

DES POLITIQUES AUX PRATIQUES D'ÉVALUATION EN BELGIQUE FRANCOPHONE : LA COHÉRENCE EN QUESTIONS

Dominique Lafontaine

Université de Liège, dlafontaine@ulg.ac.be

Mots-clés : politiques d'évaluation, épreuves externes, pratiques d'évaluation (maximum 5 mots-clés)

Résumé. A la fin des années 1990, la Belgique francophone a connu une évolution rapide de la politique scolaire en matière d'évaluation. En une quinzaine d'années, on est ainsi passé d'un système dont l'évaluation externe était totalement absente à un dispositif au maillage relativement serré. Dans cet article, nous présenterons les modèles auxquels peuvent être rattachés les différents dispositifs et outils d'évaluation en présence. La question de leur influence sur les pratiques d'évaluation en classe sera ensuite examinée. Nous concluons en interrogeant la cohérence sous des angles multiples : cohérence de la politique d'évaluation, cohérence entre politiques et pratiques d'évaluation, entre enseignement et évaluation, sur fond de débats vifs autour de l'évaluation des/par compétences.

1. Introduction

Cet article porte sur les liens entre politiques d'évaluation en Belgique francophone et pratiques d'évaluation en classe. Dans un premier temps, nous nous attacherons à fixer rapidement les balises temporelles qui jalonnent la mise en place progressive d'une politique d'évaluation en Belgique francophone. Dans un deuxième temps, nous présenterons les modèles auxquels peuvent être rattachés les différents dispositifs et outils d'évaluation en présence. La question de leur influence sur les pratiques d'évaluation en classe sera ensuite examinée. Nous concluons en interrogeant la cohérence sous des angles multiples : cohérence de la politique d'évaluation, cohérence entre politiques et pratiques d'évaluation, entre enseignement et évaluation, sur fond de débats vifs autour de l'évaluation des/par compétences.

2. Emergence d'une politique d'évaluation : repères temporels

En 1991, un rapport de l'OCDE consacré au système d'enseignement belge épingle deux faiblesses particulières de ce système – les taux de redoublement très élevés et les énormes disparités de performances entre établissements ; le même rapport pointe l'absence de culture et de politiques d'évaluation :

« Dans le système belge d'enseignement, qui repose sur des équilibres subtils, les autorités centrales responsables n'ont guère eu de moyens de préciser la qualité exacte de l'enseignement et d'évaluer sur des critères acceptés par tous, les performances réelles des différents établissements et de leurs élèves » (OCDE, 1991).

On se trouve donc, au début des années 1990, devant un vide évaluatif (en termes de politiques), un tout à l'évaluation interne et un éparpillement des exigences lié notamment à l'existence de multiples réseaux et pouvoirs organisateurs.

Ce rapport va provoquer un sursaut et dès 1994, le secrétaire général de l'Administration du Ministère de l'Education prend l'initiative de mettre sur pied un premier dispositif unifié d'évaluations externes non certificatives. Ces évaluations sont dites diagnostiques : elles interviennent en début d'année, n'ont pas de portée certificative et aident l'enseignant à identifier

les acquis et les difficultés des élèves, dans une perspective de régulation de son enseignement. Le dispositif est léger : il porte à l'époque sur le français et les mathématiques et ne concerne qu'un niveau d'étude par année.

En 1997 est promulgué un texte légal capital pour le système éducatif, le « décret Missions ». Ce décret définit les missions prioritaires (finalités) du système éducatif en Communauté française de Belgique et institue des référentiels de compétences communs pour l'ensemble des établissements, par-delà les programmes d'études propres aux différents pouvoirs organisateurs. Ce texte pose donc les jalons législatifs qui rendent possibles évaluations externes et pilotage. Dès lors qu'il existe désormais un prescrit commun, il devient possible et légitime de mesurer si les élèves possèdent les acquis attendus aux différents paliers de leur scolarité.

Par ailleurs, sans faire référence au dispositif d'évaluations externes préexistant, le même Décret instaure des commissions chargées de produire des outils d'évaluation qui seront diffusés « à titre indicatif » dans les établissements. Le Décret parlait à ce propos de « batteries d'épreuves étalonnées », ce qui aurait pu laisser présager le développement de « banques d'items » à des fins plutôt certificatives. On verra ultérieurement ce qu'il en est advenu.

Quelques années plus tard, le Décret du 27 mars 2002 vient renforcer la politique d'évaluation et de pilotage du système éducatif. Une Commission de pilotage rassemblant des représentants de tous les acteurs du monde éducatif est créée et les évaluations externes non certificatives (ci-après EENC) sont désormais inscrites dans un cadre légal. En 2006¹ et en 2009, le dispositif d'évaluations externes se trouve consolidé et renforcé par décret, en même temps qu'est prévu un développement progressif d'épreuves externes certificatives (EC) à la fin du primaire, au 1er degré du secondaire et en fin de secondaire.

Le dispositif complet est aujourd'hui celui-ci ; les évaluations internationales ont été incluses dans le tableau pour donner un aperçu général de la « pression » évaluative.

S6			Tess	Tess	Tess	Tess	Tess	Tess
S5				Lecture		Langues en S5 + une matière et une année à définir	Lecture	
S4			Sciences		Math.			une matière et une année à définir
S3								
2e com/compl	Lecture-prod. écrit	Math.	CE1D Sciences-Eveil	CE1D Lecture-prod. écrit	CE1D Langues Math.	CE1D Sciences-Eveil	CE1D Lecture-prod. écrit	CE1D Langues
2e diff.	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.
			CEB	CEB	CEB	CEB	CEB	CEB
1re diff		CEB	CEB	CEB	CEB	CEB	CEB	CEB
P6	CEB	CEB	CEB	CEB	CEB Langues	CEB	CEB	CEB
P5	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.
P4								
P3								
P2	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.	Sciences-Eveil	Lecture-prod. écrit	Math.
	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014
								2014-2015

¹ Décret du 2 juin 2006, modifié le 30 avril 2009.

Source : Service du pilotage, Administration générale de l'Enseignement et de la Recherche scientifique

En une quinzaine d'années, la Belgique francophone est donc passée d'un système où n'existait aucune évaluation externe commune à un système où les évaluations externes standardisées occupent une place importante. Deux particularités sont à pointer. Si les évaluations externes sont désormais nombreuses, elles s'inscrivent dans ce que l'on peut appeler un pilotage « bienveillant ». D'une part, les évaluations externes non certificatives (les plus nombreuses) sont sans enjeux pour les élèves et il est explicitement interdit de s'en servir pour évaluer les enseignants et les établissements. D'autre part, les évaluations externes certificatives, dont les enjeux sont un peu plus importants, ont été instaurées de manière progressive, sur une base volontaire, en veillant à associer les enseignants à leur élaboration. Leur arrivée sur le terrain s'est donc faite d'une manière que l'on peut qualifier de relativement « douce », à l'opposé de nombre de réformes antérieures perçues comme imposées par les politiques et dépossédant les enseignants de leur professionnalité.

3. Modèles d'évaluation en présence

Les épreuves d'évaluation externes certificatives et non certificatives ainsi que les outils élaborés par les « commissions d'outils » instaurées par le décret Missions (1997) sont tous coordonnés par un même service, le Service général du pilotage du système éducatif. Les dispositifs sont décrits en détail sur le site <http://www.enseignement.be> ; tous les outils et épreuves sont accessibles et téléchargeables au départ de ce site. Notre propos n'est pas ici d'entrer dans le détail de ces dispositifs, mais de caractériser à grands traits à quels modèles ils se rattachent, ceci afin de mettre en évidence d'éventuelles tensions entre les différentes évaluations.

En schématisant, on peut dire que deux modèles sont en présence. Les évaluations externes se rattachent au premier modèle, tandis que les outils des commissions se rattachent au second.

3.1. *Epreuves externes certificatives et non certificatives*

Pour rappel, les épreuves externes² ont pour finalités le diagnostic, le bilan des acquis des élèves et/ou la régulation de l'enseignement (Lafontaine, 2006). Elles sont élaborées selon les théories scientifiques de la mesure éprouvées, font appel à des procédures standardisées, font l'objet de traitements statistiques et d'une validation empirique de la qualité psychométrique. Elles comportent un nombre d'items important, parmi lesquelles des QCM, mais aussi des questions ouvertes le plus souvent à réponse courte, parfois élaborées. Elles évaluent principalement des compétences de 2^e degré et des procédures (selon le modèle de Rey, Carette, Defrance & Kahn, 2003), plus rarement des compétences de haut niveau (choisir et combiner des procédures). Ces épreuves concernent toutes les écoles et tous les enseignants du niveau concerné. Le service de pédagogie de l'Université de Liège dirigé par M. Crahay et ensuite D. Lafontaine apporte son soutien aux évaluations externes non certificatives depuis leur création ; les épreuves certificatives ne font pas appel aux services d'experts universitaires.

3.2. *Epreuves des commissions d'outils*

La finalité de ces épreuves est toute différente. Celles-ci ont avant tout une valeur d'exemple, comme indiqué sur le site <http://www.enseignement.be>
« Les outils d'évaluation, conçus pour les élèves de la première étape (2^e année primaire) et de la deuxième étape (6^e année primaire), constituent des exemples concrets d'une évaluation de compétences ».

² Les évaluations internationales s'inscrivent dans le même modèle.

Les commissions se sont mises en place progressivement à partir de 2001, avec pas mal de tâtonnements au démarrage. Pour baliser le travail des commissions, il a été fait appel à des experts universitaires : l'équipe de l'Université libre de Bruxelles (Vincent Carette, Sabine Kahn et Bernard Rey) pour le primaire, Jacqueline Beckers de l'Université de Liège pour le secondaire. Les « batteries d'épreuves étalonnées » du Décret sont ainsi devenues des « modèles » d'épreuves destinés à aider les enseignants à construire des épreuves d'évaluation de compétences. Elles sont diffusées à titre informatif sur le site <http://www.enseignement.be>

Ces outils proposent une évaluation des compétences, qui fait appel à des tâches complexes et inédites, d'une manière sensiblement distincte au primaire et au secondaire. Dans le primaire, les outils sont construits selon le modèle en 3 phases de Carette et Rey ; les outils comportent donc une perspective diagnostique. Dans le secondaire, les épreuves sont basées sur le concept de famille de tâches (dans l'enseignement général) ou de situations professionnellement significatives (dans l'enseignement professionnel) (Beckers) et n'ont pas de portée diagnostique.

Les épreuves ne sont pas étalonnées ; selon le site [enseignement.be](http://www.enseignement.be), les outils du primaire « ont été testés dans des écoles appartenant aux différents réseaux ». Pour le secondaire, rien n'est précisé à cet égard. On peut raisonnablement avancer que la validation empirique est limitée, et se fait à petite échelle.

4. Les pratiques d'évaluation en classe : usages en lien avec les épreuves externes et outils

Le moment est venu de se pencher sur la question de l'usage ou de l'appropriation par les enseignants de ces différentes épreuves et outils mis à leur disposition. Pour ce faire, nous nous fonderons sur trois sources : les rapports élaborés par l'inspection (2009 et 2010), accessibles sur le site <http://www.enseignement.be>³, les données recueillies par le questionnaire bilan envoyé par le service du pilotage après chaque campagne d'évaluations externes non certificatives (Lafontaine et Dierendonck, 2008) et une enquête d'Ouzzine, citée par Carette & Dupriez (2009).

Selon les rapports de l'inspection primaire et secondaire, « les travaux de la commission des outils d'évaluation qui n'existent pas encore pour toutes les disciplines et tous les niveaux restent peu connus et donc peu utilisés. Même dans les disciplines dans lesquelles ces outils sont davantage connus, la proportion d'utilisation reste faible ».

Quelles épreuves servent dès lors de référence aux enseignants ?

Selon Ouzzine, cité par Carette & Dupriez (2009), « les épreuves auxquelles les enseignants se réfèrent le plus étaient, pour les instituteurs primaires, les épreuves certificatives de fin de l'école primaire » (p. 41).

Selon Carette (2006, 2008), « l'évaluation des compétences telles que suggérée à travers les outils produits par les commissions d'outils d'évaluation, reste fort éloignée des pratiques effectives de la majorité des enseignants » (p. 42).

En bref, et sans réelle surprise, les épreuves les plus influentes seraient les épreuves certificatives fin du primaire, avec un possible effet en retour sur les pratiques : ce sont des épreuves à enjeux pour les élèves, les plus proches des pratiques d'enseignement, par définition et par nécessité. On ne peut en effet certifier d'une manière équitable que ce qui a été enseigné à tous les élèves, d'une manière suffisamment égale.

On sait par ailleurs que les épreuves externes non certificatives jouissent d'un succès d'estime auprès des enseignants (Lafontaine & Dierendonck, 2008) : même si ces épreuves n'ont pas

³ Voir <http://www.enseignement.be/index.php?page=24234>

d'enjeux réels pour les élèves, elles font l'objet d'une large diffusion, toutes les classes y participent, elles s'accompagnent enfin d'un suivi rapproché (résultats, pistes didactiques, suivi par l'inspection et les conseillers pédagogiques, dans le cadre de la formation continue) ; tous ces aspects contribuent à asseoir leur légitimité.

A côté de cela, on note une diffusion et utilisation restreinte – mais sans doute en augmentation – des outils produits par les commissions. Ceux-ci sont les seuls à représenter une approche « orthodoxe » de l'évaluation par compétences qui implique la mobilisation de ressources pour faire face à des tâches ou situations complexes et inédites.

Ceci nous amène inévitablement à poser la question de la cohérence entre ces différents outils et épreuves d'évaluation.

5. La cohérence en questions

Avant de poser la question de la cohérence des politiques, il n'est sans doute pas inutile de souligner qu'interviennent dans le contexte du système éducatif de la Fédération Wallonie-Bruxelles, différents experts de l'évaluation assez connus internationalement : l'équipe Carette et Rey (Université libre de Bruxelles), Jacqueline Beckers et Dominique Lafontaine (Université de Liège, mais équipes et conceptions en partie différentes) et enfin l'équipe du BIEF (Université catholique de Louvain) autour de Xavier Roegiers, François-Marie Gerard et Jean-Marie De Ketele⁴. Ces différents experts ont des conceptions que l'on peut dire, en une première approche, opposées de l'évaluation de compétences. Il ne fait pas de doute qu'à ces oppositions d'idées se conjuguent des enjeux de partage de territoires, de ressources et d'influences dont il serait vain de nier l'existence.

5.1. Des épreuves non pertinentes ?

Dans ce contexte « électrique », certains auteurs mettent en avant une forme d'incompatibilité entre épreuves externes et évaluations de compétences : « Pour les épreuves externes non certificatives et l'épreuve certificative de fin de primaire, l'évaluation des compétences ne semble pas se démarquer fondamentalement des épreuves inspirées de la pédagogie par objectifs » (Carette & Dupriez, 2009, p. 38). Dans le même esprit, De Ketele & Gerard (2005) soulignent que des épreuves comme PISA⁵, « sont fiables et valides, mais ne sont pas pertinentes par rapport à l'approche par compétences et donc par rapport aux finalités des systèmes éducatifs conçus dans cet esprit » (p. 5).

Ces auteurs ont tendance à mettre en avant l'opposition de modèles : dès lors que l'on a recours au modèle classique ou scientifique de la mesure (validité, fidélité, représentativité, standardisation, items nombreux) qui s'appuie sur l'assomption selon laquelle la compétence peut être mesurée via différents items, il n'y aurait pas d'évaluation de compétences « digne de ce nom ». De Ketele, Gerard et Roegiers, en particulier, dénoncent régulièrement la non-pertinence des évaluations externes. Ceci nous amène à questionner à notre tour la question de la non-pertinence. Si non-pertinence il y a, par rapport à quoi les évaluations externes se révèlent-elles non pertinentes ?

Ces évaluations externes, on l'a vu, ont pour finalités d'évaluer ou de certifier les acquis des élèves, en relation, bien entendu, avec le prescrit, formulé en termes de compétences. Le prescrit légal et commun (référentiels) se prononce donc sur les finalités, avec une définition « soft » ou « molle » de la compétence qui ne prend pas d'option sur la question devenue brûlante dans les

⁴ Dont on n'a pas parlé jusqu'ici, mais qui prônent une approche par compétences « orthodoxe » et une approche évaluative en cohérence avec cette approche.

⁵ Il nous semble que l'on peut, sans trahir la pensée de ces auteurs, considérer que ce qu'ils avancent à propos des épreuves PISA s'applique aussi aux évaluations externes.

années 2000 des tâches complexes et inédites. Le prescrit ne se prononce pas sur les méthodes d'enseignement susceptibles d'amener les élèves aux compétences visées ; il est même impossible qu'il en aille ainsi, en vertu de la loi sur le Pacte scolaire de 1959, qui dit que « chaque pouvoir organisateur est libre en matière de méthodes pédagogiques ». Il nous paraît dès lors difficile, si ce n'est par un artifice rhétorique qui relève de l'argument d'autorité, de disqualifier les épreuves externes certificatives et non certificatives au motif qu'elles ne seraient pas pertinentes « par rapport à l'approche par compétences et donc par rapport aux finalités des systèmes éducatifs conçus dans cet esprit » (De Ketele & Gerard, 2005) (c'est nous qui soulignons). Le système éducatif de la Fédération Wallonie-Bruxelles, jusqu'à nouvel ordre, vise certes des compétences (et des savoirs)⁶, il n'a pas opté pour une approche méthodologique par compétences et encore moins pour la forme la plus affirmée de cette approche.

5.2. Des politiques d'évaluation ambiguës ?

Compte tenu de ce qui précède, il nous est difficile de suivre l'analyse proposée par Carette & Dupriez (2009) qui, constatant que les responsables de l'enseignement mettent à la disposition des enseignants des épreuves externes et des outils traduisant toutes à leur manière les « attendus », émanant du même service du pilotage, considèrent que l'« on peut s'interroger sur les messages quelque peu contradictaires transmis par ces différentes épreuves » (Carette & Dupriez, 2009, p. 39) et soulignent sur cette base l'ambiguïté des politiques en matière d'évaluation mises en œuvre par le service du pilotage. A nouveau, au nom de quoi peut-on considérer qu'il y a ambiguïté ?

Les épreuves et outils peuvent certes, dans une optique de chercheurs, être conçus comme relevant de modèles ou d'approches différentes, voire contradictoires. Mais n'est-ce pas aller un peu vite en besogne que d'en tirer des conclusions quant aux politiques de l'administration, qui s'inscrivent dans une autre logique que la référence à des modèles théoriques ?

L'objectif du service du pilotage est double : vérifier ou aider à certifier les acquis des élèves (via les épreuves externes) et proposer des exemples d'évaluation de compétences (via les outils produits par les commissions). Il doit, dans la poursuite de ces objectifs, tenir compte d'un certain nombre de contraintes (temps de passation limité, correction rapide par des enseignants non formés à cet effet) et de ce que sont les pratiques d'évaluation habituelles des enseignants. Il serait en effet inadmissible, dans le cas des épreuves certificatives, et sans doute inadéquat, dans le cas des épreuves non certificatives, d'évaluer les élèves à l'aide d'épreuves ne correspondant que de très loin aux épreuves auxquelles enseignants et élèves sont accoutumés.

Or que sait-on des pratiques d'évaluation en classe ? Peu de choses, car il n'existe pas d'enquête récente sur le sujet. Si l'on en croit le rapport de l'inspection, les évaluations portent souvent sur les savoirs et procédures et rarement sur des tâches nécessitant la mise en œuvre de compétences. « *A quelques exceptions près, l'évaluation de compétences est peu présente et les épreuves souvent éloignées de ce qui est prescrit dans les Socles de Compétences. Ainsi, en français, constate-t-on dans les épreuves d'évaluation une prédominance accordée aux outils (grammaire) le plus souvent sans lien avec les compétences relevant du lire, de l'écrire, du parler et de l'écouter* » (Rapport de l'Inspection, 2009-2010).

5.3. Quelle cohérence privilégier, en théorie et en pratique ?

Les épreuves externes et le modèle de la mesure sur lequel elles s'appuient sont parfois présentés comme « en retard », « dépassés », non pertinents en référence à l'approche par compétences et par une extension que nous jugeons abusive, aux référentiels ; selon cette vision, les épreuves externes seraient un obstacle à la mise en place de l'approche et de l'évaluation des compétences.

⁶ Comme l'indique le titre des référentiels pour les Humanités générales et technologiques, qui s'intitulent « Compétences terminales et savoirs requis ».

Dans la perspective qui est la sienne, le service du pilotage se doit de proposer des épreuves d'évaluation externes qui obéissent à une double contrainte : d'une part, ne pas s'éloigner d'une manière radicale des épreuves habituellement pratiquées par les enseignants, d'autre part évaluer non seulement des savoirs et procédures, mais aussi des compétences, en conformité avec les référentiels. A cela s'ajoutent les exigences classiques de validité et de fidélité. En ce qui concerne les outils, qui constituent des modèles d'évaluation de compétences, il peut s'affranchir du premier type de contrainte, dans la mesure où leur utilisation est libre et choisie : s'en saisissent les enseignants qui le souhaitent. A condition qu'ils soient utilisés en classe à des fins formatives ou diagnostiques, l'exigence de validité ou de fidélité peut aussi se relâcher quelque peu.

6. En conclusion : peut-on dépasser les oppositions d'école et comment ?

Plutôt que de mettre en avant la rupture de modèle entre épreuves externes et outils (vision de chercheurs), plutôt que de les concevoir comme un frein ou un obstacle, il nous semble possible de concevoir les épreuves externes comme un premier jalon sur le chemin d'une évaluation des compétences, en insistant sur la complémentarité⁷ ou la graduation entre les épreuves et les outils. Ce faisant, on s'inscrit dans une perspective dynamique et constructive plutôt que dichotomique, qui est à nos yeux moins stérile et davantage porteuse d'ouverture.

Dans cette perspective, les épreuves externes se situeraient dans la zone de développement proche des enseignants, ce qui est sans doute souhaitable dans l'impulsion d'une dynamique de changement. Les enseignants seraient d'autant plus disposés à les adopter qu'elles tiennent compte d'un certain nombre de contraintes pratiques du métier d'enseignant (temps d'administration et de correction raisonnables).

Si l'on reprend l'exemple des évaluations en français dans le primaire, les épreuves s'ordonneraient selon un continuum savoirs-compétences qui irait des pratiques d'évaluation en classe les plus traditionnelles (souvent centrées sur les outils de la langue) aux outils des commissions qui font souvent intervenir la lecture de plusieurs documents pour écrire un texte ou résoudre un problème. Entre les deux, on trouverait l'épreuve externe certificative, qui évalue à la fois les compétences lire-écrire-parler-écouter et la maîtrise des outils de la langue, et les épreuves externes non certificatives qui portent davantage sur les compétences et moins sur les outils. Dès lors que les élèves sont, dans ce dernier cadre, invités à lire des textes entiers d'un certain degré de complexité, voire un album entier (en 2^e année primaire), il nous paraît peu justifié, ou excessivement dogmatique, de considérer qu'il n'y a pas d'évaluation de compétences du tout. Comprendre un texte nouveau est une situation complexe et inédite et l'élève doit bien, pour faire face aux tâches demandées, mobiliser des savoirs et connaissances tels que ses habiletés grapho-phonétiques, sa connaissance du vocabulaire, des genres et types textuels, sans oublier ses connaissances antérieures liées au sujet traité.

Existe-t-il, en l'état de la science, d'autres manières de réduire l'opposition de modèles ou la distance entre épreuves externes et outils des commissions ? Commençons par formuler une question qui est rarement mise sur la table : les outils des commissions pourraient-ils ou devraient-ils se rapprocher des épreuves externes, plus classiques, moins conformes à une évaluation de compétences au sens fort du terme ? N'en déplaise aux tenants de ce courant, la réponse est dans une certaine mesure oui. Les outils produits par les commissions auraient tout avantage à se

⁷ Jacqueline Beckers, lors de son intervention dans le cadre d'une table ronde, a également souligné la complémentarité entre épreuves externes et outils d'évaluation.

soumettre davantage à l'épreuve des faits⁸. Il ne s'agit pas ici d'inciter à la mise en œuvre de modèles statistiques sophistiqués, tels que les modèles de la réponse à l'item, mais, plus simplement, de soumettre les épreuves à des échantillons d'élèves en sorte que l'enseignant qui les utilise puisse avoir un point de référence en termes de pourcentage de réussite. Tester si les critères d'évaluation sont utilisés d'une manière cohérente d'un correcteur à l'autre, en particulier par des correcteurs enseignants qui ne sont pas les concepteurs des outils, donnerait d'intéressantes indications susceptibles de renforcer la confiance que l'on peut attendre de tels outils (fiabilité).

Inversement, les épreuves externes pourraient-elles et gagneraient-elles à se rapprocher des modèles d'évaluation de compétences ? La réponse par l'affirmative s'impose aussi. Dans une certaine mesure, il serait intéressant d'insérer dans les épreuves externes une ou des tâches qui seraient plus complexes. En revanche, il paraît difficile, compte tenu des exigences de validité, de fidélité et d'équité dont ce type d'épreuves peut difficilement faire fi, de changer radicalement d'approche, autrement dit de basculer dans une évaluation de compétences au sens fort du terme.

En effet, comme le souligne Bain, les évaluations externes ne peuvent faire fi de la congruence entre enseignement et évaluation. « *Evaluer des compétences complexes, c'est bien, encore faut-il que les élèves aient reçu l'enseignement adéquat!* » (Bain, 2008, p. 140).

Par ailleurs, les tenants de l'approche par compétences eux-mêmes (Beckers, 2005; Gerard, 2007) posent comme conditions à la mise en œuvre d'évaluations des compétences par tâches complexes et inédites le fait d'une part que l'ensemble des ressources aient été apprises par les élèves, d'autre part que les élèves aient été confrontés à d'autres situations de même niveau de complexité et appartenant à la même famille de situations.

Enfin, comme le souligne Chenu (2012), « le problème de l'objectivité et de la validité du concept de familles de situations n'est pas prêt d'être solutionné » (p. 52). Au terme d'une étude expérimentale et clinique portant sur la manière dont les individus regroupent les situations en famille dans le domaine de la bureautique, Chenu conclut que « non seulement les personnes n'utilisent pas les mêmes critères pour examiner les choses, mais en plus, les critères retenus ont des poids différents » (p. 51). La question de la généralisabilité reste donc entière.

En conclusion, plutôt que d'opposer les modèles et de jouer la carte de la pertinence contre celle de la validité et de la fiabilité, il nous paraît important de réaffirmer que toute épreuve ou outil d'évaluation doit répondre, au moins dans une certaine mesure, à ces trois exigences. Dans cette perspective, les outils d'évaluation gagneraient à faire davantage ou plus souvent l'objet d'une validation empirique de base (essai de terrain sur un échantillon suffisant, contrôle de la fidélité entre correcteurs du terrain, contrôle de l'équivalence des tâches supposées appartenir à la même famille...). De leur côté, les épreuves externes non certificatives, dans l'optique d'un travail dans la zone de développement proche des enseignants, pourraient progressivement incorporer des tâches plus complexes d'évaluation des compétences. En procédant ainsi par « tuilage »⁹, la distance qui sépare les épreuves et les outils se trouverait réduite, et le passage des unes aux autres moins laborieux.

⁸ Hébert, Valois et Frenette (2008), au terme d'une revue de la littérature approfondie anglophone et francophone, ont mis en évidence que « les études empiriques ayant abordé la validité d'outils alternatifs d'évaluation des apprentissages à l'enseignement primaire et secondaire au moyen de stratégies de validation dites « classiques » apparaissent absentes de la littérature francophone », alors que « la recension de la littérature anglo-saxonne a révélé que quelques chercheurs s'y sont intéressés ». (p. 5-6).

⁹ L'expression est de Linda Allal, discutante dans le symposium.

Références

- Bain, D. (1999). De l'évaluation aux compétences : mise en perspective de pratiques émergentes. In J. Dolz & E. Ollagnier (Eds). *L'énigme de la compétence en éducation* (pp. 129-145). Bruxelles : De Boeck, coll. Raisons éducatives.
- Beckers, J. (2002). *Développer et évaluer des compétences à l'école : vers plus d'efficacité et d'équité*. Bruxelles : Labor.
- Carette, V. (2008). Les caractéristiques des enseignants efficaces en question. *Revue française de pédagogie*, 132, 81-94.
- Carette, V. & Dupriez, V. (2009). La lente émergence d'une politique scolaire en matière d'évaluation des élèves. Quinze ans de transformations en Belgique francophone. *Mesure et évaluation*, 32 (3), 23-47.
- Chenu, F. (2012). *L'évaluation des compétences. De la méthodologie de la réflexivité garante de la transférabilité aux perspectives ouvertes par la didactique professionnelle*. Université de Liège, Thèse de doctorat en sciences de l'éducation.
- De Ketele, J.-M. & Gerard, F.-M. (2005). La validation des épreuves d'évaluation selon l'approche par les compétences. *Mesure et évaluation*, 28 (3), 1-27.
- Gerard, F.-M. (2007). La complexité d'une évaluation des compétences à travers des situations complexes : nécessités théoriques et exigences du terrain. *Actes du colloque international « Logique de compétences et développement curriculaire : débats, perspectives et alternative pour les systèmes éducatifs »*. Montréal : Oré, 1-15.
- Hébert, M.-H., Valois, P. & Frenette, E. (2008), La validation d'outils alternatifs d'évaluation des apprentissages : progressiste ou rétrograde ? *Actes du 20^e Colloque de l'Association pour le développement des méthodologies d'évaluation en éducation en Europe*, Université de Genève.
- Lafontaine, D. (2006). « D'un dispositif d'évaluations externes centré sur l'enseignant à un dispositif centré sur l'établissement : quelles mutations pour quels objectifs ? ». Actes du 19^e colloque international de l'Adméc, 11-13 septembre 2006, Luxembourg (en ligne sur le site <http://jemacs.uni.lu/>).
- Lafontaine, D., & Dierendonck, C. (2008). Comment les évaluations externes des acquis des élèves sont-elles perçues par les enseignants en Communauté française de Belgique et en Suisse romande ? In L. Mottier Lopez, Y.-E. Dizerens, G. Marcoux & A. Perréard Vité (Éds.). *Entre la régulation des apprentissages et le pilotage des systèmes : évaluations en tension. Actes du 20^e colloque de l'ADMEE-Europe*, Université de Genève. <http://hdl.handle.net/2268/19535>
- OCDE (1991). *Les systèmes éducatifs en Belgique : similitudes et divergences*. Bruxelles : Ministère de l'Education, de la Recherche et de la Formation.
- Ouzzine, A. (2009). *Analyse de l'influence des épreuves externes d'évaluation de compétences sur les pratiques d'évaluation*. Bruxelles : Université libre de Bruxelles. Mémoire de maîtrise complémentaire en pédagogie universitaire et de l'enseignement supérieur.
- Rey, B., Carette, V., Defrance, A. & Kahn, S. (2003). *Les compétences à l'école. Apprentissage et évaluation*. Bruxelles : De Boeck.