ensuite été estimés en intégrant l'ensemble des caractéristiques des étudiants, celles-ci expliquent 40% de la variance des notes de fin de premier semestre (modèle 4). Si aucune différence n'est relevée entre garçons et filles, l'âge exerce un impact significatif (réussite moindre pour les étudiants ayant accumulé du retard)). De même, le fait d'être boursier, qui d'une certaine manière rend compte d'une caractéristique sociale de l'étudiant, est associé à une moindre réussite (près d'un demi-point de moins, en moyenne).

La série du bac est le facteur de loin le plus discriminant. Cette variable explique à elle seule 24% de la variance des notes. La lecture des modèles montre une nette hiérarchie des séries avec les performances les plus faibles relevées pour les bacs professionnels. La série S est associée aux résultats les meilleurs, suivie de la série ES, puis de la série littéraire. Les écarts sont très élevés puisque 4 points séparent les bacs professionnels des bacs scientifiques.

Tableau 5 : Modèles généraux analysant la variance des notes au premier semestre (N =612)

		Modèle 4		Modèle 5		Modèle 6	
Modalité de référence	Modalité active	Coef.	Prob.	Coef.	Prob.	Coef.	Prob.
Constante		+3,8	0,00	+2,1	0,00	+2,2	0,00
Filière fréquentée							
Droit	AES	+2,4	0,00	+2,4	0,00	+2,4	0,00
	Psychologie	+2,5	0,00	+2,5	0,00	+2,5	0,00
Garçon	Fille	+0,2	0,33	+0,2	0,39	+0,2	0,41
Retard scolaire	Age normal	+0,6	0,00	+0,7	0,00	+0,7	0,00
Non boursier	boursier	-0,4	0,02	-0,4	0,02	-0,4	0,03
Série du bac							
Bac professionnel	Bac technologique	+0,9	0,03	+0,7	0,09	+0,7	0,09
	Bac L	+2,7	0,00	+2,3	0,00	+2,3	0,00
	Bac ES	+3,2	0,00	+2,9	0,00	+2,9	0,00
	Bac S	+4,4	0,00	+3,9	0,09	+4,0	0,00
Pas de mention	Mention AB B TB	+2,3	0,00	+2,1	0,00	+2,1	0,00
Score en compréhension de l'écrit (DALF)				+0,15	0,00	+0,15	0,00
Score au Raven (% de réussite)						-0,00	0,68
R ²		0,40		0,43		0,43	

A série du bac équivalente, la possession d'une mention, quelle qu'elle soit, crée aussi des différences : les étudiants qui en sont titulaires obtiennent 2 points de plus que les autres. Au total, le passé scolaire reste un élément déterminant pour le début du parcours universitaire. Ainsi, entre un étudiant venant d'un bac professionnel avec du retard scolaire et un étudiant possédant un bac scientifique avec mention, on relève un écart de plus de 6 points et demi sur la moyenne de fin de premier semestre. On met à nouveau en évidence à travers ces résultats l'importance du parcours scolaire antérieur pour la réussite à l'Université (Prouteau, 2009).

Le modèle 5 apporte une information importante, puisqu'à la marge de l'influence de ces facteurs classiques, le niveau académique exerce un effet autonome sur les résultats des étudiants. La valeur du coefficient (+0,15) indique qu'entre deux étudiants ayant un écart-type de différence au score du DALF (l'écart-type du score au DALF vaut 3,6 points) la différence à la moyenne du premier semestre s'évalue à plus d'un demi-point (3,6 x 0,15). Les performances en raisonnement logique n'ont pas d'effet spécifique sur les résultats scolaires. Le coefficient associé à la variable « Raven » a une valeur nulle (modèle 6). Les capacités cognitives des étudiants n'exercent donc pas d'effet autonome sur les performances universitaires ; en référence aux analyses exposées précédemment, on en déduit que ces capacités se sont exprimées essentiellement au cours de la scolarité passée.

Une manière complémentaire de mesurer la réussite en fin de premier semestre est de raisonner en termes de probabilité de réussite ou d'échec. Sur l'échantillon, la proportion d'étudiants ayant validé le semestre (ceux qui obtiennent au moins 10 sur 20) est de 32%. Un modèle de régression logistique a été estimé (tableau 6) qui intègre les mêmes facteurs explicatifs ; ce modèle présente des résultats proches des précédents.

Tableau 6 : Modèle logistique de réussite au premier semestre (N =612)

		Modèle 7		
Modalité de référence	Modalité active	Coef.	Prob.	Exp
Constante		-5,2	0,00	4,8
Filière fréquentée				
Droit	AES	+2,0	0,00	7,2
	Psychologie	+1,3	0,00	3,8
Garçon	Fille	+0,0	0,81	1,1
Retard scolaire	Age normal	+0,5	0,06	1,6
Non boursier	boursier	-0,2	0,34	0,8

Série du bac				
Bac professionnel	Bac technologique	-0,4	0,44	0,6
	Bac L	+1,0	0,05	2,8
	Bac ES	+1,4	0,00	4,3
	Bac S	+2,6	0,00	10,5
Pas de mention	Mention AB B TB	+1,6	0,00	5,1
Score en compréhension de l'écrit (DALF)		+0,10	0,00	1,1
Score au Raven (% de réussite)		+0,00	0,28	1,0
R ² de Nagelkerke		0,37		

Les chiffres mentionnés dans la colonne « Exp » du modèle 7, expriment les rapports de côte (odds ratio) associés aux coefficients, ils indiquent qu'un étudiant titulaire d'un bac S a 10,5 fois plus de chance d'être admis au semestre 1 qu'un étudiant titulaire d'un bac professionnel ; de même, un étudiant ayant obtenu une mention au bac a 5,1 fois plus de chances d'être admis qu'un non titulaire d'une mention. Le score au DALF exerce son influence puisque les étudiants ayant des performances élevées dans ce domaine ont plus de chances de réussir que les autres. En revanche, les capacités cognitives mesurées par le Raven n'apparaissent pas être un facteur autonome discriminant.

Au-delà des variables mobilisées jusqu'à présent, d'autres informations collectées à partir d'un questionnaire ont permis d'intégrer aux analyses des dimensions plus subjectives dans les conditions d'accès à l'université. C'est notamment le cas des choix d'orientation après le bac. Sur l'échantillon, 58% déclarent avoir choisi leur filière en fonction d'un projet professionnel, 48% par intérêt ou par goût, 13% par curiosité, 5% par défaut et seulement 3% sur la base de la réussite probable dans les disciplines choisies. On notera que les choix multiples sont relativement limités puisque 22% des étudiants mentionnent deux choix, 2% trois choix et seulement 0,5% quatre choix. Parmi les différentes modalités, une seule exerce une influence significative sur la moyenne du premier semestre, il s'agit de l'orientation décidée en fonction d'un choix professionnel. Quand on intègre cette variable au modèle 6 précédent, on relève un effet positif et significatif (au seuil de 1%) de cette modalité de choix d'orientation. Les étudiants qui déclarent avoir décidé de leur avenir scolaire sur la base d'un projet professionnel ont, toutes choses égales par ailleurs, une note moyenne supérieure de près d'un demi-point aux autres (coefficient de +0,49). L'adéquation de l'orientation à un projet professionnel contribue donc pleinement à la réussite.

3.2.L'analyse de la réussite en fin d'année universitaire

Ces estimations porteront sur des effectifs plus réduits (69 étudiants ont abandonné après les partiels de janvier), les faibles résultats aux examens de janvier semblent être le facteur le plus déterminant de l'abandon puisque les étudiants concernés avaient une moyenne de 5,4 sur 20 à l'issue du premier semestre.

Un premier constat est la très forte corrélation existant entre les deux prises d'informations (premier et second semestre), bien que les disciplines évaluées ne soient pas les mêmes (les UE des deux semestres sont différentes), le classement des étudiants n'évolue guère puisque la corrélation relevée sur l'ensemble de l'échantillon est de +0,83 et ne varie guère dans les trois filières : +0,85 en droit, +0,85 en psycho et +0,84 en AES. C'est donc près de 70% de la variance des notes du second semestre qui est expliquée par les performances du premier semestre. Les déterminants de la réussite au second semestre sont comparables à ceux relatifs au semestre précédent, la série du bac, la mention, l'âge apparaissent toujours comme des facteurs déterminants de la réussite. On relèvera également l'influence positive des performances en compréhension de l'écrit et l'absence d'effet du score en raisonnement logique. Par ailleurs, le choix d'orientation n'exerce pas non plus d'effet sur les résultats au second semestre.

Bien que les échelles de mesure soient différentes, nous avons estimé un modèle de progression (modèle 9) entre les résultats des deux semestres. En effet, ce modèle fournit une indication sur les progrès relatifs des étudiants en cours d'année en termes de position au sein de l'échantillon. Le modèle explique les trois-quarts de la variance des notes du second semestre. Certaines variables affichent des coefficients significatifs : mention et série du bac. Il reste à présenter les résultats de l'ensemble de l'année universitaire : notes moyenne de l'année (modèle 11). La moyenne générale sur l'année est de 9,4 avec un écart-type de 3. Les estimations du modèle 11 confirment les analyses précédentes mais on relèvera cependant que les performances des étudiants des bacs technologiques et professionnels ne se distinguent plus.

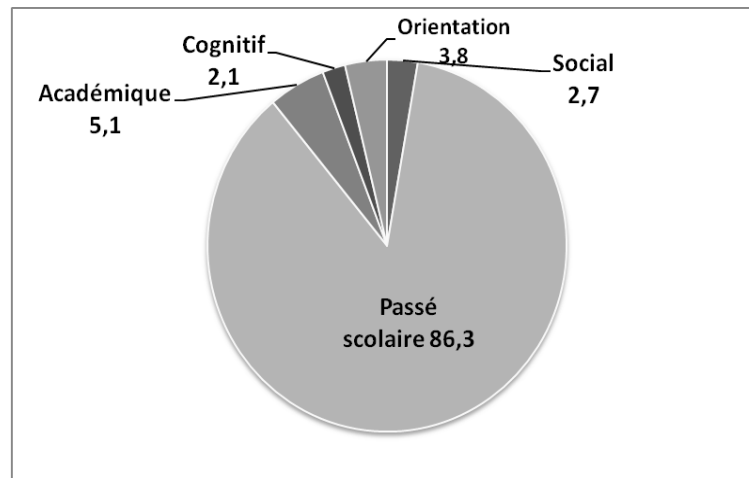
Tableau 8 : Modèle analysant la variance des notes de l'année

		Modèle 11 (N = 543)	
Modalité de référence	Modalité active	Coef.	Prob.
Constante		2,8	0,00
Filière fréquentée			
Droit	AES	+2,2	0,00
	Psychologie	+2,9	0,00
Garçon	Fille	+0,2	0,38
Retard scolaire	Age normal	+0,7	0,00
Non boursier	Boursier	-0,5	0,02
Série du bac			
Bac professionnel Bac technologique	Bac L	+1,9	0,00
	Bac ES	+2,7	0,00
	Bac S	+3,7	0,00
Pas de mention	Mention AB B TB	+2,1	0,00
Score en compréhension de l'écrit (DALF)		+0,14	0,00
Score au Raven (% de réussite)		+0,00	0,47
R ²		0,47	

Nous nous sommes enfin intéressés conjointement à l'ensemble des déterminants de la réussite en fin de première année avec une régression logistique et un phénomène intéressant apparaît. L'effet du niveau académique (score en compréhension de l'écrit) s'atténue puisque le coefficient attaché à cette variable n'est significatif qu'au seuil de 8%. En revanche, les performances au Raven ont tendance à exercer une influence positive sur les résultats, la probabilité d'erreur associée au coefficient étant de 10%. On ne peut bien sûr ici parler de tendance puisque le seuil des 5% d'erreur n'est pas atteint. Une autre information est fournie dans ce modèle ; les modalités de choix d'orientation universitaire jouent un rôle sur la réussite (ce qui n'était pas le cas dans les estimations précédentes relatives aux notes de fin d'année) : les étudiants qui ont effectué leur choix en fonction d'un projet professionnel ont une probabilité de réussite supérieure de 1,7 fois à celles des autres étudiants.

En conclusion

Les éléments et analyses exposés précédemment montrent que ce sont principalement les variables classiques liées au parcours scolaires des étudiants (retard scolaire, série et mention du bac) qui déterminent le succès au terme de la première année passé à l'université. Si l'on tente de rendre compte du poids de chaque groupe de facteurs sur la réussite, on obtient la situation visualisée par le graphique suivant. Sur la base de modèles logistiques et d'autres modèles intermédiaires qui intègrent successivement les groupes de variables explicatives, on peut chiffrer la contribution propre de chacun de ces groupes à l'explication des écarts de réussite entre étudiants (sur la base du R² de Nagelkerke). L'ensemble des variables explique environ 30% de la variance des résultats (en déduisant la contribution de l'appartenance à une filière) ; le graphique décompose cette variance en termes de pourcentages selon les différents groupes de facteurs.



Graphique 2 : Contribution des caractéristiques des étudiants à leur réussite en fin de première année universitaire

On remarque le poids très fort du passé scolaire (retard scolaire, série et mention du bac), le niveau académique (DALF) représentant 5% de cette variance expliquée, les modalités de choix d'orientation 3,8%, les caractéristiques sociales (le fait d'être boursier ou non) moins de 3% et les performances cognitives (Raven) seulement un peu plus de 2%. Les compétences académiques et les capacités cognitives ne jouent donc à elles seules qu'un rôle limité pour expliquer les différences de réussite entre étudiants, l'essentiel de l'influence de ces variables s'étant exprimé auparavant tout au long de la scolarité. En conclusion, il apparaît que les parcours des étudiants à l'entrée à l'université sont très marqués par la nature de leur scolarité passée et ces résultats interrogent plus largement sur les procédures d'orientation et les choix d'études à l'issue de l'enseignement secondaire. Le fait que le choix de la filière repose sur un projet professionnel soit un facteur associé à la réussite est un résultat important dans la mesure où il vient nuancer les déterminismes liés au passé scolaire. Au-delà des forts effets liés à la série du baccalauréat, l'influence autonome du niveau de compétence en compréhension de l'écrit montre que les compétences acquises, quelle que soit la filière fréquentée dans l'enseignement secondaire, contribuent à la réussite à l'université. Enfin, si les capacités cognitives ne semblent pas, à l'issue de cette recherche, exercer un effet autonome sur les résultats des étudiants, il reste à démontrer que cela reste vrai dans la suite du parcours universitaire au niveau de la licence ou du master.

ⁱ Chaque score est la moyenne centrée réduite des scores centrés réduits obtenus à chaque épreuve.

ⁱⁱ Les corrélations négatives viennent du fait que l'échelle des scores de vitesse de traitement est différente des deux autres échelles (RAVEN et mémoire de travail) puisque, pour cette dimension, les scores les plus élevés correspondent aux performances les plus faibles.

L'APPRENTISSAGE PAR LES COMPETENCES

LE RÔLE DE L'ÉVALUATION AU COURS DE L'APPRENTISSAGE PAR COMPÉTENCE

Zoualfakar JAMMOUL

Laboratoire : ICAR (Interactions, Corpus, Apprentissages, Représentations)

Ecole doctorale : ED 485(EPIC) : Education, Psychologie, Information, Communication

Université Lumière Lyon2

zoualfakar@hotmail.fr

Mots-clés : *apprentissage, compétence, situation, évaluation, dispositif.*

Résumé. *Le concept de compétence professionnelle qui a émergé au cours des années quatre-vingt du siècle dernier a conduit à une série de changements concernant les méthodes d'élaboration des référentiels des programmes d'apprentissage, d'enseignement et d'évaluation. Aussi, les chercheurs en sciences de l'éducation ont proposé un certain nombre de démarches didactiques à propos des compétences professionnelles qui dépendent essentiellement de l'effort (mental et manuel) de l'apprenant et du suivi par l'enseignant. Mais la question est de savoir comment nous concevons l'apprentissage ; comment nous évaluons le travail des élèves au cours de l'apprentissage ; quel est le rôle de l'évaluation dans l'acquisition et la validation des compétences. Dans cet article, nous allons essayer d'aborder le lien entre les méthodes d'apprentissage et le rôle de l'évaluation au cours de l'apprentissage.*

Introduction :

L'évaluation est l'un des aspects les plus importants de l'apprentissage. En effet, nous ne pouvons pas mettre un plan d'apprentissage sur une durée à long ou court terme, sans connaître le niveau des élèves ciblés et nous ne pouvons pas progresser dans les processus d'apprentissage sans connaître les progrès des élèves au cours du processus d'apprentissage puisque si les élèves n'atteignent pas l'objectif de chaque processus, ils auront une difficulté lors de la phase suivante. Par ailleurs, l'obtention du diplôme, validant l'acquisition des compétences pour l'élève, est validée par une évaluation finale qui permet également à l'enseignant d'évaluer la réussite de sa démarche pédagogique et l'efficacité de ses outils. Nous pouvons donc dire que la relation entre les méthodes d'évaluation et les méthodes d'enseignement dans la situation de l'apprentissage par compétences est réciproque. Dans cet article, nous allons aborder la question des méthodes d'enseignement et d'évaluation à partir du dispositif que nous avons utilisé dans le contexte de notre recherche doctorale en sciences de l'éducation. Les résultats obtenus suite à notre observation sur le terrain concernant l'acquisition des compétences au sein de la formation initiale professionnelle dans la filière restauration vont nous permettre de proposer une démarche pédagogique visant à améliorer la situation d'apprentissage.

1. La définition de compétences dans une situation problème :

Nous observons qu'il existe de nombreuses définitions du terme « compétence ». Certains se concentrent sur les savoirs, savoirs- faire et savoirs-être ; d'autres sont axés sur la réalisation de la tâche dans une situation de travail ou dans une situation d'apprentissage similaire à une situation de travail où l'apprenant doit agir face à une tâche complexe ou à une situation problème. Par exemple, la compétence selon Labrauffe est la capacité à mettre en œuvre des connaissances, savoir-faire et comportements en situation. (LABRUFFE, 2009, p. 101). Cette définition ressemble à la définition de Gonnin-Bolo qui a souligné que la compétence est la mise en œuvre en situation professionnelle de capacités qui permettent d'exercer convenablement une fonction ou une activité. (GONNIN-BOLO, 2005, p. 37). Quant à Vecchi, il définit une compétence comme une aptitude à agir efficacement dans une situation complexe en utilisant des acquis élémentaires (ensemble de capacités, d'attitudes, de connaissances notionnelles). (VECCHI, 2011, p. 95). Aussi, Scallon a résumé les compétences par les points suivants :

- Qualité globale de la personne ;
- Intégration appropriée des savoirs, savoirs- faire et savoirs-être ;
- Un système de connaissances conceptuelles et procédurales ;
- Etat de la personne ;
- Ensemble intégré d'habilités ;
- Capacité d'action. (SCALLON, 2007, p. 104).

Malgré les différentes définitions apportées au concept de compétence, nous relevons des termes communs : savoirs, savoirs- faire, savoirs- être, capacité, attitude, connaissance, activité, démarche, situation, tâche, action, complexe, comportement, performance et habilité. Cela montre la souplesse de ce terme et sa complexité, mais nous pouvons dire que la compétence est l'action de l'apprenant en face d'une tâche, action ou un problème dans une situation complexe lorsque l'apprenant doit analyser la tâche et la situation, qu'il doit ensuite remobiliser ses connaissances « théoriques et pratiques » pour sélectionner ce qu'il croit être efficace dans cette situation. Cette situation impose à l'élève de combler le manque concernant ses connaissances afin d'élaborer un plan de travail et de fixer les objectifs finaux sans oublier de choisir ou de tenir compte des paramètres qui serviront à l'évaluation, positive ou négative, de la mise en œuvre du plan et des objectifs. L'apparition de ce concept apportait, selon Romainville 1996, trois changements importants sur la situation apprentissage :

- Le passage « d'un apprentissage centré sur les matières à un apprentissage centré sur l'élève apprenant » ;
- Le passage « d'un apprentissage centré sur des acquis peu mobilisables à un apprentissage centré sur un potentiel d'action » ;
- Le passage « d'un apprentissage de connaissances à un apprentissage de savoir-faire, de savoir-réfléchir ». (DELORY, 2002, p. 22).

L'apparition des compétences a conduit à des changements substantiels au rôle de l'enseignant et de l'élève ainsi que des changements radicaux dans les méthodes d'apprentissage, où nous constatons que la majorité des programmes d'apprentissage se concentre sur un apprenant en face d'un problème ou d'une situation complexe. La situation-problème, c'est la résolution qui permet de finaliser des séquences d'apprentissage cognitif. Ce qui est nouveau, c'est la création de situations au cours desquelles la production de nouvelles capacités n'est pas distincte de la mobilisation de ces capacités. (BARBIER, 1996, p. 38). Donc, selon le concept de compétence, l'apprentissage doit être fait dans la situation complexe ou problème, mais comment se définit une situation problème et quelles sont ses caractéristiques, selon B. Douady (1988, P.99) et P. Mérieux (1987). La situation d'apprentissage doit répondre aux caractéristiques suivantes : cette situation propose une tâche à accomplir qui fait problème à l'élève parce qu'il ne dispose pas pour le moment de tout ce qui lui est nécessaire pour s'en acquitter. Et, ce qui manque à l'élève, c'est précisément ce que l'on a prévu comme apprentissage, de telle sorte que sa réalisation passe par la maîtrise d'un objectif sous-jacent. (GILLET, 1997, p. 121). Scallon a indiqué que dans une pédagogie de situation, le déroulement des activités est structuré d'après une succession de problèmes à résoudre, de tâches complexes à accomplir ou de projets concrets à réaliser. Chaque problème doit faire appel à une multitude d'éléments appartenant au répertoire cognitif et affectif de l'élève. Dans ce type de pédagogie, le point de départ est la saisie par l'élève d'un problème complexe entendu dans un sens très général. (SCALLON, 2010, p. 34). Selon cette logique, l'objectif n'est pas la complexité de l'apprentissage mais l'incitation des élèves à réfléchir, à rechercher, à analyser, à renforcer leur autonomie et initiative afin d'améliorer leur capacité d'agir en face d'une situation problème. Mais si l'apprentissage, selon cette logique, se concentre sur le travail de l'élève dans une situation problème, quel est alors le rôle de l'enseignant au cours d'une situation d'apprentissage. En réalité, le rôle de l'enseignant, en fonction de cette logique, a connu des changements puisqu'il transpose des informations et les évalue. En fait, il organise des situations complexes, invente des problèmes et des défis, propose des énigmes ou des projets ; son rôle est très important, mais il ne tient plus le devant de la scène et monopolise plus la parole. (MORLAIX, 2009, pp. 80-81).

2. Les démarches d'apprentissage par problème.

Nous ne pouvons pas parler de l'évaluation des compétences selon la logique d'apprentissage en situation sans expliquer au début certaines démarches concernant la construction des compétences par l'apprentissage en situation problème, ce que nous allons tenter d'expliquer en référence à G. Meyer et M.N Simonard puis Guy Le Boterf et l'IUFM.

2.1 La démarche de G. Meyer et M.N Simonard.

G. Meyer et M.N Simonard ont développé un modèle d'apprentissage par le problème qui comprend quatre étapes essentielles selon l'ordre suivant :

- *Le conflit cognitif* : l'élève se trouve en face d'une tâche complexe qu'il va essayer de résoudre ; ce stade permet à chaque élève de diagnostiquer ses connaissances et ses expériences passées ainsi que sa motivation. C'est un travail individuel ;
- *Le conflit socio- cognitif* : Cette étape comprend le travail au sein d'un petit groupe de quatre élèves avec des conseils pour comparer les résultats obtenus avec les résultats de l'étape précédente. Elle s'appuie principalement sur l'interaction et l'échange d'opinions, l'interprétation et les explications entre les membres du groupe afin d'obtenir une réponse consensuelle ;

- *La métacognition* : Cette phase, animée par le formateur ou l'enseignant, est le lieu de présentation d'une série de questions comme par exemple : quelle est votre réponse ? Comment en êtes-vous arrivé là ? Qu'est-ce qui vous permet de dire cela ? Cette phase est nécessaire pour que l'étudiant prenne conscience de sa réponse, de ses démarches et des informations qu'il utilise pendant la résolution de la tâche complexe « problème » ;
- *L'évaluation* : cette étape permet de repérer les difficultés et les étapes qui doivent être surmontées afin de réussir dans une réelle situation de travail. Cette évolution formative consolide l'apprentissage. (GILLET, 1997, pp. 121-122).

2.2 La démarche de Guy Le Boterf.

En fait, le modèle de G. Meyer et M.N Simonard n'est pas trop éloigné du modèle de Guy Le Boterf, qui comprend aussi quatre étapes principales pour l'apprentissage par une tâche complexe en vue de la construction des compétences selon l'ordre suivant :

- *Le moment de l'expérience* vécue correspondant à l'action du sujet dans une situation donnée, à la mise en application de ses savoirs ;
- *Le moment de l'explication* correspondant à la verbalisation, à la description de l'action réalisée permettant au sujet une mise à distance, un début de réflexion, une prise de conscience des faits: étapes menant à une représentation de l'action ;
- *Le moment de la conceptualisation* correspondant à l'élaboration d'un savoir pragmatique fondé sur la construction de schèmes opératoires, d'une structure conceptuelle répondant à différentes situation ;
- *Le moment du transfert* ou de la transposition correspondant à la mise à l'épreuve des schèmes opératoires. Le transfert consiste en l'accommodation des schèmes à la nouvelle situation ce qui implique une remise en cause des représentations afin de reconstruire une adéquation. (LE BOTERF, 2007, pp. 117-128).

2.3 La démarche de L'IUFM.

La démarche didactique de l'IUFM de Rouen proposée lors de la formation des enseignants de la filière restauration, est fondée sur l'action de l'enseignant et de l'élève ; cette démarche comprend trois étapes essentielles :

- *La première étape comprend* : l'explication, la démonstration des techniques et des gestes professionnels ;
- *La deuxième étape comprend* : l'analyse de l'expérience simple vécue par les élèves, les résultats, les réflexions sur les méthodes utilisées ;
- *La troisième étape comprend* : l'appréhension de la part des élèves des contraintes de production et de commercialisation, la justification des produits retenus et pouvoir apporter sa valeur ajoutée à la réalisation ; (IUFM : hra.spip.ac-rouen.fr/IMG/ppt/Seminaire_formateur_de_formateur.ppt).

Suite à la présentation des trois méthodes, nous remarquons que lors de la construction de la compétence professionnelle au cours de l'apprentissage, nous devons utiliser trois types d'évaluation. La première est l'évaluation diagnostique où l'élève doit se rappeler ses expériences « théoriques-pratiques » antérieures en vue de résoudre le problème auquel il est confronté et que l'enseignant a choisi en fonction du niveau de l'élève. La deuxième est l'évaluation formative où les élèves doivent être conscients de leurs démarches et de leurs nouvelles expériences « théoriques et pratiques » pour résoudre un problème ou réaliser une tâche complexe dans une situation problème qui permet à l'enseignant de mesurer l'état d'avancement de la procédure d'apprentissage. La troisième, c'est l'évaluation sommative où les élèves peuvent mesurer les connaissances acquises au cours de leur apprentissage en comparant avec leur situation initiale.

Nous allons donc clarifier le rôle de ces trois évaluations pendant la période d'apprentissage.

3. L'évaluation au cours d'apprentissage en situation problème.

3.1 L'objectif de l'évaluation.

« Évaluer c'est facile ; quand on a précisé l'objectif qu'on poursuit ; quand on a donné des critères d'évaluation ». (MEYER & SIMONARD, 1990, p. 92). Certains pourraient penser que l'objectif de l'évaluation est d'obtenir un certificat ou déterminer le niveau de réussite des élèves, mais en fait, cela est une partie de l'évaluation. L'évaluation en fonction de l'apprentissage est une méthode d'enseignement ou du moins une partie essentielle des méthodes de l'enseignant qui ne peut pas commencer les processus d'apprentissage sans connaître le niveau du groupe cible. Et, l'évaluation au cours du processus d'apprentissage est essentielle d'abord pour voir l'état d'avancement et d'efficacité des méthodes utilisées ; l'évaluation finale va nous aider à connaître les connaissances accumulées. Nous allons donc essayer d'expliquer les quatre types d'évaluation et d'illustrer leur importance au cours du processus d'apprentissage, à savoir : l'évaluation diagnostique, formative, l'auto-évaluation et l'évaluation sommative. En effet, Jean Cardinet fixe à l'évaluation quatre buts fondamentaux :

- Améliorer les décisions relatives à l'apprentissage de chaque élève ;
- Améliorer la qualité de l'enseignement général ;
- Informer sur sa progression l'enfant et ses parents ;
- Décerner des certificats nécessaires à l'élève et à la société. (GALIANA, 2005, p. 25).

Or, que pouvons-nous évaluer au cours de l'action de formation ? Le Boterf distingue « trois niveaux d'effets d'une action de formation : l'évaluation des acquis (connaissances, capacités...) en fin de formation, l'évaluation des pratiques construites et mises en œuvre pour gérer les situations professionnelles, l'évaluation des impacts sur les performances d'une unité (sur les paramètres sensibles à l'action de formation » (LE BOTERF, 2011, p. 540). Cela signifie que la relation entre les méthodes d'apprentissage et d'évaluation est interactive. L'apprentissage ne peut évoluer sans l'évaluation qui ne peut se faire sans l'identification des objectifs et des méthodes d'apprentissage. Wittgenstein résume ainsi cette situation : « dis-moi comment tu cherches, je te dirai ce que tu cherches », « ce que vous évaluez dépend du concept de compétence que vous utilisez. Selon ce que vous voulez évaluer, si les personnes sont compétentes ou ont des compétences, vous n'observez pas la même chose. » (LE BOTERF, 2008, p. 102). Pour évaluer, nous devons donc d'abord fixer les objectifs et le plan d'apprentissage, et pour fixer le plan d'apprentissage et les méthodes d'enseignement et d'évaluation nous devons connaître le niveau des élèves c'est-à-dire l'évaluation diagnostique.

3.2 L'évaluation diagnostique.

Le diagnostique ne porte pas sur une petite partie, mais prend en considération l'ensemble organisé d'une situation. (MEYER & SIMONARD, 1990, p. 97). Il est lié à toutes les étapes et les processus d'apprentissage. En éducation, l'évaluation diagnostique a été liée à un « test », mais pas au type et à la nature du test qui sont les plus importants, c'est la manière d'employer les résultats pour améliorer la qualité de l'apprentissage et donc le but ultime de développer les compétences des élèves. Selon le dictionnaire de l'évaluation et de recherche en éducation De Landsheere (1992) « les tests diagnostiques ont pour objectif de découvrir les faiblesses et les habitudes défectueuses dans tous les domaines de l'apprentissage scolaire ». (REY & all, 2006, p. 38). L'évaluation diagnostique a une fonction préventive qui nous permet de définir les éléments qui pourraient affecter la qualité de l'apprentissage à l'avenir, (à titre d'exemple, les intérêts des élèves, leur motivation, leur expérience ou leur maturité). Bien qu'elle contribue à déterminer les modalités et les méthodes d'enseignement au cours d'apprentissage en conformité avec le niveau général de groupe cible, elle prend en compte également les différences individuelles et permet alors de repérer les élèves qui ont des difficultés particulières et qui ont besoin de soutien individuel afin de rendre l'homogénéité entre le niveau des membres du groupe. Scallon souligna que nous pouvons utiliser l'évaluation diagnostique au cours du processus d'apprentissage quand nous affrontons l'un des problèmes persistant et après l'utilisation de toutes les méthodes de nature pédagogique possible pour la résoudre et rechercher dans les aspects extérieurs qui peuvent être mis en cause : l'état de santé de l'élève, son milieu familial, ses intérêts et sa motivation. (SCALLON, 2010). Ainsi, nous pouvons définir les objectifs de l'évaluation diagnostique :

- Déterminer le niveau général du groupe cible ;
- Identifier les difficultés de chaque élève ;
- Aider l'enseignant à mettre le plan d'apprentissage ;
- Aider l'enseignant à choisir les méthodes et les moyens d'apprentissage ;
- Aider les élèves à découvrir et à accomplir leurs expériences nécessaires pour réussir ;
- Permettre à l'enseignant d'améliorer le niveau des élèves au cours de l'apprentissage à travers la comparaison entre les résultats de l'évaluation diagnostique et formative et le produit final de l'apprentissage à travers la comparaison entre les résultats de l'évaluation diagnostique et sommative.

3.3 L'évaluation formative.

Si l'objectif de l'évaluation diagnostique est de mesurer le niveau des élèves pour établir le plan de l'apprentissage, l'objectif de l'évaluation formative est de connaître les progrès de l'élève au cours de l'apprentissage et de mesurer l'efficacité des méthodes et des outils de l'enseignant. « L'évaluation formative viserait à servir l'apprentissage et l'enseignement, et non la sélection des élèves et la sanction des études. En ce sens, elle serait centrée sur les processus, et non sur les seuls produits de l'apprentissage. » (MORRISSETTE, 2009, p. 13). Et, Andrade et Cizek définissent les objectifs de l'évaluation formative :

- Identifier les forces et les faiblesses de l'élève ;
- Aider l'élève à réfléchir à ses propres processus d'apprentissage et à guider ses démarches en vue de progresser ;
- Croître son autonomie et sa prise de responsabilité face à ses apprentissages ;
- Orienter la planification de l'enseignement ». (MORRISSETTE, 2010, pp. 4-5).

Par ailleurs, Talbot souligna que si l'évaluation formative fournit à l'élève des informations utiles sur et pour ses apprentissages, elle apporte aussi à l'enseignant des indications primordiales sur ses pratiques d'enseignement notamment du point de vue de son rôle dans la régulation des apprentissages. L'évaluation formative intéresse donc à la fois l'apprentissage et l'enseignement. » (TALBOT, 2009, p. 9). Et, comme les métiers de l'enseignement sont devenus actuellement très complexes, l'enseignement doit prendre en compte les diversités des niveaux des élèves et se concentrer aussi sur l'amélioration du niveau collectif du groupe cible. En fait, Vial dit que l'évaluation est un dispositif d'apprentissage en trois phases : l'orientation de la tâche en exhibant ses normes, l'exécution de la tâche en y appliquant les critères procéduraux pour la rendre plus efficace et l'entraînement au contrôle des procédures exhibées. C'est un dispositif d'apprentissage du contrôle des outils de la tâche, d'entraînement progressif à l'intériorisation de ses normes. » (VIAL, 2000, p. 102). En réalité, l'enseignant ne peut pas élaborer un plan d'apprentissage sans l'évaluation diagnostique et ne peut pas avancer les processus de l'apprentissage sans l'évaluation formative, donc, la relation entre les méthodes d'enseignement et les processus de l'apprentissage est réciproque et interdépendante.

Enfin, nous pouvons dire que l'évaluation formative permet aux élèves et aux enseignants de :

- Connaître les progrès réalisés sur le plan de l'apprentissage ;
- Connaître le niveau de développement des élèves, collectivement et individuellement ;
- Connaître l'efficacité des outils, des méthodes et du rôle de l'enseignant ;
- Connaître les difficultés collectives et individuelles des élèves ;
- Modifier le plan de l'apprentissage selon les difficultés rencontrées par les élèves ;
- Rendre l'apprentissage plus flexible et facilement adaptable aux besoins des élèves.

Nous pouvons dire aussi que l'enseignant peut bénéficier de l'évaluation formative comme de l'auto-évaluation parce qu'elle lui montre les failles de son plan, de son rôle, de ses méthodes et de ses outils.

3.4 L'auto-évaluation.

L'auto-évaluation est efficace si les objectifs, les règles à atteindre sont clairement définies au préalable avant l'action. Et, l'enseignant, au centre de l'auto-évaluation, fixe les règles que les apprenants se doivent de respecter afin d'accomplir la tâche. Aussi, ce type d'évaluation permet aux élèves de mesurer l'acquisition de leurs compétences, pour les enseignants, de réajuster le dispositif d'enseignement. (MILLET & all, 2010) & (VIAL, 2001).

Elle peut donc d'après cette approche apporter des avantages pour les élèves concernant :

- Le développement de compétences métacognitives – les élèves apprennent à mieux rectifier le travail qu'ils sont en train de faire en vue d'en améliorer la qualité ;
- La responsabilisation accrue des élèves à l'égard de leur propre apprentissage, en raison du nombre accru de possibilités d'autoréflexion ;
- L'aptitude des élèves à la pensée critique ;
- L'amélioration de l'aptitude à la résolution de problèmes. (Ontario, 2007).

3.5 L'évaluation sommative.

L'évaluation sommative intervient après un ensemble de tâches d'apprentissage constituant un tout, correspondant par exemple à un chapitre du cours, à l'ensemble du cours d'un trimestre ; les examens périodiques et les interrogations d'ensemble sont donc des évaluations sommatives. C'est à partir d'une évaluation sommative que se dresse le plus souvent un classement des élèves entre eux. (REY & all, 2006, pp. 37-38). « L'évaluation sommative devrait idéalement survenir au terme d'un long processus d'enseignement et d'apprentissage afin de sanctionner, certifier, pour chaque étudiant, le degré de maîtrise des objectifs pédagogiques visés. Si elle devait se réaliser comme une évaluation rigoureusement terminale, l'évaluation sommative pourrait rendre compte des compétences telles qu'elles se présentent au moment où les jugements interviennent, c'est-à-dire à la fin de la période d'enseignement et d'apprentissage envisagée. En outre, la distinction à maintenir entre l'évaluation formative et l'évaluation sommative serait on ne peut plus nette. » (SCALLON, 2010, p. 17). Selon G. De Landsheere, l'évaluation sommative peut aussi avoir pour objet de déterminer dans quelle mesure un programme fonctionne bien dans son ensemble. » (REY & all, 2006, pp. 37-38). L'évaluation sommative est donc la dernière étape de l'apprentissage où nous observons l'amélioration définitive du niveau des élèves et le résultat final du processus d'apprentissage. Certains l'utilisent pour une classification des élèves entre eux et délivrer le diplôme qui reflète aussi l'atteinte des objectifs de l'apprentissage.

4. Le dispositif de l'évaluation.

Dans le contexte de notre recherche doctorale en sciences de l'éducation, nous avons effectué un stage d'observation auprès d'un lycée professionnel, en formation initiale « filière restauration ». Nous avons établi une grille d'observation qui comprend 12 items essentiels pour repérer les faiblesses des élèves selon les paramètres compris dans le référentiel de compétences de la formation.

		A	B	C	D
1	Savoir-être				
2	Savoirs				
3	Savoirs –faire				
4	Autonomie				
5	Initiative personnelle				
6	Raisonnement pourquoi				
7	Explication comment				
8	Rapidité				
9	Réalisation de la tâche				
10	Echange avec l'enseignant				
11	Echange entre élèves				
12	Transfert des compétences sur le terrain au cours du stage				

Tableau 1 : le dispositif de compétence

Pour lire ce tableau : A = tout à fait insuffisant ; B = plutôt insuffisant ; C = plutôt suffisant ; D = tout à fait suffisant.

L'objectif de cette grille d'observation est de repérer les difficultés des élèves en situation d'apprentissage et leurs faiblesses en les croisant avec les méthodes d'enseignement et les contraintes sur le terrain afin de contribuer à améliorer la situation d'apprentissage. Pour atteindre cet objectif, nous avons observé pendant une année scolaire 68 élèves et nous avons effectué deux observations pour chaque élève selon que l'action a été réalisée au préalable ou pas. L'élaboration de cette grille d'observation est fondée essentiellement sur les éléments suivants :

- L'analyse du référentiel ministériel qui comprend essentiellement un groupe de compétences (savoirs-être, savoirs, savoirs- faire) ; suite à cette analyse, nous avons retenu 20 compétences concernant la technologie culinaire, que les élèves doivent maîtriser à la fin de leur formation initiale. (MED, 1998) ;
- L'analyse des démarches didactiques : IUFM et Guy Le Boterf ;
- Notre observation sur le terrain (lycée, entreprise) ;
- Les échanges sur le terrain avec les enseignants, les élèves et le tuteur.

Ces 12 items renvoient aux connaissances théoriques et pratiques que l'élève doit posséder afin de réaliser son action : avoir la capacité d'argumenter, de verbaliser, de réagir et de prendre l'initiative dans une situation d'apprentissage, de transférer les compétences dans une nouvelle situation, sans oublier l'interaction avec les autres membres de son équipe. Ensuite, nous avons élaboré pour chaque compétence les paramètres qui nous permettent de récolter les données sur le terrain.

Suite à notre observation sur le terrain, nous avons retenu la démarche pédagogique des enseignants du lycée professionnel et interroger par la suite son efficacité au vue de l'acquisition des compétences professionnelles au sein de la formation initiale.

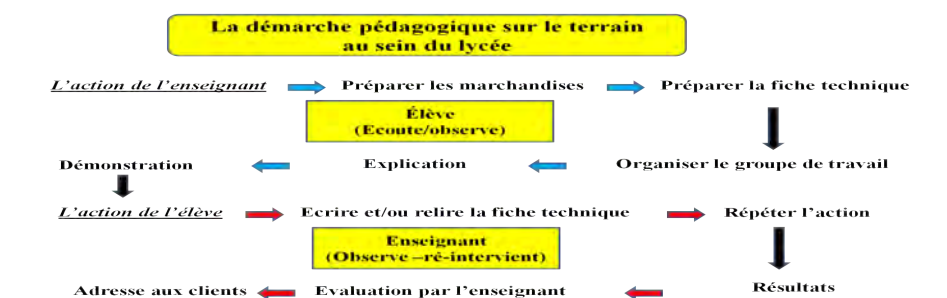


Figure 1 : la démarche pédagogique de l'enseignant

Nous avons utilisé le logiciel SPSS pour analyser nos données et appliqué « ANOVA, LSD multiples comparaisons et le test en % ». Suite à l'analyse, nous observons les résultats suivants :

4.1 Acquisition des compétences si l'action n'a pas été réalisée au préalable.

Nous avons observé qu'il y a une différence significative selon le test ANOVA (Tableau 2 - ANOVA 1) entre les élèves de « BEP1, BEP2, BAC1, BAC2 » selon que l'action n'a pas été réalisée au préalable concernant les compétences «Savoirs-être ; savoirs ; savoirs- faire ; autonomie ; raisonnement « pourquoi » ; explication « comment » ; rapidité » parce que le Sig 0,000 < 0,05. En revanche, il n'y a pas de différence significative entre les élèves de « BEP1, BEP2, BAC1, BAC2 » selon que l'action n'a pas été réalisée au préalable concernant les compétences « Initiative personnelle ; réalisation de la tâche ; échange entre élèves » parce que le Sig 0,362 ; 0,084 ; 0,349 > 0,05 et par rapport à la compétence 10 « échange avec l'enseignant » où 100 % des élèves « BEP1, BEP2, BAC1, BAC2 » ont le même niveau, c'est pourquoi le logiciel de SPSS n'analyse pas cette compétence.

ANOVA	1	2	3	4	5	6	7	8	9	10	11
Sig	,000	,000	,000	,000	,362	,000	,000	,000	,084		,349

Tableau 2 - ANOVA 1- l'action des élèves n'a pas été réalisée au préalable

Pour lire le tableau ci-dessus : 1= savoirs-être ; 2 = savoirs, 3 = savoirs - faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11 = échange entre les élèves.

Suite à l'application du test « LSD multiples comparaisons » (Tableau 3 « LSD multiple comparaisons »1), nous avons observé :

- Qu'il y a une différence significative entre les élèves de BEP1 d'une part et les élèves de BEP2, BAC1 et BAC2 d'autre part concernant les compétences «savoirs-être ; savoirs ; savoirs- faire ; autonomie ; raisonnement « pourquoi » ; explication « comment » ; rapidité » ;
- Qu'il y a une différence significative entre les élèves de BEP2 et BAC1 concernant la compétence 2 « savoirs » ;
- Qu'il y a une différence significative entre les BEP2 et les BAC2 concernant les compétences « savoirs ; raisonnement « pourquoi » ; explication « comment » et entre BAC1 et BAC2 concernant « raisonnement « pourquoi » ; explication « comment ».

		1	2	3	4	5	6	7	8	9	10	11
(I)classe	(I)classe	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig
BEP1	BEP2	,000	,000	,000	,000		,000	,000	,000			
	BAC1	,000	,000	,000	,000		,000	,000	,000			
	BAC2	,000	,012	,000	,000		,000	,000	,000			
BEP2	BAC1	,502	,009	1,000	,680		,554	1,000	,789			
	BAC2	,905	,009	,606	,659		,013	,017	,392			
BAC1	BAC2	,458	,533	,560	,901		,023	,008	,468			

Tableau 3 : « LSD multiple comparaisons » 1 : l'action des élèves n'a pas été réalisée au préalable

Pour lire le tableau ci-dessus : 1 = savoirs-être ; 2 = savoirs, 3 = savoirs- faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11= échange entre les élèves.

Suite à l'application du test de pourcentage (Tableau 4 « en % » 1), nous observons que la majorité absolue de notre échantillon n'a pas acquis les compétences « savoirs, savoirs- faire ; autonomie ; initiative personnelle ; raisonnement « pourquoi » ; explication « comment » ; rapidité ; réalisation de la tâche ; échange entre les élèves » pour la réalisation d'une action jamais réalisée au préalable. Par contre, nous observons que la majorité des élèves de notre échantillon ont acquis les compétences « savoirs-être et les échanges avec l'enseignant ».

En %	1	2	3	4	5	6	7	8	9	10	11
Tout à fait insuffisant	14,7	58,8	52,9	44,1	97,1	50,0	52,9	47,1	10,3		61,8
Plutôt insuffisant	17,6	41,2		55,9	2,9	47,1	44,1	47,1	86,8		32,4
Plutôt suffisant	25,0		47,1			2,9	2,9	5,9	2,9		5,9
Tout à fait suffisant	42,6									100,0	

Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0
Insuffisant en %	32,3	100	52,9	100	100	97,1	97,1	94,2	97,1		94,2
Suffisant en %	67,6		47,1			2,9	2,9	5,9	2,9	100,0	5,9

Tableau 4 « en % » 1 : l'action des élèves n'a pas été réalisée au préalable

Pour lire le tableau ci-dessus : 1 = savoirs-être ; 2 = savoirs, 3 = savoirs- faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11 = échange entre les élèves.

4.2 Acquisition des compétences si l'action a été réalisée au préalable.

Nous avons observé qu'il y a une différence significative selon le test ANOVA (Tableau 5 – ANOVA 2) entre les élèves de « BEP1, BEP2, BAC1, BAC2 » selon que l'action a été réalisée au préalable concernant les compétences « savoirs-être ; savoirs ; savoir-faire ; autonomie ; raisonnement « pourquoi » ; explication « comment » ; rapidité ; réalisation de la tâche ; » parce que le Sig 0,000 et 0,001 < 0,05. En revanche, il n'y a pas de différence significative entre les élèves pour la « réalisation de la tâche » entre BEP1, BEP2, BAC1, BAC2 » selon que l'action a été réalisée au préalable concernant les compétences « initiative personnelle ; échange entre élèves » parce que le Sig 0,617 ; 0,398 > 0,05 et par rapport à la compétence 10 « échange avec l'enseignant » où 100 % des élèves « BEP1, BEP2, BAC1, BAC2 » ont le même niveau, c'est pourquoi le logiciel de SPSS n'analyse pas cette compétence.

ANOVA	1	2	3	4	5	6	7	8	9	10	11
Sig	,000	,000	,000	,000	,617	,000	,000	,000	,001		,398

Tableau 5 « ANOVA » 2 : l'action des élèves a été réalisée au préalable

Pour lire le tableau ci-dessus : 1 = savoirs-être ; 2 = savoirs, 3 = savoirs- faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11 = échange entre les élèves.

Suite à l'application du test « LSD multiples comparaisons » (Tableau 6 « LSD multiple comparaisons » 2), nous avons observé :

- Qu'il y a une différence significative entre les élèves de BEP1 d'une part et les élèves de BEP2, BAC1, BAC2, d'autre part concernant les compétences « savoirs-être ; savoirs ; savoirs- faire ; autonomie ; raisonnement « pourquoi » ; explication « comment » ; rapidité » ;
- Qu'il n'y a pas une différence significative entre les élèves de BEP2 et BAC1 concernant « savoirs-être ; savoirs, savoirs- faire ; autonomie ; raisonnement « pourquoi » ; explication « comment » ; rapidité ; réalisation de la tâche » ;
- Qu'il y a une différence significative concernant la compétence « réalisation de la tâche » entre les BEP2 et BAC2 et entre les BAC1 et BAC2.

		1	2	3	4	5	6	7	8	9	10	11
(I)classe	(I) classe	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig	Sig
BEP1	BEP2	,000	,000	,000	,000		,000	,000	,000	,049		
	BAC1	,000	,000	,000	,000		,000	,000	,000	,017		
	BAC2	,000	,000	,000	,000		,000	,000	,000	,000		
BEP2	BAC1	,464	1,000	,516	1,000		,835	,677	,834	1,000		
	BAC2	,093	,606	,908	,671		,506	,505	,266	,026		
BAC1	BAC2	,213	,560	,647	,632		,349	,260	,150	,013		

Tableau 6 « LSD multiple comparaisons » 2 : l'action des élèves a été réalisée au préalable

Pour lire le tableau ci-dessus : 1 = savoirs-être ; 2 = savoirs, 3 = savoirs- faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11 = échange entre les élèves.

Suite à l'application du test de pourcentage (Tableau 7 « en % » 2), nous observons que la majorité de notre échantillon n'a pas acquis les compétences « savoirs, savoirs- faire ; autonomie ; initiative personnelle ; raisonnement « pourquoi » ; explication « comment » ; rapidité ; réalisation de la tâche ; échange entre les

élèves » pour une action réalisée au préalable. Par contre, nous observons que la majorité des élèves de notre échantillon ont acquis les compétences « savoirs-être et les échanges avec l'enseignant ».

En %	1	2	3	4	5	6	7	8	9	10	11
Tout à fait insuffisant		10,3	20,6	32,4	98,5	23,5	23,5	22,1	5,9		64,7
Plutôt insuffisant	20,6	47,1	45,6	48,5	1,5	39,7	41,2	39,7	83,8		33,8
Plutôt suffisant	26,5	42,6	33,8	19,1		36,8	35,3	38,2	10,3		1,5
Tout à fait suffisant	52,9									10,0	
Total	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	10,0	100,0
Insuffisant en %	20,6	57,4	66,2	80,9	100	63,2	64,7	61,8	89,7		98,5
Suffisant en %	79,4	42,6	33,8	19,1		36,8	35,3	38,2	10,3	10,0	1,5

Tableau 7 « en % » 2 : l'action des élèves a été réalisée au préalable

Pour lire le tableau ci-dessus : 1 = savoirs-être ; 2 = savoirs, 3 = savoirs- faire ; 4 = autonomie ; 5 = initiative personnelle ; 6 = raisonnement « pourquoi » ; 7 = explication « comment » ; 8 = rapidité ; 9 = réalisation de la tâche ; 10 = échange avec l'enseignant ; 11= échange entre les élèves.

Donc malgré les différences entre les élèves de « BEP1, BEP2, BAC1, BAC2 » concernant les compétences « savoirs, savoirs- faire ; autonomie ; initiative personnelle ; raisonnement « pourquoi » ; explication « comment » ; rapidité ; réalisation de la tâche » selon que l'action a été réalisée ou pas au préalable, nous observons que la majorité des élèves de notre échantillon n'ont pas acquis ces compétences.

4.3 Démarche pédagogique proposée.

Au vu des résultats recueillis et des démarches pédagogiques existantes, nous avons proposé une démarche adaptée au contexte d'apprentissage, en tenant des contraintes de fonctionnement de l'établissement, pour permettre une meilleure acquisition des compétences et formation des élèves en fin de formation initiale.



Figure 2 : la démarche pédagogique proposée

En fait, nous ne pouvons pas développer l'apprentissage sans l'évaluation et sans connaître les faiblesses de notre plan, de nos outils, nos moyens et nos méthodes, donc il n'y a pas d'apprentissage sans évaluation et vice-versa.

Perspectives :

1. Les enseignants devraient avant de mettre le plan d'apprentissage, choisir leurs méthodes et leurs moyens, faire une évaluation diagnostique. Certains enseignants pensent qu'ils connaissent le niveau de leurs élèves donc ils n'effectuent pas ce type de l'évaluation, mais elle est très importante car elle leur permet de sonder et de découvrir les compétences sous-jacentes chez les élèves et permet également aux élèves de découvrir les compétences nécessaires pour réussir ;
2. L'évaluation formative est importante pour repérer les difficultés, les faiblesses des élèves ainsi que les failles du projet des enseignants et pour montrer l'inadéquation des méthodes et des moyens au cours de l'apprentissage ;

3. L'auto-évaluation est très importante, car elle encourage les élèves à faire l'analyse et l'autocritique. Elle donne l'occasion aux enseignants de réajuster leur projet pédagogique ;
4. L'évaluation sommative est importante car elle donne une indication de la réussite du projet et du processus d'apprentissage. Elle amène également à revoir le développement du projet, des méthodes et démarches en vue d'obtenir de meilleurs résultats ; ce dispositif doit être flexible et ajustable pour leur permettre d'être impartial et objectif à la fois.

Conclusion :

L'observation des élèves sur le terrain pendant une année scolaire a permis de comparer les niveaux des élèves « BEP1, BEP2, BAC1, BAC2 » pour repérer les différences significatives en fonction de l'acquisition de compétences selon que l'action a été réalisée au préalable ou pas. Il s'avère que l'étape la plus difficile dans la construction du dispositif d'observation était l'élaboration des paramètres qui nous permettent de repérer les faiblesses et les difficultés des élèves en situation d'apprentissage étant donné qu'il faut tenir compte des objectifs à atteindre pour établir les paramètres d'observation. Par ailleurs, les enseignants ne perçoivent pas les paramètres de la même manière que l'observateur.

Nous devons donc pour améliorer la situation d'apprentissage discuter avec les enseignants, les élèves et les tuteurs concernant les résultats que nous avons obtenus grâce à ce dispositif qui peut être un outil d'évaluation formative et d'auto-évaluation. Cependant, il faut tenir compte également des contraintes de fonctionnement de l'établissement pour adapter au mieux les méthodes d'enseignement en apprentissage professionnel.

Bibliographie

- BARBIER, J.-M. (1996). Les évolutions de la formation et leurs enjeux. Dans J.-C. RUANO-BORBALAN, *Savoir former: Bilans et perspectives des recherches sur l'acquisition et la transmission des savoirs* (pp. 35-46). Paris : Les éditions DEMOS.
- DELORY, C. (2002). *L'évaluation des compétences dans l'enseignement fondamental : de quoi parle-t-on ?* Dans L. PAQUAY & ALL, *L'évaluation des compétences chez l'apprenant : Pratiques, Méthodes et Fondements* (pp. 21-35). Louvain : Presses Universitaires de Louvain.
- GALIANA, D. (2005). *Mémento de l'évaluation : analyser et améliorer sa pratique de l'évaluation*. Dijon : Educagri éditions.
- GILLET, P. (1997). *Construire la formation : Outils pour les enseignants et les formateurs, "collection pédagogie"* (éd. 4). Paris : ESF.
- CONNIN-BOLO, A. (2005). *Entreprises et écoles : une rencontre de professionnels*. Lyon : Institut National de Recherche Pédagogique.
- IUFM de Rouen. (s.d.). *Séminaire formateur de formateur*. Consulté le 07 05, 2010, sur hra.spip.ac-rouen.fr/IMG/ppt/Seminaire_formateur_de_formateur.ppt.
- LABRUFFE, A. (2009). *Les nouveaux outils de l'évaluation des compétences*. La Plaine Saint-Denis : AFNOR.
- LE BOTERF, G. (2007). *Construire les compétences individuelles et collectives* (éd. 4). Paris : EYROLLES.
- LE BOTERF, G. (2011). *Ingénierie et évaluation des compétences* (éd. 6). Paris : Editions d'Organisation.
- LE BOTERF, G. (2008). *Repenser la compétence pour dépasser les idées reçues : 15 propositions* (éd. 2). Paris : Organisation.
- MED. (1998). *Baccalauréat Professionnel Restauration*. Paris : Ministère de l'Education Nationale.
- MEYER, G., & SIMONARD, M.-N. (1990). L'évaluation diagnostique à l'école. Dans C. DELORME, & G. AVANZINI, *L'évaluation en questions* (éd. 3, pp. 91-103). Paris : ESF.
- MILLET, J.-G., & all. (2010). *Evaluation : trappe ou tremplin*. Paris : Editions d'Organisation.
- MORLAIX, S. (2009). *Compétences des élèves et dynamique des apprentissages*. Rennes : Presses Universités de Rennes.
- MORRISSETTE, J. (2009). *Manières de faire l'évaluation formative des apprentissages : Analyse interactionniste du savoir-faire d'enseignants du primaire*. Sarrebrouk : Universitaires Européennes.
- MORRISSETTE, J. (2010). Un panorama de la recherche sur l'évaluation formative des apprentissages. *Mesure et évaluation en éducation* (Vol. 33, N°3), pp. 3-4.
- Ontario. (2007). *Accroître la capacité*. Paris : Secrétariat de la littératie et de la numératie.
- REY, B., & all. (2006). *Les compétences à l'école : apprentissage et évaluation*. Bruxelles : De Boeck.
- SCALLON, G. (2007). *L'évaluation des apprentissages dans une approche par compétences* (éd. 2). Bruxelles : De Boeck.
- SCALLON, G. (2010). *L'évaluation formative*. Bruxelles : De Boeck.
- TALBOT, L. (2009). *L'évaluation formative : Comment évaluer pour remédier aux difficultés d'apprentissage*. Paris : Armand Colin.
- VECCHI, G. D. (2011). *Evaluer sans dévaluer et évaluer les compétences*. Paris : Hachette.
- VIAL, M. (2000). *Organiser la formation : Le pari sur l'auto-évaluation*. Paris : L'Harmattan.
- VIAL, M. (2001). *Se former pour évaluer: Se donner une problématique et élaborer des concepts* (éd. 1). Bruxelles : De Boeck.

L'APPROCHE PAR COMPETENCES ET L'EVALUATION CLINIQUE OBJECTIVE STRUCTUREE EN SOINS INFIRMIERS.

Corinne Bay, Haute Ecole Robert Schuman Libramont Belgique corinne.bay@hers.be

Mots-clés : Approche par compétences (APC) ; Evaluation Clinique Objective Structurée (ECOS) ; Soins infirmiers.

Résumé - Contexte : pour répondre aux exigences de l'évaluation et aux effets de la docimologie, des enseignants du baccalauréat en soins infirmiers de la Haute Ecole Robert Schuman ont choisi de modifier la stratégie d'évaluation des étudiants. La mise en place d'une Evaluation Clinique Objective Structurée (ECOS) a permis de favoriser l'approche par compétences mais aussi de réformer les méthodes pédagogiques des enseignants. **But :** l'enquête exploratoire descriptive avait pour finalités de mettre en évidence les changements observés dans la pratique des enseignants et d'évaluer le désir de maintenir la pérennité du projet mis en place en 2009. **Méthode :** la population visée était les enseignants (n=25) de la section qui encadrent les étudiants durant leur cursus et qui participent activement à leur évaluation clinique. Un questionnaire individuel composé de dix questions a été soumis par voie électronique. Une analyse quantitative et qualitative a permis de comparer les résultats pour en extraire les différences et similitudes. **Résultats :** l'unanimité est présente en ce qui concerne la décision de maintenir le concept et l'organisation de l'ECOS. Les trois quart des enseignants affirment rendre plus active leur pédagogie et intégrer l'évaluation au service de l'apprentissage, même si le temps nécessaire est incontestablement un élément qui plaide en défaveur de l'avancée du projet. **Conclusion :** L'ECOS est un projet prometteur, favorable à l'évaluation formative et formatrice tout en ayant le bénéfice de dynamiser une équipe d'enseignants en soins infirmiers.

1. Introduction

L'Evaluation Clinique Objective Structurée (ECOS) trouve ses origines avec R.Harden, en 1975, dans le milieu universitaire et a principalement été exploitée dans la discipline médicale, pour évaluer les futurs médecins en fin de cursus. Cette évaluation s'est ensuite étendue aux disciplines telles que la médecine vétérinaire, la pharmacologie, l'ostéopathie,... Elle se développe depuis peu dans le domaine paramédical, en particulier au Canada. Depuis 2000, l'ECOS constitue la partie pratique de l'examen professionnel d'admission de l'Ordre des Infirmiers et Infirmières du Québec, en vue de l'obtention du droit de pratique. Depuis lors, il est reproduit dans certains établissements de formation paramédicale canadiens (OIIQ, 2000).

Dès la rentrée académique 2009-2010, le corps professoral de la Haute Ecole Robert Schuman (HERS) appuyé par la direction ont opté pour cette méthode d'évaluation. Il s'agit de la première Haute Ecole belge à entamer cette pratique au sein de la section Baccalauréat en Soins Infirmiers (BSI). Le projet consiste à développer une évaluation formative et formatrice qui permette à la fois de motiver les étudiants à apprendre et de repérer les étudiants qui sont en difficulté ; de dynamiser l'équipe des enseignants et de l'aider à adapter ses pratiques pédagogiques grâce à l'évaluation de l'enseignement.

Pour intégrer l'évaluation au service de l'apprentissage, différentes étapes ont été indispensables : la création d'une « ECOS d'intégration ¹ » en cours d'année pour tous les étudiants; la possibilité pour les maîtres de formation pratique (MFP) de prendre connaissance et conscience du taux de réussite pour chaque station afin d'orienter leurs objectifs et méthodes d'apprentissage ; la mise en

¹ ECOS d'intégration ou ECOS « blanc » : ECOS qui se déroule dans les conditions identiques à l'évaluation certificative mais les résultats ne sont pas pris en compte dans la note d'enseignement clinique.

place de séances de remédiation pour favoriser l'implication de chaque étudiant dans son apprentissage. La cohérence du corps professoral se voit renforcée par l'introduction progressive des données probantes (Evidence Based Nursing) dans les cours.

La présentation, ci-dessous, d'une étude exploratoire descriptive vise à informer les modifications de pratiques pédagogiques favorisées par l'introduction de l'ECOS. Il existe actuellement peu de références bibliographiques mais surtout peu d'études, dans la discipline en soins infirmiers, permettant de justifier notre choix.

2. Motivations

Jusqu'en 2006, l'évaluation clinique des étudiants de la section soins infirmiers se faisait par une épreuve certificative, réalisée le dernier jour du dernier stage. Dans le cadre de la démarche « Qualité », l'analyse de cette épreuve a montré qu'elle provoquait de nombreuses insatisfactions de part et d'autre : trop grande différence entre les situations cliniques choisies, subjectivité des interrogateurs, manque d'utilisation de critères directement liés au référentiel de formation par compétences progressivement créé, ... On ne retrouvait pas dans cette épreuve les caractéristiques attendues d'un instrument d'évaluation (validité, reproductibilité, valeur éducative, ...). En particulier, cette épreuve ne témoignait pas de la mise en œuvre de la compétence 8 du CAPAES² : « Maîtriser et utiliser les outils d'évaluation des apprentissages adaptés à l'enseignement dispensé et pourvoir répondre de ses choix » (Décret 2002).

En 2007-2008, la première alternative proposée par les enseignants a été de créer une épreuve dite « intégrée » en complémentarité à l'épreuve certificative clinique. Celle-ci permettait à tous les étudiants de réfléchir sur une situation clinique identique et d'y travailler par écrit. L'équité était certes assurée, de même qu'une plus grande objectivité (grille de correction identique pour tous les enseignants), mais très vite, on a constaté un taux massif d'échecs et la persistance de nombreuses incohérences dans ce qui était demandé aux étudiants. Cette expérience nous a ainsi prouvé les difficultés de la docimologie (DeLandsheere, 1976).

La remise en question de nombreux enseignants et la recherche d'une plus grande cohérence entre l'évaluation et les objectifs de la formation nous ont alors guidés vers l'ECOS. L'expérience d'un enseignant de notre établissement également chargé de cours à la faculté de médecine de l'Université Catholique de Louvain (UCL) nous rapporte son enthousiasme et des avis positifs quant au vécu d'une ECOS pratiquée dans la discipline médicale. Cela a motivé notre volonté d'expérimenter cette pratique avec nos étudiants de soins infirmiers. D.Giet et V.Massart impliqués dans le projet ECOS de l'Université de Liège (DUMG) sont venus nous partager leurs expériences et conseils. Cette rencontre fut décisive pour notre équipe. Depuis septembre 2009, nous avons introduit l'ECOS dans notre pratique pédagogique en BSI. La formation professionnalisante nécessite d'intégrer une pédagogie active pour former des infirmier(e)s « réflexif(ve)s ». D'autant que le milieu professionnel de la santé attend du milieu de formation que les étudiants finalistes soient suffisamment compétents pour être efficaces et entrer en fonction sans porter de préjudice au patient.

Nous tenons aussi à souligner que l'élaboration du modèle ECOS a été l'occasion de remplacer cette épreuve intégrée par une évaluation qui soit davantage en cohérence avec l'approche par compétences. Elle a également permis de mettre en œuvre une pratique évaluative centrée sur le processus d'apprentissage plutôt que sur les seules performances de l'étudiant. La réflexion à partir de l'ECOS a permis d'examiner en vue d'équilibrer, à travers les contenus définis et les méthodes pédagogiques choisies, d'une part les connaissances à maîtriser, d'autre part les manières de les mobiliser pour développer des habiletés et des compétences transférables. Dans un certain nombre de cours, qu'ils soient théoriques ou pratiques, les enseignants proposent des situations-problèmes, à partir desquelles des notions-clés sont dégagées grâce à la triade contextualisation - décontextualisation - recontextualisation. Les étudiants sont donc confrontés à de véritables problèmes à résoudre et à des opérations mentales qui initient des transferts de connaissances et de compétences. Le besoin de résoudre un problème ou d'accomplir une tâche

² CAPAES : Certificat d'Aptitude Pédagogique Approprié à l'Enseignement Supérieur, titre pédagogique exigé pour enseigner dans les hautes écoles et l'enseignement supérieur de promotion sociale (type court et type long) en Belgique.

devient ainsi le moteur de l'apprentissage des étudiants. Ceux-ci font alors l'expérience que le transfert ne constitue pas seulement la phase finale de l'apprentissage, mais qu'il peut être présent tout au long de celui-ci (Tardif, 1999) et que ce n'est « pas le simple transfert d'une compétence acquise de A vers B, mais la capacité de « se transporter » à travers des situations discontinues et hétérogènes, sans cesse à reconstruire » qui est en jeu (Astolffi in Morisette, 2000).

Les pédagogues sont d'accord pour dire que l'évaluation doit être au service de la formation favorisant la motivation des étudiants à apprendre. L'ECOS doit permettre d'évaluer des compétences, parties de compétences ou ressources visées par la formation, dans des situations ou familles de situations données. L'objectif principal est de vérifier si l'étudiant est capable de transférer les savoirs acquis dans sa formation et de prouver ses compétences dans la clinique quotidienne (pratique).

A la HERS, les enseignants responsables de l'équipe ECOS ont décidé de développer une plus grande expertise en ce qui concerne l'évaluation des apprentissages, ce qui nécessite de se préoccuper à la fois du contenu et du processus des évaluations. Cette étude exploratoire vient asseoir notre choix et favorise le développement de l'approche par compétences dans la mouvance de formation des professionnels de la santé.

3. Conceptualisation

3.1 Qu'est-ce que l'ECOS ?

Le principe de l'ECOS est de faire parcourir aux étudiants une succession de stations d'évaluation dans une durée déterminée ce qui constitue un circuit. On distingue cinq stations pour les premières et deuxièmes BSI et un total de huit stations dont une est dite « repos » pour les étudiants de troisième année. La durée des stations est définie en fonction de la complexité de ce qui est demandé, mais elle est uniformisée : sept minutes ou quatorze minutes. L'étudiant dispose d'une minute pour passer d'une station à l'autre. Un enseignant qui a le rôle de « gestionnaire du temps » est présent dans le couloir, pour gérer le timing et signifier par une sonnerie les changements de station ou la fin d'un circuit.

Pour chaque station, la présence de trois intervenants est indispensable:

- Un étudiant qui prend le rôle du soignant,
- Un co-évaluateur (= acteur) qui joue le rôle du soigné,
- Un évaluateur qui objective l'interaction soignant/soigné à l'aide d'une grille d'évaluation fiable et validée.

L'étudiant reçoit un parcours prédéfini et personnalisé, pour éviter que deux étudiants ne se trouvent en même temps, au même moment, dans la même station. A l'entrée de chaque station, l'étudiant reçoit les instructions précises quant à ce qui est attendu et au temps dont il dispose pour effectuer la tâche ou l'ensemble de tâches demandées (OIIQ, 2010).

La construction des situations est réfléchie par un groupe d'enseignants de la HERS impliqués dans le projet ECOS. Comme le préconise G.Scallon (2007) le choix des situations est important et doit rester centré sur les ressources (savoirs, savoir-faire et savoir-être) mobilisées par l'étudiant à la réalisation de la tâche demandée. Celles-ci sont ensuite testées, avant l'application pour les étudiants en ECOS d'intégration et/ou certificatif, par des pairs ne faisant pas partie de ce groupe de travail. Dès cet instant, un consensus se met en place pour réviser les objectifs de la station et rendre l'évaluation la plus objective possible. Le nombre de station est défini par les compétences attendues par année en lien avec le curriculum de formation.

La validation des stations est la dernière phase de construction et s'effectue avec l'aide de pairs professionnels de terrain. Les cadres de santé des hôpitaux partenaires dans la formation de nos étudiants consacrent une journée pour venir finaliser le travail avec les membres actifs du projet.

L'ECOS est déterminée par les caractéristiques mêmes d'une évaluation. Il faut qu'elle réponde aux exigences d'objectivité, soit reproductible d'un étudiant à l'autre, détermine les apprentissages et doit être réalisée à l'aide d'une grille d'évaluation fiable et validée. Selon, C.Brailovsky (1998) le choix de l'outil mais surtout sa qualité sont responsables de la prise de décision dans l'évaluation.

3.2 Utilisation de l'ECOS pour développer l'aspect formatif dans l'apprentissage

Dés 2009, pour exploiter de manière optimale cet outil et ne pas rester dans une évaluation sommative certificative, nous avons décidé de mettre en place une ECOS d'intégration par année de formation. Les étudiants de première BSI disposent de deux périodes de stage, la première dure deux semaines et la seconde trois semaines, pour un total de 180 heures de pratique clinique sur l'année académique. La stratégie choisie pour faire passer l'ECOS d'intégration est de le réaliser avant chacun des stages. La première épreuve dite « blanche » se déroule en novembre et est centrée sur la préparation de l'approche relationnelle de l'étudiant face à un patient puisqu'il est vierge de tout contact intra/extra hospitalier. Elle est aussi en rapport avec la matière vue aux cours qui reste peu important à cette période de l'année. La seconde épreuve est planifiée en mars centrée cette fois sur le référentiel de formation et permet d'évaluer ce qui est attendu de l'étudiant en fin d'année académique.

ETUDIANT	ENSEIGNANT	R E F L E X I V I T E
L'aider à réaliser un diagnostic de son apprentissage Le motiver à apprendre Le préparer pour son stage L'initier à l'évaluation ⇒ Développer l'autonomie	Identifier l'étudiant en difficulté pour y remédier Réaliser un « état des lieux » : - de sa méthode d'apprentissage - de l'outil d'évaluation	

Figure 1 : Synthèse des objectifs de l'ECOS d'intégration à la HERS.

En ce qui concerne les étudiants de 2^{ème} et 3^{ème} BSI, l'ECOS d'intégration est programmée avant leur dernier stage. L'objectif est de permettre aux étudiants de se situer dans leur apprentissage et ainsi de mettre en place une remédiation éventuelle avant l'épreuve clinique certificative.

Comme le préconise P.Perrenoud (2010), pour aider l'étudiant au développement des compétences, il est important de donner une place prédominante à l'évaluation formative et de la rendre formatrice tout au long de sa formation. Aussi, par cette approche, nous répondons notamment aux critères de l'évaluation formative telle que définie par L.Allal (1979). Elle a pour but de préparer les étudiants aux exigences de l'épreuve certificative, de faire prendre conscience de leur niveau dans la formation et de les préparer à leur stage clinique. Il est également important de leur montrer l'intérêt de ne pas se satisfaire de ce que l'on sait, de ne pas se limiter à la théorie des cours ou aux savoirs procéduraux, de chercher à aller plus loin, de se poser des questions,... et de nourrir ainsi la motivation des étudiants à apprendre (Viau, 2006).

Du côté professoral, que ce soit pour les ECOS d'intégration ou certificatif, tous les évaluateurs reçoivent la station complète (situation clinique, objectifs, consignes pour l'acteur,...) et la grille d'évaluation sous format électronique dans un tableur Excel[®]. La grille d'évaluation est ainsi critériée par des indicateurs (signes observables) pour en limiter la subjectivité. La consigne pour l'évaluateur est d'encoder sous une forme binaire. La personne doit mettre le chiffre 1 dans la colonne correspondante à l'item si l'étudiant répond à l'indicateur et mettre le chiffre 0 s'il n'y répond pas. Il est à préciser que la pondération des items n'est pas binaire et qu'il existe une mesure dite « incontournable³ » se rapportant à la finalité de la station influençant la réussite de celle-ci. L'épreuve terminée, le comptage informatique grâce à des formules dans le tableur Excel[®], permet de visualiser directement le score obtenu par l'étudiant. Tous les résultats sont centralisés par station dans un même fichier afin de les traiter pour les analyser (figure 2).

Le traitement des données s'effectue en passant par différentes étapes en vue d'obtenir un maximum de renseignements. Il est question de faire émerger les résultats globaux par station pour

³ On entend par « incontournable » un critère qualitatif de décision qui se distingue par une pondération plus élevée.

l'ensemble des étudiants puis par classe et de terminer par les indicateurs (items d'évaluation) de chacune des grilles d'évaluation.

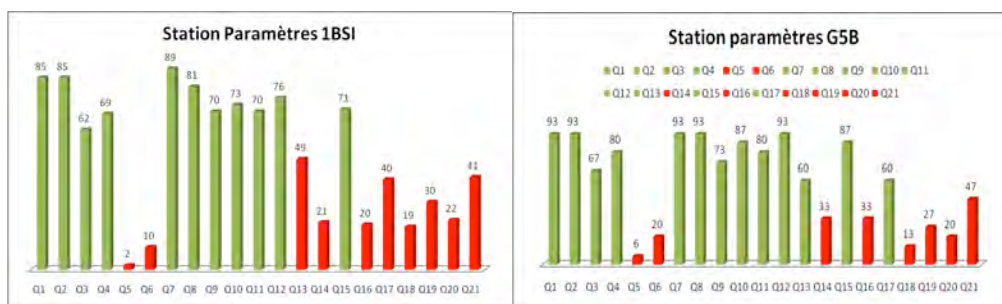


Figure 2 : Illustrations des résultats exprimés en % de réussite qui sont présentés aux enseignants avec une légende en annexe reprenant les différents indicateurs.

Par cet exemple illustré, on peut constater que la classe G5B réussit à 60% l'item correspondant au Q13 alors que l'ensemble des 1^{ères} BSI occupe un taux de réussite à 49% pour le même item. Par contre, si la comparaison se fait sur les items Q5 et Q6 tous les étudiants sont en posture d'échec. On comprend dès lors que l'analyse des résultats peut déboucher sur des réflexions et réactions intéressantes parmi l'ensemble des enseignants. Dans l'exemple cité, l'enseignant du groupe G5B peut partager sa méthode pédagogique à ses pairs qui donnent le cours et ensemble ils vont essayer d'uniformiser leurs pratiques sans pour autant les standardiser ni les stéréotyper.

Les résultats généraux permettent de situer les étudiants pour l'année académique en cours et de faire une comparaison avec les années précédentes. Cette situation diagnostique apporte une fonction régulatrice à l'évaluation.

Les résultats par classe assurent à l'enseignant, d'une part une vision globale du niveau de sa classe mais d'autre part lui permet de pointer les lacunes éventuelles liées à son enseignement. Le taux de réussite par station permet à chaque enseignant de se remettre en question par rapport aux objectifs de cours et aux méthodes pédagogiques utilisées. Ils fournissent également un diagnostic par étudiant, nécessaire à la mise en place d'une remédiation individuelle.

Pour terminer, les résultats obtenus par item d'évaluation pour chacune des grilles assurent une validation de l'outil pour les personnes responsables de la construction des situations d'évaluation.

L'organisation du feed-back pour les étudiants se fait en deux temps, le premier consiste à présenter les résultats généraux par année et pour la classe tandis que le second est individuel. Grâce au premier temps, l'enseignant peut au besoin apporter des compléments d'information à l'ensemble de la classe sur base des conséquences constatées. Si nous reprenons l'exemple de la figure 1, les résultats pour les items Q5 et Q6 sont-ils à l'origine d'une lacune dans la matière théorique ou pratique enseignée au cours ? Le second temps, est consacré à l'étudiant qui analyse ses résultats personnels et entame un consensus avec l'enseignant pour trouver des pistes d'amélioration. Il est invité à se remettre en question et de mettre les suites de ce debriefing en application en stage.

Comme le souligne G.Scallon (2007), l'intégration de l'évaluation dans le processus d'apprentissage est bien plus qu'un simple recensement des acquis de l'étudiant et implique la nécessaire mise en application d'un feed-back constructif. Ce retour sur l'action permet d'entrer dans la démarche réflexive (Gallagher et Saintonge, 2011). Les enseignants invitent les étudiants à prendre conscience des actions passées et de réfléchir à « comment » ils s'y prendraient s'ils devaient intervenir à nouveau sur la situation.

4. Méthodologie

L'étude descriptive exploratoire que nous avons effectuée est basée sur un questionnaire envoyé par voie électronique au corps professoral de la section baccalauréat en soins infirmiers de la HERS. La population visée est composée des enseignants donnant cours de pratique professionnelle qui encadrent les étudiants en stage clinique et qui participent aux ECOS (n=25).

L'enquête a pour finalité de dresser un portrait des modifications des pratiques pédagogiques effectives depuis la mise en place du projet, soit après un an et demi. Elle est constituée de dix questions factuelles : trois sont à énumération graphique demandant au participant d'indiquer son appréciation se rapportant au concept et à l'organisation de l'ECOS ; cinq sont dichotomiques avec la possibilité de se justifier librement par avis constructif visant à mettre en avant les facteurs influençant les pratiques pédagogiques; une question par ordre de rang pour identifier la perception de la valeur ajoutée du projet et enfin la dernière est une liste de pointage pour déterminer l'avis favorable ou non à la pérennité du projet.

Un délai d'une semaine a été donné pour y répondre. Ce laps de temps limité a été respecté par l'ensemble des participants, seuls deux personnes ont rendu leur questionnaire avec deux jours de retard. Aucun rappel n'est formulé et 18 questionnaires peuvent être exploités.

Le recueil des données par questionnaire individuel, sans assistance, a été préféré à une entrevue individuelle pour permettre à chaque participant de prendre le temps nécessaire pour y répondre. Cette méthode laisse la possibilité à chaque intervenant de le réaliser au moment où il le désire sans contrainte de rendez-vous ou d'horaire particulier. Une discussion collective ou groupe de discussion aurait sans doute influencé la fiabilité et la faisabilité de l'enquête. Les réponses obtenues ont été centralisées et traitées dans une base de données par une seule et même personne, dans la plus stricte fidélité et confidentialité, minimisant le risque d'interprétation et d'erreur.

L'analyse des données a principalement été axée sur une interprétation descriptive. Les données complémentaires formulées en justificatif pour certaines questions dichotomiques permettent une analyse qualitative.

Les résultats de cette enquête ont été présentés à l'ensemble des enseignants de la section pour établir un consensus et contribuer à l'évolution du projet.

5. Résultats

Le taux de participation est de 72% (18 réponses/25) et l'ensemble des questionnaires a été exploitable. Les sept personnes qui n'ont pas répondu à la demande ne se sont pas justifiées. Nous pouvons nous satisfaire de ce rapport si nous nous référons aux écrits de MF.Fortin (2006) qui déclare que le taux de réponse est en moyenne de 25 à 30% lorsqu'on utilise la stratégie de l'envoi d'un questionnaire.

Une large majorité (n=15/18) des enseignants se sont dit satisfaits voire très satisfaits du concept et de l'organisation de l'ECOS. Ils ont affirmé que d'un point de vue pédagogique ce concept favorise :

- « - Une actualisation des cours et une évaluation en accord avec les finalités du cursus,
- Une réelle intégration de la théorie avec la pratique,
- Une cohésion d'apprentissage entre les formateurs,
- Un processus d'intégration plus rapide des pairs novices,
- Une amélioration de la créativité dans la construction des cours,
- Un réel partenariat avec les pairs de terrains,
- Une complémentarité de l'évaluation certificative hors contexte hospitalier,
- Une stimulation des apprenants à s'autoévaluer et à aller revoir leurs cours et leur pratique ».

D'un point de vue organisationnel, l'unanimité était présente pour dire que la coordination du projet était efficace, le « timing » respecté, le matériel nécessaire aux mises en situations accessible et acceptable pour rendre celles-ci le plus crédible possible.

La question relative à la perception de la valeur ajoutée a été difficilement exploitable car elle était demandée sous forme d'ordre de rang et la consigne n'a pas été comprise. Nous préférons ne pas la prendre même si la valeur ajoutée reste sous entendue car près d'un tiers des enseignants estiment que c'est « *Sa capacité de jugement nécessaire à résoudre un (des) problème(s) de soins* » qui est testée principalement par ce type d'évaluation, suivie de « *Sa capacité à adapter son attitude face à un patient dans une situation problème* », ensuite « *Sa capacité à adapter son attitude face à un patient dans une situation problème* », puis « *Sa capacité d'habileté face à un*

patient dans une situation problème à résoudre », et enfin « Sa capacité cognitive dans diverses situations problèmes à résoudre ».

A la question se rapportant au choix des mises en situation des stations, une minorité (n=4 /18) des enseignants se sont dit moyennement satisfaits des mises en situation. Ils déplorent la discordance entre le rôle de l'acteur et la réalité des faits tout en admettant que l'authenticité empirique absolue est illusoire. Le reste de la population interrogée s'est dit satisfaite (n=7/18) voire très satisfaite (n=7/18), justifiant leur part d'insatisfaction par la configuration de la station (chambre du patient, matériel...) représentant à leur avis un grand décalage avec la réalité de terrain. Pour eux, le choix des mises en situation est judicieux et les séquences utilisées corroborent la réalité empirique. Cependant, l'ensemble des personnes interrogées disent que le stress généré par l'évaluateur/acteur reste très proche de celui généré par l'enseignant sur le terrain en situation réelle. Deux tiers (n= 12/18) des répondants estiment que la présence de plusieurs ECOS d'intégration est au bénéfice de l'apprentissage des étudiants. Il apparaît que la majorité (n=14/16) des participants ont modifié leur pratique d'enseignement. Deux personnes ne se sont pas prononcées en justifiant qu'elles ne donnent pas cours et leur réponses n'ont pas été prises en compte. Les enseignants qui ont fait évoluer leurs cours mettent l'accent sur :

- Une préférence à la pédagogie active par l'adaptation des stratégies d'apprentissage (exemple : jeux de rôles),
- Une meilleure compréhension par la pratique de « mini ECOS » au sein du cours tout en déplorant le temps que cela prend,
- L'intégration de situations semblables en stage : lien entre théorie et pratique,
- L'utilisation des grilles d'évaluation comme support de cours,
- Les matières mal ou non intégrées sur base des lacunes observées lors des séquences de débriefing,
- L'intégration des étudiants dans l'acquisition de leurs connaissances via des travaux de groupe suivie de séquences de réflexion sur leurs actions,
- L'utilisation et l'intégration des données probantes dans les cours.

Enfin, pour la question relative à la pérennité du projet, seules trois personnes désirent modifier cette évaluation en justifiant qu'il est impératif d'affiner les grilles d'évaluation et d'augmenter la pondération des critères dits « incontournables ». Les autres répondants (n=15/18) expriment que c'est un très bon outil d'évaluation complémentaire à la formation. Les arguments avancés et les points forts sont :

- L'encouragement des étudiants à :
 - Réviser ses cours,
 - Se questionner quant à ses pratiques,
 - «Partager» leur vécu, leur expérience,... avec ses pairs,
 - Se situer dans son apprentissage.
- La possibilité donnée de :
 - Développer une plus grande équité entre les étudiants par rapport à l'évaluation certificative en stage,
 - Assurer plus de cohérence dans la formation,
 - Mettre en œuvre diverses adaptations en lien avec les cours.

6. Discussion

Les résultats obtenus sont satisfaisants et montrent un intérêt majeur de la population interrogée. Il eut été intéressant de questionner les sept personnes qui n'ont pas répondu à la demande. Une entrevue aurait peut-être permis d'apporter des données plus précises sur la procédure du changement et d'insister sur les facteurs influençant les pratiques pédagogiques. L'exploration des caractéristiques des enseignants (âge, expertise dans l'enseignement, détention ou non d'un titre pédagogique tel que le CAPAES,...) constitue un élément qui aurait été déterminant pour apporter davantage de crédit à cette recherche afin de vérifier l'influence de celles-ci sur les modifications apportées dans la pratique des participants.

La finalité de cette enquête était de dresser un portrait des modifications des pratiques pédagogiques et secondairement, s'assurer de la pérennité du projet. Les résultats montrent une diversité de l'utilisation de l'outil d'évaluation tant comme ressource pour construire un cours ou

le mettre en application pour rendre la pédagogie plus active que comme base de sa propre évaluation de l'enseignement.

Grâce à l'ECOS d'intégration, la majorité des répondants pensent que les étudiants sont suffisamment préparés pour l'épreuve certificative et que tout est mis en place pour limiter le stress. Comme l'exprimait P.Perrenoud (2010) lors d'une formation reçue au sein de notre établissement : « il n'y a pas d'évaluation sans stress... ». Certains enseignants craignent que le temps passé à l'école pour réaliser les épreuves d'intégration n'empiète sur le temps qui pourrait être valorisé en stage. A nos yeux, l'intérêt de l'ECOS est justement d'offrir à l'étudiant la possibilité de se tromper sans pour autant risquer de mettre en danger un patient. Sur base de ses erreurs, à la condition qu'il bénéficie d'un débriefing constructif, l'étudiant peut se remettre en question et transférer ses apprentissages lorsqu'il sera confronté à une situation semblable sur le terrain. Il est d'autant plus pertinent et sécuritaire d'utiliser la forme curriculaire que la mise en situation réelle empirique est trop souvent interrompue par des imprévus liés au patient.

7. Conclusion

Cette étude exploratoire visant à évaluer un type nouveau d'évaluation a eu comme bénéfice premier de dynamiser une équipe d'enseignants. En effet, elle a permis de dégager un consensus sur la méthode d'évaluation et de stimuler la mise en œuvre de pratiques de pratique pour améliorer notre cohérence pédagogique. Il est intéressant de constater que plus des trois quart des répondants ont choisi de modifier leur méthode d'apprentissage depuis la mise en place du projet. La conviction pour ceux-ci de répondre aux exigences de l'approche par compétences en est-elle le moteur ? Cette question mériterait d'être approfondie. Du côté des étudiants, on a constaté une certaine tendance à la résistance au changement, principalement basée sur la peur de l'échec, même s'ils ont très vite compris l'intérêt de cette évaluation. Sur base de leurs commentaires oraux on peut dire aujourd'hui que les étudiants sont devenus favorables à l'ECOS. Nous lions cet effet aux débriefings réalisés après chaque ECOS d'intégration mais surtout aux modifications perçues dans les différents cours concernés. Ceci pourrait faire le cadre d'une prochaine étude.

Contributions et remerciements

Cette étude n'aurait pu être sans la participation de tous les enseignants de la HERS impliqués de près ou de loin dans le projet ECOS. Je tiens à leur exprimer toute ma gratitude et à exprimer des remerciements particuliers à C.Bolly, L.Canautte, C.Lebas, A.Pirson, J.Schelfhaut et E.Tison, qui se sont engagés dans la rédaction et la relecture de cet article.

Bibliographie

- Allal, L. (1979). *Stratégies d'évaluation formative : conceptions psychopédagogiques et modalités d'applications*. In Allal, L. Cardinet, J. & Perrenoud, P. *L'évaluation formative dans l'enseignement différencié*. Bern : Peter Lang.
- Astolfi, J.P. (1994). in Morissette, R., (2002). *Accompagner la construction des savoirs*, Montréal : Chenelière Education.
- DeLandsheere, G. (1976). *Evaluation continue et examens précis de docimologie*, Bruxelles.
- Fortin, M.F., Côté, J. et Filion, F. (2006) ; *Fondements et étapes du processus de recherche*, Québec, Canada : Chenelière éducation.
- Harden, R., Stevenson, M., Downie, W., Wilson, G. (1975) *Assessment of clinical competence using objective structured examination*. BMJ; 1:447-451.
- Leclercq, D. (2006-2007). *Méthodes d'action pédagogique et d'évaluation en grands groupes*, IFRES des Formasup
- OIIQ : <http://www.oiiq.org/infirmieres/etudiants/examen.asp>.
- Ordre des infirmières et infirmiers du Québec, (2010). *Guide de préparation à l'Examen professionnel de l'OIIQ*, Canada,
- Perrenoud, P. (2010). *Ateliers pédagogiques*, Documents non publiés. Haute Ecole Robert Schuman Belgique :
- Scallon, G. (2007). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles : de boeck.
- Tardif, J. (1999) *Le transfert des apprentissages*. Montréal : Les éditions logiques (p 34-35).

Gallagher,F. et Saintonge, L. (2011). *Formation interne sur la réflexivité*. Documents non publiés. Haute Ecole Robert Schuman Belgique :

Viau, R. (2006)., *La motivation des étudiants à l'université, mieux comprendre pour mieux agir*. Université de Sherbrooke, Canada

Cadre Légal belge. Décret du 17 juillet 2002 instituant le CAPAES, modifié par le décret du 2 juin 2006.

UNE EVALUATION « PAR LA BANDE » DES COMPETENCES EN MILIEU PROFESSIONNEL. FORMATEURS D'ADULTES ET CHARGES D'INSERTION SOUS EVALUATION

Cédric Frétné

Université Paris Est Créteil Val de Marne, Laboratoire REV-CIRCEFT
cedric.fretigne@u-pec.fr

Mots-clés : évaluation des compétences, formation professionnelle d'adultes, insertion professionnelle

Résumé. Le texte met l'accent sur la manière dont les évaluations de dispositifs d'insertion sociale et professionnelle contribuent incidemment à la production d'une mesure et à l'énoncé d'un jugement sur la qualité du travail des chargés d'insertion. Aussi indirectes soient-elles (ce n'est pas leur vocation première), les appréciations sur les résultats obtenus (les taux d'insertion de « bénéficiaires » accompagnés dans leur « parcours d'insertion ») sont clairement référées à l'activité ordinaire des chargés d'insertion, à leur professionnalisme (à la mesure de leur investissement professionnel et au jugement sur leurs compétences). L'analyse des modalités d'évaluation de cinq dispositifs d'insertion sociale et professionnelle en France permet de repérer dans quelle mesure l'évaluation des dispositifs contribue également à évaluer la qualité de la prestation fournie par les chargés d'insertion et, in fine, leur travail lui-même.

Introduction

La question de l'évaluation connaît, en sciences humaines et sociales, un regain d'intérêt certain. Dernièrement, des numéros spéciaux de revue ont entrepris de faire état de la « passion évaluative » (Nouvelle revue de psychosociologie, 2010) qui gagne différents espaces sociaux ou d'interroger ce qu'évaluer signifie, en termes de normes de référence, valeurs véhiculées et pratiques effectives (Cahiers internationaux de sociologie, 2010). De cet ensemble de travaux et réflexions ressort une interrogation forte visant à distinguer le plus précisément possible deux acceptions possibles de l'évaluation : une *mesure* se voulant énoncée en extériorité, de manière neutre et objective, visant à fixer la contribution effective à la production de richesses ou au rendu d'un service ; une *valeur*, référant à des normes comportementales attendues, pointant l'état d'esprit général, l'effort consenti et appréciant l'adhésion au projet (qu'il soit d'entreprise, associatif ou dans le cadre d'un service public).

Dans bien des cas toutefois, ce n'est pas le travail lui-même qui fait l'objet de l'évaluation, mais bien son résultat. Par une opération qui relève de la subsomption, de la conformité des résultats aux objectifs fixés *ex ante* est inférée la qualité du travail proprement dit. Inversement, de résultats jugés insatisfaisants naît un doute sur la (piètre) qualité du travail et une interrogation sur les compétences et le professionnalisme de la personne considérée.

Partant de ce cadre d'analyse général, nous souhaitons mettre ici l'accent sur la manière dont les évaluations de dispositifs d'insertion sociale et professionnelle contribuent incidemment à la production d'une *mesure* et à l'énoncé d'un *jugement* sur la qualité du travail des chargés d'insertion. Aussi *indirectes* soient-elles (ce n'est pas leur vocation première), les appréciations sur les résultats obtenus (les taux d'insertion de « bénéficiaires » accompagnés dans leur « parcours

d'insertion») sont clairement référées à l'activité ordinaire des chargés d'insertion, à leur professionnalisme (à la *mesure* de leur investissement professionnel et au *jugement* sur leurs compétences).

L'analyse des modalités d'évaluation de cinq dispositifs d'insertion sociale et professionnelle¹ en France constitue le matériau d'enquête qui soutient cette communication. Il s'agit plus précisément de repérer dans quelle mesure l'évaluation des dispositifs, telle qu'annoncée dans les cahiers des charges et au moment de sa mise en œuvre, contribue également (et dans quelle mesure) à évaluer la qualité de la prestation fournie par les chargés d'insertion et, *in fine*, leur travail lui-même.

La « philosophie de l'évaluation en débats »

L'évaluation est une activité sociale qui pénètre de plus en plus fortement au cœur des organisations, quelles que soient leur raison sociale et les finalités qu'elles poursuivent. L'évaluation des politiques publiques tient toutefois une place singulière dans cet ensemble. Élément catalyseur de la diffusion de cette forme d'activité sociale, le champ de l'action publique est le lieu de toutes les expérimentations, de tous les engendremens méthodologiques. Dans le but d'en proposer une critique raisonnée, Danilo Martuccelli (2010) s'est dernièrement employé à distinguer les formes élémentaires de cette philosophie de l'évaluation. Sans revenir sur l'ensemble de ses développements, on peut souligner avec le chercheur les principes qui structurent et légitiment le recours à l'évaluation, évaluation la plus fréquemment fondée sur des indicateurs quantitatifs de performance.

« Tout est mesurable » constitue la pierre angulaire de cette philosophie. L'irréductible hétérogénéité des univers de sens est battue en brèche par le postulat d'une possible (et nécessaire) comparaison des éléments signifiants. Sur cette base, des règles d'équivalence sont forgées, des constructions par analogie établies, de la consubstantialité trouvée.

« Évaluer garantit le respect d'un principe de justice sociale ». Sur le terrain des grands principes démocratiques, ensuite, l'évaluation est présumée dire en toute objectivité *ce que chacun vaut réellement*. Échappant ainsi à toute appréciation biaisée au regard des pratiques de népotisme et plus généralement de favoritisme, l'évaluation s'offrirait comme le point d'appui méthodologique de la méritocratie.

« Évaluer génère une juste répartition des pouvoirs » constitue le troisième atout reconnu à cette pratique sociale. Rendant transparent ce qui restait opaque, *technicisant* à des fins d'objectivation ce qui semblait subjectif et orienté, l'évaluation permet de dégager une élite (des individus, des organisations) crédible, compétente, à laquelle il paraît légitime de confier les rênes du pouvoir.

« Évaluer est gage d'efficience dans les organisations ». Fondée sur un principe d'allocation rationnelle des ressources, l'évaluation fait gagner en efficience, rend optimal le rapport entre les *inputs* engagés et les résultats observés. Économie d'énergie, gain de temps, restriction des coûts financiers, l'évaluation est fondatrice d'un modèle de développement qui fait la chasse aux gaspillages, aux temps morts de l'activité, aux « grosseurs » dans les organisations.

« Évaluer est source d'efficacité ». Efficientes certes, mais tout bonnement efficaces, les activités conduites suivant les principes de l'évaluation gagnent en pertinence et rendent les services voulus. La pratique de la comparaison systématique, dite du *benchmarking*, conduit à l'importation d'idées nouvelles, à l'abandon de routines jugées improductives et à l'intronisation de préceptes managériaux à prétention universelle.

¹ Un dispositif d'insertion sociale par le logement, un dispositif d'insertion pour les bénéficiaires du Revenu de solidarité active (RSA), un dispositif d'insertion pour des jeunes suivis par l'Aide sociale à l'enfance (ASE) et/ou la Protection judiciaire de la jeunesse (PJJ), un dispositif dédié à la mobilisation professionnelle pour des jeunes chômeurs, un dispositif pour l'insertion par l'emploi de chômeurs de longue durée adultes.

« Il y a une dimension performative de l'évaluation ». Qu'advient-il des pratiques de tous les jours quand un pilotage par l'aval devient manifeste, lorsque l'on sait que l'activité présente sera, un jour prochain, mesurée à l'aune d'indicateurs quantitatifs ? L'idée promue est ainsi qu'en prévision de la *prochaine* évaluation, l'activité des individus et des organisations soit de manière croissante organisée par référence aux attentes qui seront formulées par l'évaluateur. C'est pourquoi le rapprochement temporel des évaluations est également pratiqué afin de transformer le rapport à l'activité ordinaire et de maintenir salariés et/ou organisations sous pression.

« Évaluer accroît la légitimité des organisations ». L'évaluation est présumée affermir la légitimité des dirigeants et contribuer à la production d'une société pacifiée. Technicité, scientificité et objectivité des procédures, outils et indicateurs sont tenues pour les meilleurs gages de l'efficacité, de l'efficience et de la transparence des prises de décision. La légitimité des élites et de leurs actions gagnerait ainsi à être fondée en toute rationalité, par la commande initiale d'évaluations en tout genre et sur tout sujet. *Last but not least*, de ces processus de rationalisation pourrait naître une société (enfin) rationnelle ou plus rationnelle, coupant court avec les romantismes surannés, le bricolage institutionnel, des décisions politiques irrationnelles, des stratégies économiques construites en aveugle, etc.

La philosophie de l'évaluation a donc de solides arguments à faire valoir. On pourrait, avec Martuccelli, les décortiquer voire les démonter un par un. Il suffit toutefois, pour notre propos, d'en avoir évoqué les principes. Ce court développement rend explicite la difficulté, pour les acteurs sociaux, de s'en affranchir. S'ils subodorent bien que, sous ses beaux atours, l'évaluation remplit d'autres fonctions que celles explicitement affichées, ils se sentent assez fréquemment écrasés sous le poids des grands principes qui en régissent le fonctionnement officiel. On verra toutefois, dans la suite de ce texte, les insatisfactions que suscite sa mise en œuvre effective.

Des professionnels sous évaluation

On sait, en sociologie du travail, la difficulté à apprécier la contribution productive propre à chacun. De longue date maintenant, la discipline a mis l'accent sur la nécessaire prise en compte des formes de coordination et de coopération qui rendent très largement inappropriées les évaluations individuelles des salariés. Ce qui vaut pour l'emploi industriel vaut plus encore dans le cadre d'activités de service, fondées sur un principe de « collaboration » interactionnelle entre le vendeur et le client, le fonctionnaire et l'utilisateur, etc. Concernant spécifiquement les métiers relationnels, Lise Demailly (2008, p. 19) a souligné qu'ils « articulent étroitement une expertise, un positionnement éthique et des pratiques d'interaction ».

Comment évaluer, *en propre*, l'activité de travail de ces professionnels de la relation ? Pour résoudre les problèmes liés à la complexité inextricable des compétences mobilisées, la solution adoptée consiste de fait à recourir à une évaluation indirecte, fondée sur une logique de dérivation. Évaluant l'efficacité des résultats d'un dispositif public, les évaluateurs en inféreront l'efficacité de la réponse apportée par le mandataire. Et pour rendre compte de l'efficacité à géométrie variable des opérateurs financés sur fonds publics, les évaluateurs procéderont à une deuxième inférence les conduisant à se prononcer sur l'efficacité du travail des formateurs et des chargés d'insertion eux-mêmes. Le saut qualitatif opéré conduit les professionnels concernés à mentionner qu'ils ne se « retrouvent » ni dans les procédures déployées, ni dans les résultats annoncés. Il leur semble, ainsi que leur activité effective n'est pas prise en compte, mieux, qu'elle fait office de point aveugle de la démonstration. Dès lors, on comprend que puisse naître des regrets voire de l'amertume à l'endroit de démarches d'évaluation qui rendent si infidèlement compte de l'exercice professionnel dans ses attributs les plus ordinaires. Dans ce cas de figure, l'évaluation est rarement vue comme formative mais présente tous les atours de la sanction. L'évaluation est redoutée, par avance délégitimée, si possible tenue à distance. Et l'on en arrive parfois à la situation que l'évaluation visait précisément à éviter : le repli sur le quant à soi et l'apathie comme réponses à l'injonction évaluatrice.

Le propos ne consiste pas ici à prendre fait et cause pour tel ou tel segment de la chaîne de l'évaluation. Il s'agit très clairement de marquer une rupture avec les logiques « vues d'en haut » insistant sur l'incurable « résistance au changement » d'acteurs de premier niveau pris dans leurs routines et réticents à l'idée même de modifier les cadres de leur action. Mais il s'agit également d'échapper aux logiques « vues d'en bas » dénonçant des démarches d'évaluation découplées des « réalités du terrain », aveugles et sourdes aux spécificités des activités d'accompagnement d'usagers/bénéficiaires/allocataires. L'idée qui prévaut est à la fois plus simple dans son principe et plus complexe dans sa mise en œuvre. Il s'agit de repérer *ce qu'évaluer veut dire*, ce qui signifie, entre autres, de s'intéresser aux questions suivantes : qu'est-ce que l'évaluation engage en termes de connaissances produites ? Que donne-t-elle à voir de l'activité effective des opérateurs ? Quels usages sociaux en font les différents protagonistes ? Sur quels points les controverses se nouent-elles et sur quels terrains les accords s'établissent-ils ?

Les évaluations des différents dispositifs

Un dispositif visant l'« accompagnement social » de bénéficiaires du RSA

Un premier dispositif porte sur l'accompagnement social de bénéficiaires du Revenu de solidarité active (RSA) impulsé par un Conseil général. Réglementé par le code des marchés publics, l'appel à projets stipule les modalités d'évaluation attendues... des « évolutions » du bénéficiaire durant la période de suivi et non celles de la structure ayant emporté le marché et encore moins celles du professionnel assurant la prestation. Pour faire, des outils sont imposés au prestataire avec obligation de les renseigner périodiquement. Un outil de suivi mensuel voisine ainsi avec un support offrant un récapitulatif des interventions sur l'année civile.

Bien qu'on ne trouve pas trace, dans l'appel à projets, d'éléments précis en matière d'évaluation, des indicateurs implicites apparaissent lorsque sont mentionnées les obligations contractuelles du soumissionnaire. Il s'agit en premier lieu de s'assurer que tous les bénéficiaires de l'action disposent bien de l'ensemble des droits sociaux auxquels ils peuvent légitimement prétendre. Il s'agit ensuite, et dans chaque cas, d'élaborer un projet personnel et/ou professionnel prenant en compte les capacités existantes de l'allocataire. Enfin, ce projet ne saurait être standardisé, dupliqué à l'identique d'une personne à l'autre. Il doit impérativement être personnalisé, de même que l'accompagnement à sa réalisation. Néanmoins, par méfiance envers la « religion » du projet personnalisé, il est bien rappelé que cette démarche ne constitue pas une fin en soi mais n'est qu'un moyen en vue de permettre à chaque allocataire d'« atteindre des objectifs réalistes d'amélioration de sa situation personnelle ou professionnelle ». Ces exigences trouvent leur traduction dans des écrits (une fiche individuelle récapitulative des éléments du parcours) et la communication de données chiffrées (« orienter au moins 50 % des bénéficiaires suivis sur des actions visant la construction ou la réalisation d'un parcours professionnel adapté ou vers l'emploi »). C'est donc une synthèse des parcours, allocataire par allocataire, qui est demandée. Les données individuelles sont, quant à elles, agrégées sur une base informatique, permettant au Conseil général de s'assurer qu'au moins 50 % des « mesures » ont donné lieu, au terme du parcours, à une orientation à finalité professionnelle.

Dans la pratique administrative, l'évaluation se réduit assez fréquemment au « contrôle de service fait ». Dans le cas présent, trois ensembles de productions doivent être réalisées. En premier lieu, une application informatique nommée EPI (« Évaluation des parcours d'insertion ») doit être alimentée pour la production de données statistiques par agrégation d'informations individuelles. Des indicateurs d'activité pourront ainsi être produits, à l'échelon départemental, par territoire, par profil d'allocataire et, surtout, par organisme financé. L'établissement d'indicateurs de résultats attendus est ensuite rendu possible : nombre de personnes sorties avec un projet parmi l'ensemble, avec une ventilation possible entre projet social et familial d'une part, projet professionnel de l'autre. Le rédacteur du cahier des charges prend alors garde d'ajouter que « ces indicateurs pourront être complétés, le cas échéant, par tout autre indicateur jugé utile par le Département et le

porteur de projet ». « Rapports intermédiaires et finaux d'évaluation pédagogique » forment une deuxième série d'éléments qui doivent être communiqués. Chacun fait état de l'ensemble des démarches initiées avec l'allocataire et argumente les évolutions et autres adaptations opérées durant l'accompagnement du parcours. Des annexes techniques et financières dûment renseignées sont enfin censées permettre un contrôle de la bonne exécution de la convention.

Dans les faits, c'est une logique plutôt énumérative qui préside à la synthèse opérée. Sous EPI mais également dans les « synthèses qualitatives », une succession d'informations est livrée sans réel « liant » entre elles. La juxtaposition l'emporte sur la coordination, donnant le sentiment d'un tronçonnage contraire à tout ce que la rhétorique du parcours et de sa continuité donne à entendre. À raisonner par ailleurs suivant un schéma unilinéaire pour ne pas dire évolutionniste, on oublie (ou ignore) les périodes stationnaires et les moments de « régressions » dans le parcours d'insertion. Ce que les professionnels de l'insertion énoncent tous à longueur d'entretiens semble proprement inaudible et irrecevable dans le cadre des pratiques d'évaluation mises en œuvre par les financeurs. Une grande partie du *réel du travail*, rendu invisible dans le cadre de l'évaluation, consiste à travailler au maintien à flot des bénéficiaires durant les périodes de transition où le doute le dispute à l'angoisse face à un avenir incertain. Et ce travail là n'est semble-il jamais (ou presque) reconnu, au double sens du terme. Non reconnu au sens de non répertorié. Mais non reconnu, au sens également de non valorisé. C'est dire que l'évaluation qualitative fait ici figure de parent pauvre. Les dispositifs et procédures mis en œuvre par le Conseil général pour apprécier, dans le cadre de l'action considérée la qualité et l'efficacité de l'intervention sont au fond assez peu soucieux de promouvoir des éléments qualitatifs, coordonnés et présentés en contexte.

Un dispositif départemental liant éducation, formation et insertion

Co-piloté par un Conseil général et une Protection judiciaire et la jeunesse, l'ambition du dispositif soutenu dans le cadre du fonds d'expérimentation sociale pour la jeunesse est d'opposer une force contraire à l'ensemble des freins qui nuisent le plus généralement aux démarches d'insertion de jeunes dits « en difficultés ». Sont ainsi rappelés que les « effets de seuil » ou les « effets de frontière » limitent leur accès à des formations en raison de leur âge, de leur statut, de leur lieu d'hébergement, du type de financement obtenu. De même, des « effets d'éviction » (pour ne pas dire des formes de discrimination) interdisent parfois que leurs droits à l'insertion soient rendus opérants. Enfin, des « effets d'émiettement » de l'offre rendent délicat l'accompagnement de ces jeunes dans leurs démarches d'insertion. L'ambition ici poursuivie vise à « faire accéder les jeunes à l'ensemble des dispositifs existants et à réduire au minimum les sorties "sèches" sans perspectives réelles d'autonomie de vie ». Pour ce faire, il s'agit d'abord d'améliorer les actions existantes : mieux organiser et fédérer l'ensemble des acteurs du territoire concernés par ce sujet ; rendre plus accessible l'ensemble des ressources et éléments d'informations disponibles, pour les jeunes et leurs familles. Plus spécifiquement, l'objectif est d'améliorer la réussite scolaire et de favoriser l'insertion professionnelle de jeunes fortement marqués par l'échec scolaire et par les difficultés d'insertion professionnelle.

Conformément aux attendus de l'appel à projets, un organisme indépendant assure l'évaluation du dispositif. L'objectif affirmé de l'évaluation est d'« identifier ce qui favorise et fait obstacle à l'entrée dans le dispositif départemental éducatif et de formation et d'insertion en direction des adolescents et des jeunes en situation de déscolarisation ». Elle entend procéder à l'analyse des parcours des jeunes depuis leur entrée dans le dispositif, « en accordant un intérêt particulier à la phase de définition d'un projet, mais aussi aux abandons ou aux modifications du projet ». Là encore, rien n'est dit d'une éventuelle évaluation de l'action des professionnels engagés dans le dispositif. Il reste que la réponse à l'appel d'offre précise qu'un outil d'évaluation de la situation individuelle des jeunes sera élaboré en concertation avec les professionnels concernés par leur accompagnement et en accord avec le comité de pilotage chargé de suivre l'expérimentation, afin notamment de responsabiliser l'ensemble des participants. L'outil imaginé, un questionnaire, comprend d'abord l'exposé de la situation actuelle en termes de formation, d'emploi ou d'hébergement. Les évolutions enregistrées entre les temps t et t+1 ont vocation à repérer les

situations qui se sont améliorées et celles qui se sont dégradées. Au regard des éléments de réponses, l'évaluateur se propose ensuite d'identifier les besoins d'accompagnement des jeunes.

Au final, trois dimensions méritent ici être prises en considération. En premier lieu, l'outil est-il renseigné ? En second lieu, qui le renseigne et que comprend-il effectivement ? Est-il donc un outil partenarial comme initialement envisagé et procure-t-il des informations exploitables ? Enfin, se montre-t-il ajusté aux objectifs écrits dans la réponse à l'appel à projets et permet-il une évaluation fiable des résultats ? De fait, moins de la moitié des questionnaires ont été retournés (60 sur 140). Le traitement, à la fois qualitatif et quantitatif (par recodage et catégorisation des réponses à certaines questions ouvertes), conduit à interroger l'adhésion des professionnels à la démarche. De fait également, des divergences de point de vue sont bien apparues entre les différentes personnes (le jeune, le référent, les responsables légaux et les autres professionnels), notamment au temps t. L'analyse des écarts a été considérée comme un point *a priori* sensible de l'évaluation mais promoteur. Au regard des finalités pratiques investies dans ce dispositif, elles sont présumées permettre de relever et de lever des incompréhensions mutuelles, de dégager des pistes d'intervention profitables à l'insertion sociale et professionnelle des jeunes.

Au final, l'évaluateur s'interroge lui-même : la multiplicité des indicateurs qualitatifs et la perspective chronologique ne rendront-elles pas l'évaluation très largement inextricable par défaut d'instrumentation ? En effet, la comparaison terme à terme des deux questionnaires pour chacun des jeunes est parfaitement possible. On y verra transparaître, quand les questionnaires sont dûment renseignés, les évolutions enregistrées. Mais comment opérer sur des données agrégées correspondant aux 60 questionnaires ? Aucun codage informatique n'ayant été initialement prévu, le dépouillement et l'exploitation des données s'opéreront à coup sûr de manière particulièrement empirique. Quel crédit les acteurs du dispositif eux-mêmes donneront-ils à cette opération d'évaluation ? Comment en particulier concevoir le partage d'informations entre un professionnel reconnu, disposant d'un code d'éthique, membre d'une association professionnelle et soumis à des règles d'activité codifiées (assistant de service social, éducateur spécialisé, etc.) et des « intervenants sociaux » aux titres et fonctions changeants, aux compétences malaisément décelables et à la légitimité réduite ? Si la « médiation familiale » fait l'objet, depuis, 2004, d'une reconnaissance institutionnelle par la délivrance d'un diplôme d'État de niveau II (Divay, 2009), il n'existe rien d'équivalent pour les « nouveaux métiers » de l'intervention sociale dans le champ de l'insertion sociale et professionnelle. Et si le repli sur le « colloque singulier » et l'argument juridique du devoir de confidentialité constituent autant de pare-feux avancés par les professions réglementées du social pour couper court à tout réel travail partenarial, c'est assurément aussi pour échapper, via l'évaluation du dispositif, à l'évaluation de la qualité de leur travail.

Un dispositif d'insertion dans le cadre d'un Foyer de Jeunes travailleurs (FJT)

Un certain nombre de bénéficiaires d'un Contrat d'insertion dans la vie sociale (CIVIS) disposent de places en Foyer de jeunes travailleurs (FJT) financées par le Conseil général et l'État pour des jeunes inscrits et suivis en Missions locales. Dans le cas présent, la durée de leur hébergement en FJT est fixée à 9 mois maximum. Il apparaît que leur situation personnelle au regard de l'emploi et du logement est globalement marquée par une forme d'instabilité sinon de précarité. D'un point de vue qualitatif, cela se traduit par des éléments de diagnostic pointant assez invariablement une série de difficultés. Les professionnels mentionnent dans leurs « rapports sociaux » des *problèmes* de langue, d'insertion professionnelle. *Sans* soutien familial pour nombre d'entre eux, ils sont caractérisés comme *manquant* de réseaux sociaux, de revenus (stables), de ressources en termes de formation et de compétences à faire valoir.

Un grand constat d'impuissance accompagne ces comptes rendus. De fait, leur situation au regard de l'emploi ne s'est pas améliorée au terme des 9 mois de parcours. Quant à la situation en matière de logement, les chiffres sont, de l'avis même des professionnels accablants. Suivant la typologie *ad hoc* du FJT, parmi les 14 sortis en 2009, 4 sont sans logement, 3 en Centre d'hébergement et de réadaptation sociale (CHRS), 1 en FJT, 4 chez des tiers et 2 seulement ont eu accès à un logement

autonome. Avec cette conclusion, il ne s'agit pas d'incriminer l'association porteuse du FJT ni de mettre en question les compétences des intervenants professionnels qui exercent en son sein. Au demeurant, jusqu'à présent, les financeurs de l'action CIVIS s'en tenaient, en matière d'évaluation, à des données factuelles relatives au nombre de jeunes hébergés dans l'année et à la durée de leur séjour. Il s'agit donc moins d'une évaluation que d'un contrôle *ex post* du respect de la convention en termes numériques. Les taux d'insertion (quels que soient les outils de mesure et les indicateurs retenus) sociale et professionnelle ne sont donc pas encore mis en avant pour juger de la pertinence de renouveler la convention. Il apparaît toutefois que le nouveau cahier des charges auquel l'association devra répondre comporte désormais, et pour les années futures, des indications concernant le « résultat obtenu par le dispositif ». Une attente de résultats quantitatifs en termes de devenir professionnel figure en première ligne des nouvelles exigences, ce qui ne laisse pas d'inquiéter les professionnels du FJT.

Pour les équipes, les projets des résidents sont en effet très imprécis en début de séjour et rarement finalisés neuf mois plus tard. Des « sorties sans solution » sont assez généralement observées et les intervenants craignent que la possible centration de l'évaluation sur ces seules données quantitatives rende peu justice à leur action au quotidien. Par ailleurs, professionnels de l'action sociale ou de l'animation au sens large du terme, ils se reconnaissent mal outillés pour assurer ce travail d'accompagnement vers l'emploi. Et, telle une litanie, revient sans cesse la question de la durée du contrat de séjour. Comment, dans des délais si brefs et au regard des « difficultés » cumulées des jeunes escompter une sortie positive en quelques mois ? Les intervenants insistent sur la nécessité exprimée par les jeunes eux-mêmes de « se poser » après des années chaotiques où ils ont connu des instabilités voire des ruptures familiales accompagnées de violences, des résidences successives et temporaires, des scolarités heurtées marquées par de nombreux changements d'établissement et de filières, des emplois précaires sous statut varié (CDD, intérim, emploi non déclaré).

On peut toutefois se demander si le tournant annoncé de l'évaluation quantitative ne sera pas particulièrement dévastateur des pratiques professionnelles jusque là soigneusement protégées des regards inquisiteurs, sans que les professionnels y gagnent de reconnaissance en retour. Si le financeur constate, année après année, que l'insertion professionnelle n'est pas au rendez-vous (et que la trajectoire résidentielle ne conduit que très marginalement au logement autonome), quelles conclusions en tirera-t-il ? Le risque est grand qu'il incrimine la structure taxée d'incompétence et ses professionnels jugés peu efficaces. Alors même qu'une évaluation qualitative de l'insertion sociale des résidents saisie au prisme de leurs « compétences à habiter » permettrait de créditer le FJT et ses professionnels d'une véritable habileté dans le champ de compétences qui est le leur (l'insertion par le logement), l'évaluation quantitative du devenir social et professionnel *ex post* fait totalement abstraction de l'action au quotidien des intervenants. Si les termes du nouveau cahier des charges se confirment, tout sera donc fait comme si les jeunes n'avaient aucune propriété sociale dirimante à l'entrée au FJT. D'inemployables selon une formule à la mode, il sera attendu que des résidents peu diplômés, peu expérimentés, chômeurs, dépourvus de réseaux sociaux *qui comptent* se muent en quelques mois en jeunes munis d'un projet d'insertion réaliste, employables sinon en emploi, disposant par surcroît d'un logement autonome. Et il sera attendu que les professionnels du FJT se muent en agents d'insertion... dans l'emploi.

Un dispositif d'« accompagnement individualisé vers l'emploi »

Dans le cadre d'une action financée par un Plan d'insertion dans l'Emploi (PLIE), un contrôle des « services faits » tient lieu semble-t-il d'évaluation. À cette aune, l'opérateur retenu donne pleinement satisfaction... puisqu'il accueille plus de bénéficiaires que le nombre initialement prévu et que le référent de parcours réalise plus d'entretiens que ceux prévus de manière forfaitaire dans la convention de partenariat. Restent toutefois en suspens plusieurs questions sur lesquelles nous nous arrêterons. Si, dans le cas présent, la reddition de compte et le contrôle des services faits valent semble-t-il évaluation, quelle place la *mesure* laisse-t-elle à l'appréciation de la *valeur* de l'activité d'accompagnement ? Dans la continuité, on peut également se demander quel sort est-il

fait, dans l'évaluation-mesure, au travail d'accompagnement proprement dit et à la valeur qui est le sien en matière d'accompagnement ?

Le cahier des clauses administratives particulières précise l'attente à l'endroit des répondants au marché public. L'opérateur « référent de parcours » occupe normalement une fonction de conseiller en insertion professionnelle. Son activité vise à établir, avec les bénéficiaires, un parcours d'insertion « individualisé », « réaliste », « progressif » et « adapté ». Dans les faits, le référent de parcours rencontré souligne que son travail est à la fois contraint par des orientations quantitatives en provenance des financeurs de l'action et par la nécessaire prise en compte, singulière, de chaque bénéficiaire. Entre traitement de masse (800 entretiens par an sont attendus à raison d'une moyenne de 10 par bénéficiaire) et individualisation des réponses (signature d'un contrat d'objectif adapté à chacun), le professionnel organise ainsi son activité ordinaire et ajuste ses pratiques de telle sorte qu'elles satisfassent cette double injonction. En résumé, il compose entre une routinisation des pratiques qui lui permet de gagner en efficacité et de « faire du nombre » et une écoute attentive qui l'autorise à adapter l'accompagnement aux besoins exprimés par chacun.

Ni linéaires ni mécaniques, les processus qui conduisent la personne à entrer dans une démarche d'auto-direction sont précisément détaillés rencontre après rencontre. La grande majorité des informations couchées sur papier échappent aux exigences formulées par le PLIE. Et pourtant, elles sont, pour le professionnel, les éléments *effectivement* constitutifs du parcours *réel* d'insertion sociale et professionnelle. Qu'en est-il des relations entretenues avec les partenaires également en charge de participer à l'accompagnement ? Où en sont les personnes concernées par des problèmes financiers dans leurs démarches avec les assistants de service social ? Des places en crèche ont-elles été trouvées pour les enfants en bas âge à moins que des assistantes maternelles n'aient été employées ? Les droits relatifs à la prise en charge des frais de garde par la CAF ont-ils été ouverts ? Les associations ayant compétence pour soutenir des étrangers dans l'obtention ou le renouvellement de leur titre de séjour ont-elles été saisies ? Pour les personnes placées sous main de justice (ou ayant des enfants dans ce cas), qu'en est-il de leur situation actuelle ou comment y remédier à meilleur compte ? Des dépôts de dossier de logement ou de relogement auprès de bailleurs privés ou d'offices HLM ont-ils été réalisés en cas de situation locative insatisfaisante et les aides au logement auxquelles la personne peut prétendre ont-elles été réclamées ?

Dans le cas présent, l'évaluation se résume en une *mesure* plus ou moins précise de l'efficacité de l'action. Il s'agit de dire si les objectifs fixés ont bien été atteints, par un contrôle de conformité par rapport au cahier de charges principalement. La *valeur* du travail produit est donc réduite à sa *mesure*. Dit autrement, le résultat obtenu semble suffire à apprécier la qualité de la prestation délivrée par le référent de parcours. On comprend mieux alors pourquoi les données qualitatives produites de manière surabondante par le professionnel pour satisfaire aux objectifs qui lui sont communiqués restent inexploitées par le financeur. S'ils servent, au quotidien, pour réaliser l'accompagnement demandé, ils ne sont jugés d'aucune utilité par le PLIE pour apprécier/évaluer la réponse apportée par l'opérateur. L'activité du référent de parcours (son investissement subjectif dans l'action notamment) est considérée sous le seul angle d'une série d'effets observés qui sont imputés à son action, au premier chef, les « solutions positives ». De même que les personnes accompagnées sont pensées et comptabilisées dans le cadre de l'évaluation en unités de « solutions positives », l'activité du professionnel est réduite et résumée à cet indicateur dont le chiffrage importe plus que tout. Son activité réelle, nécessairement complexe, est compactée en indicateurs quantitatifs finalement assez frustes.

Un dispositif de « mobilisation professionnelle » de jeunes en insertion

Si « motiver » et « mobiliser » les jeunes autour d'un « projet » personnel est le cœur de cible du dispositif ici considéré, comment les professionnels procèdent-ils pour ce faire et, rapportées aux préoccupations portées dans cette communication, comment apprécient-ils les résultats de leurs démarches ? Comment également le financeur de l'action, un Conseil régional en la circonstance,

évalue-t-il le service rendu par l'opérateur qui a remporté le marché public ? Et qu'en est-il de l'évaluation de l'activité des professionnels proprement dits ?

L'Espace dynamique d'insertion (EDI) s'inscrit dans un vaste programme proposé par un Conseil régional en vue de la formation et de l'insertion professionnelle des personnes âgées de 16 à 26 ans révolus, sorties sans qualification du système scolaire, résidant sur son territoire et en situation de demandeur d'emploi. Même si cela peut sembler tautologique, le Conseil régional tient à rappeler, dès le préambule du cahier des clauses techniques particulières que « l'insertion professionnelle des jeunes est bien la finalité [ultime] du dispositif régional d'insertion ». La formation et ses dispositifs afférents ne sont donc que les outils nécessaires à la réalisation de cet objectif. Dans ce cadre, « l'insertion professionnelle signifie prioritairement l'acquisition d'une qualification professionnelle et/ou d'une certification et/ou l'accès à l'emploi durable ». C'est à cette aune, semble-t-il, que sera jugée la qualité de la prestation fournie par les détenteurs des différents lots du marché public.

Tout un travail visant à canaliser l'énergie de jeunes issus de milieux populaires est assurément conduit. Il s'agit moins de lutter contre la « tchatche » et la mise en avant du capital physique qui structurent leurs pratiques et représentations sociales ordinaires que de les inviter à ouvrir une parenthèse le temps de la formation. Les tentatives d'oppositions frontales aux *habitus* intériorisés et revendiqués (de *bad boys* notamment) échouent lamentablement et nombre de professionnels ont renoncé à cette stratégie du passage en force. Les professionnels étant bien fréquemment eux-mêmes issus de cet univers, ils savent, plus souvent qu'à leur tour, combien ces référents identitaires participent de l'équilibre de vie de ces jeunes. Il ne s'agit donc pas de chercher à les « transformer » du tout au tout, mais d'introduire une dose de « pragmatisme » dans leur approche des milieux professionnels et du monde du travail. Les convaincre que le « racisme » n'explique pas tout et que la discrimination a (parfois) bon dos. Leur rappeler que la forme juridique du contrat de travail introduit la subordination, ce qui diffère certes de l'esclavage, de la domination sans limite ou de l'arbitraire incontrôlé, mais qui appelle de leur part le respect de la hiérarchie, l'application de procédures et de règlements, l'adoption de marques de civilité voire de déférence.

C'est donc une expérience marquée par la « concrétude » des relations quotidiennes qui signe, du point de vue des professionnels, la qualité du travail. Et c'est sur cette base qu'il leur paraît nécessaire d'évaluer leur contribution aux « évolutions » des jeunes en matière de « socialisation », « mobilisation » ou « insertion ». L'ineffable « motivation » est ainsi appréciée à l'aune de considérations qui renvoient au registre du savoir d'action. Les formateurs et chargés d'insertion s'accordent à reconnaître que la formalisation de cet ensemble de compétences sociales relève de la gageure. Tout en dénonçant les indicateurs retenus par le financeur pour apprécier la qualité du service rendu (les indicateurs de placement), ils reconnaissent la difficulté à quantifier différemment ce que le passage par l'EDI *fait* aux jeunes. Tout en regrettant que leur activité effective soit si peu restituée dans les documents produits à des fins d'évaluation, eux-mêmes peinent à décrire, par une *évaluation qualitative interne*, la valeur-ajoutée de leur travail. Dans le cadre de notre collaboration avec un responsable d'EDI, nous avons eu toutes les peines du monde à l'amener à écrire sur son activité ordinaire. Non qu'il se sentait illégitime pour le faire, non même qu'il n'en voyait pas l'intérêt, mais il peinait à traduire les savoirs d'action mis en œuvre en des formulations objectivées et, par certains côtés, totalisantes.

Une dimension qualitative apparaît dans les exigences formulées par le Conseil régional. Toutefois, elle ne porte pas sur le parcours des jeunes et les modalités pédagogiques mises en œuvre pour assurer socialisation, production de projets professionnels et insertion. Elle consiste à s'assurer que les moyens humains et matériels dévolus à l'action sont conformes (en nombre certes, mais aussi en qualité) aux attendus du cahier des charges. Pour n'en donner que deux brefs exemples, un responsable pédagogique a-t-il été désigné ? Ou alors, l'équipe pédagogique comprend-elle, comme indiqué dans le cahier des charges, des formateurs disposant d'une connaissance des publics visés, d'un diplôme de l'enseignement supérieur (licence ou équivalent), en sciences de l'éducation ou en formation pour adultes (pour au moins une personne de l'équipe)

ou une expérience professionnelle de trois ans dans le domaine de l'insertion des jeunes demandeurs d'emplois ou dans le champ professionnel d'un module abordé dans le cadre des parcours proposés ?

Évaluer certes, mais quoi évaluer ? Il semble que le Conseil régional ne sache pas bien lui-même sur quels éléments faire porter sa prescription dans le cadre de l'EDI. S'agit-il d'évaluer les acquis des bénéficiaires de l'action ? Clairement, c'est sur ce terrain qu'est porté le débat par le financeur puisqu'il a entrepris de mener une réflexion sur l'« évaluation des stagiaires ». Il reste que la discussion engagée pose question aux organismes de formation. Très explicitement, ils se demandent s'il ne s'agit pas plutôt d'autoproduire les outils de leurs propres évaluations. Un courrier adressé au directeur de la formation professionnelle du Conseil régional par le président de l'union régionale des organismes de formation insiste ainsi sur l'inquiétude manifestée par ses mandants à l'endroit d'une telle initiative. Loin de les satisfaire et de les rassurer, la réponse apportée a contribué à aviver les craintes puisqu'il est fait mention en son cœur à la nécessité « de développer ce type de démarche visant la qualité d'accueil et d'accompagnement des stagiaires ». Las, ce n'est donc pas l'évaluation des stagiaires qui semble recherchée mais bien celle des organismes de formation soumissionnaires via l'évaluation de la qualité de leurs prestations d'accueil et d'accompagnement. Et les formateurs s'interrogent également : sera-ce l'occasion, même incidemment, d'introduire une évaluation de leur travail ? On le sait, « l'évaluation ne peut [donc] être totalement dissociée d'un jugement, lequel se réfère de près ou de loin à des règles, des valeurs ou des croyances plus ou moins claires » (Amado et Enriquez, 2010, p. 7). Rien de très surprenant donc, dans un contexte où le travail sur les « compétences sociales » est si difficile à objectiver, que d'observer des formes d'attentisme des directions d'organismes de formation et de repli sur un certain quant à soi des formateurs et autres chargés d'insertion. L'enjeu est de taille : Dans le cadre d'une telle évaluation, la qualité du service rendu serait-elle reconnue et la professionnalité des interventions (et des intervenants) confirmée ? Avec, toujours, cette lancinante question : quels outils seront-ils mobilisés pour apprécier et mesurer cette activité assez peu formalisable du travail avec, pour et sur autrui qui porte pour nom l'« insertion professionnelle et sociale » ?

Conclusion

Certes l'évaluation n'est pas une pratique scientifique. Assurément idéologique, le premier enjeu est de savoir si elle est dialogique, partagée, ou si elle est mise au service des dominants et contribue ainsi la réification des rapports sociaux. Certes, l'évaluation est souvent quantophrénique. Assurément normative, le deuxième enjeu est de savoir si elle fait une place aux dimensions formatives ou si elle s'en tient aux orientations sommatives. Certes, l'évaluation est plutôt attentive aux effets, résultats, impacts. Assurément, elle se préoccupe peu de la « boîte noire » que constitue pour elle l'activité réelle et son flux quotidien. L'enjeu, bien connu depuis vingt ans au moins, est alors de procéder à une contre-évaluation « centrée prioritairement sur l'analyse du dispositif de formation lui-même, de ses composantes, de ses *articulations majeures*, et pas seulement sur la mesure de ses performances et l'analyse de données comptables » (Froissart et Aballéa, 1988, p. 13). Cette évaluation, qualitative, « repose sur l'idée que les performances d'une action, sa réussite, ses résultats, ne peuvent s'apprécier sans prendre en compte un ensemble d'éléments constitutifs de la conception, de la mise en œuvre et du déroulement de la formation elle-même. ».

Ce qui vaut pour les activités de formation vaut également pour les activités d'insertion. Tous les professionnels rencontrés dans notre travail de « recherche dans l'action » l'ont énoncé, chacun avec leurs mots : « l'évaluation est aveugle à notre activité de tous les jours » ; « l'évaluation ne rend pas compte de ce que l'on fait » ; « nos démarches avec les personnes reçues sont un impensé de l'évaluation », etc. Parmi les clés pour le débat, la principale concerne assurément la reconnaissance du travail effectué par les professionnels au contact de publics dits en difficulté. Tant que les évaluations réalisées leur paraîtront fortement éloignées de leur « cœur de métier », tant qu'elles apparaîtront ignorantes de leur activité ordinaire, il y a fort à parier qu'ils n'y

adhéreront que du bout des lèvres. La reconnaissance n'est pas un vain mot. Les travaux en sciences sociales ont bien montré, ces dernières années, combien les acteurs sociaux luttent pour la reconnaissance (Honneth, 2000), sont soucieux des marques de *reconnaissance* que l'on peut leur témoigner. Ils accordent également une grande attention au respect de leur *dignité* et sont attachés à la promotion de leur *estime de soi* (Payet et Battagay dir., 2008). Nulle surprise donc à constater les réserves que les formes actuelles d'évaluation suscitent chez les formateurs d'adultes ou les chargés d'insertion.

Bibliographie

Amado, G. & Enriquez, E. (2010). Introduction au dossier « La passion évaluative », *Nouvelle revue de psychosociologie*, 8, 7-10.

Divay, S. (2009). La médiation sociale : « un nouveau métier » plus de dix ans après son émergence ? In D. Demazière & C. Gadéa (Ed.), *Sociologie des groupes professionnels. Acquis récents et nouveaux défis* (pp. 242-253). Paris : La découverte.

Froissart, C. & Aballéa, F. (1988). *L'évaluation qualitative. Une approche appliquée à la formation*. Noisy-le-Grand : ADEP éditions.

Honneth, A. (2000). *La lutte pour la reconnaissance*. Paris : Le Cerf.

Martuccelli, D. (2010). Critique de la philosophie de l'évaluation. *Cahiers internationaux de sociologie*, CXXVIII-CXXIX, 27-52.

Payet, J.-P. & Battagay, A. (Ed.) (2008). *La reconnaissance à l'épreuve. Explorations socio-anthropologiques*. Villeneuve d'Ascq : Presses Universitaires du Septentrion.