

---

## Les évaluations externes à faibles enjeux : les enseignants confrontés à eux-mêmes ?

*Low-stakes testing and teaching practices: a mirror effect?*

Esteban Rozenwajn et Xavier Dumay

---



### Édition électronique

URL : <http://journals.openedition.org/rfp/4977>

DOI : 10.4000/rfp.4977

ISSN : 2105-2913

### Éditeur

ENS Éditions

### Édition imprimée

Date de publication : 31 mars 2016

Pagination : 71-90

ISSN : 0556-7807

### Référence électronique

Esteban Rozenwajn et Xavier Dumay, « Les évaluations externes à faibles enjeux : les enseignants confrontés à eux-mêmes ? », *Revue française de pédagogie* [En ligne], 194 | 2016, mis en ligne le 01 juin 2020, consulté le 04 janvier 2021. URL : <http://journals.openedition.org/rfp/4977> ; DOI : <https://doi.org/10.4000/rfp.4977>

---

© tous droits réservés

# Les évaluations externes à faibles enjeux : les enseignants confrontés à eux-mêmes ?

Esteban Rozenwajn  
Xavier Dumay

Depuis les années 1990, l'évaluation externe des élèves est progressivement devenue en Europe continentale un outil pour réguler les pratiques d'enseignement. Pour ses promoteurs, la confrontation aux résultats de l'épreuve externe susciterait chez les enseignants un processus réflexif aboutissant à une amélioration de la qualité de l'enseignement. Au travers d'une étude par questionnaire menée auprès d'enseignants belges du secondaire inférieur confrontés à l'épreuve certificative de la fin du 1<sup>er</sup> degré (CE1D), cet article tente d'éclairer l'impact des dispositifs d'évaluation sur les pratiques d'enseignement. L'objectif de l'étude est de préciser les changements survenant dans les pratiques enseignantes, mais aussi de mieux comprendre les processus cognitifs intervenant dans le traitement des informations liées aux épreuves externes. L'analyse des données recueillies illustre un lien généralement faible entre les pratiques d'enseignement et l'évaluation externe, fort probablement en raison du fait qu'une majorité d'enseignants ne perçoit pas les résultats de l'évaluation externe comme un feedback de performances. Indépendamment des scores des élèves toutefois, la façon dont l'enseignant perçoit l'épreuve externe ainsi que l'influence sociale de son entourage professionnel semblent jouer un rôle dans les changements rapportés. Ces variables entretiennent néanmoins des liens beaucoup plus importants avec l'intention de changement qu'avec les changements effectifs, permettant de supposer que la faiblesse des modifications comportementales des enseignants ne dépend peut-être pas d'un manque de volonté de leur part.

**Mots-clés (TESE) :** évaluation externe, pratique pédagogique, résultat d'évaluation, politique en matière d'éducation.

## Introduction

L'évaluation externe, considérée ici comme toute épreuve conçue par des institutions relativement indé-

pendantes des établissements scolaires mais néanmoins imposée à l'ensemble des élèves d'une région et d'un niveau déterminés, fait l'objet d'un engouement politique particulièrement prononcé depuis les

années 1990. Pendant longtemps, la fonction majeure de ce type de dispositifs a été d'évaluer de manière standardisée et commune tous les élèves d'une cohorte donnée à différents moments-clés de leur scolarité. À l'heure actuelle toutefois, le rôle des épreuves externes s'est étendu de manière considérable à des fonctions régulatrices du système éducatif (Mons, 2009). À cet égard, certains auteurs et/ou acteurs politiques voient en ces dernières un instrument susceptible de contribuer à l'amélioration des pratiques d'enseignement. Pour ceux-ci, l'idée est que les résultats des élèves à ces épreuves servent de feedback aux enseignants et les engagent dans un processus réflexif visant à ajuster leurs pratiques d'enseignement et à améliorer ainsi les performances des élèves. Les mécanismes au travers desquels les résultats de l'évaluation externe pourraient produire une telle amélioration des pratiques font cependant débat. Pour certains auteurs (Klieme, 2004; Thélot, 2002), la réflexivité spontanée des acteurs de terrain serait suffisante pour susciter le changement. La simple confrontation aux résultats issus des épreuves devrait alors permettre de stimuler la tendance à l'optimisation des pratiques. Pour d'autres (Hanushek & Raymond, 2002; Phelps, 2005) au contraire, la simple confrontation aux résultats ne serait pas suffisante. Considérant la motivation à éviter des sanctions ou à obtenir des récompenses comme déterminante du comportement humain, ceux-ci estiment alors qu'il convient d'associer des incitants explicites aux résultats afin qu'ils soient en mesure de réguler les pratiques des acteurs.

Ces deux conceptions se traduisent dans les faits par la distinction régulièrement opérée entre dispositifs d'évaluation externe à forts et faibles enjeux (Dupriez & Mons, 2011). Principalement caractéristiques des pays anglophones, les premiers supposent que les résultats soient associés à des conséquences potentiellement importantes non seulement pour les élèves, mais aussi pour la carrière de l'enseignant ou pour le fonctionnement de l'établissement (production de plans d'amélioration, réorganisations du travail, changement de la direction...). Par contraste, davantage caractéristiques du continent européen, les seconds ne reposent généralement pas sur des conséquences formelles associées aux résultats pour les enseignants et/ou les établissements.

Mais si les dispositifs d'évaluation externe sont en forte expansion depuis les années 1990, qu'en est-il de leur efficacité? Influencent-ils effectivement les pratiques enseignantes, et surtout, de quelles manières?

Le lien entre l'évaluation externe et les pratiques enseignantes est au cœur de notre étude. Réalisée en Belgique francophone, par rapport au dispositif d'évaluation externe imposé aux élèves à la fin de la deuxième année secondaire, la présente étude répond à un double objectif. Le premier est de faire un état des pratiques et changements rapportés par les enseignants en lien avec l'évaluation externe en distinguant, sur la base de la littérature existante, des changements de nature diverse (adaptations des méthodes pédagogiques, changements curriculaires, utilisation plus instrumentale des matériaux de l'évaluation...). Le second est d'éclairer plus spécifiquement la relation entre les résultats de l'évaluation externe et les changements dans les pratiques enseignantes en prenant en compte les mécanismes cognitifs susceptibles d'intervenir dans leur interprétation. L'étude corrélationnelle menée auprès d'un échantillon d'enseignants du secondaire inférieur (n = 253) indique d'abord que, conformément à la littérature empirique existante, les changements suscités par l'évaluation externe sont plutôt instrumentaux et curriculaires que centrés sur les méthodes pédagogiques proprement dites. Des régressions multiples et en pistes causales permettent ensuite de mettre en évidence le fait que les résultats de l'épreuve externe ne contribuent que très faiblement aux changements dans les pratiques enseignantes. Et pour cause probable, la plupart des enseignants ne considèrent pas les résultats de l'évaluation externe comme une mesure légitime de leurs compétences, c'est-à-dire comme un feedback personnel. Indépendamment des scores obtenus par les élèves, ce sont des facteurs liés à la façon dont l'enseignant perçoit l'épreuve elle-même et à la cognition sociale qui semblent davantage déterminants dans l'évolution des pratiques enseignantes.

Dans la suite de cet article, nous commençons par proposer un état de la littérature concernant le lien entre évaluation externe et pratiques enseignantes d'une part, et le lien plus spécifique entre résultats et pratiques enseignantes d'autre part. La revue de la littérature sera ensuite complétée par une conceptualisation des mécanismes cognitifs impliqués dans l'interprétation des résultats qui précédera la présentation de l'étude par questionnaire que nous avons réalisée. Les apports de l'étude au débat concernant la régulation des pratiques d'enseignement par les dispositifs d'évaluation externe à faibles enjeux seront finalement discutés.

## Revue de la littérature

### Évaluations externes et pratiques enseignantes

L'étude des changements de pratiques liés à l'évaluation externe dans différents contextes aura principalement contribué à faire émerger le concept du « *teaching to the test* » pour souligner le fait que les épreuves externes génèrent une tendance à l'alignement de l'enseignement sur les exigences de l'épreuve. Derrière cette notion se cachent notamment des pratiques de focalisation des contenus enseignés sur ceux de l'épreuve et l'utilisation de l'épreuve comme source d'inspiration à des fins internes. Dans sa métasynthèse qualitative, Au (2007) constate en ce sens que près de 70 % des études nord-américaines recensées rapportent que l'évaluation standardisée aurait eu pour conséquence la « contraction » curriculaire, entendant par-là l'exclusion des matières non couvertes par l'épreuve externe. Dans une certaine mesure, cette tendance à la focalisation de l'enseignement sur les contenus de l'épreuve semble traverser les différents contextes puisque la planification de l'enseignement en fonction du contenu testé et la réduction du temps attribué aux matières non testées apparaissent également dans des études menées en France et en Suisse (Dutercq & Lanéelle, 2013; Longchamp & Gilliéron, 2009).

En ce qui concerne l'utilisation de l'épreuve à des fins internes, plusieurs études montrent la tendance à mobiliser ou à concevoir du matériel similaire ou identique au contenu de l'épreuve (Amrein-Beardsley, Berliner & Rideau, 2010; Pedulla, Abrams, Madaus *et al.*, 2003; Smith, 1991). Valli et Buese (2007) constatent par exemple dans une étude longitudinale que l'achat de manuels scolaires dont le format et la présentation sont similaires aux épreuves standardisées est une pratique courante dans les établissements du Maryland. À nouveau toutefois, il semble que le phénomène ne se cantonne pas aux dispositifs d'évaluation externe à forts enjeux. En ce sens, des entretiens réalisés en Belgique francophone auprès d'enseignants du primaire et du secondaire par Rosenwajn (à paraître) permettent d'établir un constat similaire quant à l'utilisation des épreuves externes à des fins internes.

Mais si les tendances au *teaching to the test* illustrent l'influence manifeste des épreuves externes sur les contenus enseignés, ces dernières semblent moins à même de stimuler l'innovation pédagogique. Parmi les études recensées par Au (2007), la grande majorité

(84 %) de celles qui rapportaient des changements dans les méthodes pédagogiques indiquaient le renforcement d'une pédagogie « traditionnelle » centrée sur la transmission de savoirs de l'enseignant vers l'élève. Opérant la distinction entre les contenus (ce que les enseignants enseignent) et les méthodes (comment les enseignants enseignent), Diamond montre par exemple que les enseignants de l'État de Chicago ayant participé à son étude étaient enclins à reconnaître que « l'évaluation et les standards de performances influencent le contenu qu'ils enseignent par contraste avec leurs méthodes » (2007, p.305). Un constat similaire a été posé pour les épreuves externes à faibles enjeux en Europe, où Dierendonck et Fagnant (2010) rapportent qu'une minorité d'enseignants (29 %) déclarent que les épreuves standardisées ont eu de réelles retombées sur leur travail quotidien.

Sur la base des tendances majeures observées dans la littérature relative aux pratiques enseignantes en contexte d'épreuves externes, nous distinguerons donc dans le cadre de cette étude une catégorie générale relative à l'intensité des changements pédagogiques opérés par les enseignants en réponse aux épreuves externes, et deux catégories de pratiques plus spécifiquement portées sur les changements opérés sur les contenus au travers de l'alignement curriculaire et de l'utilisation instrumentale de l'épreuve à des fins internes.

### L'utilisation des résultats

Si les études présentées précédemment décrivent les changements de pratiques opérés par des enseignants confrontés aux épreuves externes, elles laissent cependant ouverte la question de l'usage des résultats et de leur rôle éventuel dans l'orientation des pratiques. Dans quelle mesure la modification des pratiques enseignantes est-elle liée, tel que le soutiennent certains discours normatifs sur l'épreuve externe, à l'interprétation des résultats des élèves ? C'est en particulier dans les contextes des dispositifs d'évaluation externe à faibles enjeux, où la négligence des enseignants à l'égard des résultats apparaît comme une tendance dominante, que cette question a été soulevée par la recherche (Maier, 2009; Schildkamp & Kuiper, 2010; Vanhoof, Verhaeghe & Van Petegem, 2012; Verdière, 2013). Hellrung et Hartig (2013) soutiennent dans leur revue de la littérature sur l'utilisation des feedback en milieu scolaire que jusqu'à 90 % des enseignants ne font pas usage des données issues des épreuves en

langue et en mathématiques auxquelles ils sont confrontés. Plus spécifiquement, Vanhoof, Verhaege et Van Petegem (2012) constatent dans leur étude que rares sont les enseignants en Communauté flamande de Belgique qui mobilisent les résultats pour une réflexion critique à l'égard de leurs pratiques en classe. Dans le même ordre d'idées, Verdière a montré sur la base d'entretiens réalisés en région lilloise que, pour « la quasi-totalité des enseignants [...] rencontrés, [les] indicateurs statistiques permettant l'évaluation des élèves ou des établissements sont connus, mais ont peu de liens concrets avec leurs pratiques professionnelles » (2013, p. 83).

À ce stade, donc, deux constats majeurs peuvent être soulevés à partir de la littérature relative à l'impact de l'évaluation externe sur les pratiques enseignantes et de celle relative à l'utilisation des résultats. D'une part, si les épreuves externes semblent avoir un impact pédagogique, celui-ci est plutôt limité aux contenus enseignés. D'autre part, les résultats issus des épreuves externes paraissent peu déterminants dans la régulation des pratiques enseignantes. Ces constats nous invitent à examiner la question suivante : comment expliquer les effets de l'évaluation externe sur les contenus enseignés s'ils ne semblent pas déterminés par les résultats de l'épreuve ? Et plus généralement, comment rendre compte de la relative incapacité des résultats à influencer les pratiques enseignantes ? Quelles que soient les réponses à ces questions, nous suggérons qu'elles demandent une meilleure compréhension des mécanismes cognitifs impliqués dans le traitement des informations liées aux épreuves.

## Épreuves externes et traitement de l'information

À cet égard, si l'on suit les discours normatifs entourant les dispositifs d'évaluation externe, il semble que l'enjeu de ceux-ci consiste à faire des résultats un feedback de performances – à savoir une information fournie par un agent externe relative à la performance d'un individu dans une tâche donnée (Kluger & DeNisi, 1996) – pour les enseignants. L'idée fondamentale qui réside derrière l'utilisation du feedback à des fins d'apprentissage repose sur la stimulation fournie par une dissonance négative, à savoir un feedback inférieur aux attentes préalables. Il est attendu de l'individu confronté à une telle situation qu'il adapte son comportement de façon à réduire l'écart entre les résultats obtenus et les objectifs de performances qu'il s'est préalablement fixés.

Pourtant, il est apparu que tout feedback négatif ne menait pas systématiquement à une modification du comportement, et ce, principalement en raison de l'interprétation qui en est faite. Dans le cadre particulier du feedback constitué par les résultats de l'évaluation externe, les indications fournies par Bickman et Riemer (2011) nous mènent à envisager deux mécanismes importants : la validité perçue du feedback et l'attribution causale. En effet, avant que les résultats de l'épreuve externe puissent constituer un feedback et exercer une influence éventuelle sur les pratiques, il faut qu'ils soient perçus comme tels, que l'enseignant s'estime être en mesure de les modifier. C'est en ce sens que l'attribution causale est un élément déterminant dans l'interprétation des résultats de l'évaluation externe, et plus particulièrement les attributions causales internes dans lesquelles l'enseignant reconnaît une part de responsabilité personnelle dans la production des résultats tout en estimant être en mesure d'influencer ces causes par son comportement.

Remarquons aussi que les résultats de l'évaluation externe constituent potentiellement un feedback relatif à l'identité professionnelle et donc susceptible de porter atteinte à l'image de soi. Leur interprétation risque donc d'être soumise à l'interférence de facteurs motivationnels (Jordan & Audia, 2012). Le rejet de la validité des résultats ou de la responsabilité personnelle pourrait en effet permettre aux enseignants de préserver une image de soi favorable. Ceci est d'autant plus plausible dans le cas du feedback constitué par les résultats des élèves, puisqu'il s'agit d'une variable particulièrement ambiguë quant à ses causes. Il est dès lors aisé pour l'enseignant de recourir à des attributions causales externes minimisant sa part de responsabilité personnelle dans la production des résultats, pour préserver sa propre image.

Les théories relatives aux feedback de performances reposent cependant en grande partie sur des indications fournies par des études expérimentales très contrôlées qui diffèrent radicalement du contexte dans lequel un enseignant reçoit les résultats de l'épreuve externe. Dans l'idée de dépasser cette limite, nous avons estimé pertinent de prendre en compte une série de facteurs supplémentaires en s'inspirant de la théorie du comportement planifié (Ajzen, 1991), dont l'enjeu est la prédiction du comportement humain en milieu « naturel ». Outre l'attribution causale, cette dernière insiste sur l'intervention de facteurs sociaux et attitudeux dans la planification du comportement. C'est ainsi que nous avons également pris

en considération l'attitude des enseignants à l'égard de l'épreuve externe et l'influence sociale exercée par l'entourage professionnel direct des enseignants (collègues, direction).

Bien que n'ayant pas été conceptualisée comme telle, l'attitude semble avoir retenu l'attention d'un certain nombre d'études cherchant à comprendre la façon dont les enseignants perçoivent les épreuves externes (Dierendonck, 2008; Dierendonck & Fagnant, 2010; Lafontaine, Soussi & Nidegger, 2009; Maier, 2009, 2010). Considérée comme une « tendance psychologique qui s'exprime par l'évaluation favorable ou défavorable d'une entité particulière » (Eagly & Chaiken, 1998, p. 269), l'attitude (vis-à-vis de l'évaluation externe) risque d'être d'autant plus déterminante que l'autonomie qui caractérise le métier d'enseignant fait en sorte que l'assentiment de ces derniers soit une condition nécessaire à l'efficacité des réformes en milieu scolaire (Dupriez, 2015).

Quant à l'influence sociale, l'hypothèse est que l'interprétation du feedback se déroule dans un environnement où les interactions inévitables, plus ou moins fréquentes, entre les acteurs qui composent l'établissement scolaire, sont marquées par des jeux d'influence sociale (Cialdini & Trost, 1998). En particulier, la position qu'entretiennent les directions et collègues de l'établissement à l'égard de l'épreuve externe, telle que perçue par un enseignant déterminé, n'est sans doute pas sans influence sur l'interprétation des résultats posée par ce dernier et/ou sur son comportement.

En somme, nous nous attendons à ce que les résultats des évaluations externes influencent les pratiques enseignantes uniquement dans certaines conditions : la perception d'une dissonance négative, la reconnaissance que les résultats des élèves constituent un feedback personnel, et d'une marge de liberté pour les influencer. Nous suggérons aussi qu'un enseignant sera d'autant plus enclin à modifier son comportement qu'il entretient une attitude favorable à l'égard de l'épreuve externe et qu'il perçoit que certains acteurs de son entourage professionnel y accordent également de la valeur. Enfin, eu égard à la difficile prédiction du comportement à partir du traitement de l'information (Ajzen & Fishbein, 2005), nous nous intéressons non seulement aux changements effectifs, mais aussi aux intentions de changement.

## Méthode

Afin de répondre à nos questions de recherche qui, rappelons-le, portent d'une part sur les changements de pratiques suscités par l'évaluation externe et d'autre part sur le lien entre l'interprétation des résultats des élèves et ces changements, nous nous sommes orientés vers une étude par questionnaire menée en Communauté française de Belgique, qui a récemment imposé une épreuve externe certificative standardisée (le CE1D) à la fin du premier cycle du secondaire (13-14 ans).

La raison pour laquelle nous nous limitons à cette épreuve régionale en particulier se justifie tout d'abord par le fait que le système scolaire belge, pour des raisons qui sortent du cadre de cet article, est scindé sur la base communautaire depuis 1989<sup>1</sup>. Le développement des épreuves externes ne s'est donc pas fait en Belgique au niveau national, mais au niveau communautaire. En outre, notre choix d'analyser cette épreuve en particulier se justifie par le fait que les résultats du CE1D sont généralement plus variables et moins favorables, ce qui permet de supposer une diversité plus importante de réactions des enseignants qui y sont confrontés.

Notons aussi que l'épreuve certificative CE1D constitue bien un enjeu important pour les élèves puisque sa réussite détermine leur passage au niveau supérieur, mais qu'aucune conséquence formelle n'est associée aux résultats pour les enseignants et les établissements. Pour ces derniers donc, le CE1D apparaît comme une épreuve à faibles enjeux qui n'est pas même assortie d'une obligation d'utilisation au-delà de la certification des élèves. Les documents parlementaires relatifs au décret qui encadre ces épreuves et les interventions publiques des ministres de l'Enseignement permettent toutefois de soutenir l'idée que l'enjeu implicite de ces épreuves est également de contribuer à la régulation des pratiques enseignantes<sup>2</sup>.

1 Pour de plus amples développements sur ce point, nous renvoyons au Cahier du CRISP intitulé *Le système scolaire* (Draelants, Dupriez & Maroy, 2011).

2 Nous pointons notamment la déclaration de la ministre de l'Enseignement de la Communauté française, Mrs Schyns, qui déclarait le 29 août 2016 à l'antenne de la RTBF que « ces épreuves externes, elles ont quand même pour objectif de permettre à la fois aux élèves, mais aussi aux enseignants de s'évaluer. D'évaluer leur pratique par rapport aux attendus. »

## Procédure

La population cible de notre étude était constituée de l'ensemble des enseignants confrontés aux épreuves certificatives de la fin du premier degré de l'enseignement secondaire en Belgique francophone. C'est ainsi que les 472 établissements secondaires recensés en Communauté française furent contactés par téléphone entre avril et juin de l'année scolaire 2014-2015. En cas de liaison établie avec la direction de l'établissement, il lui était proposé de contribuer à l'étude en invitant les enseignants de son équipe éducative à compléter le questionnaire en ligne. Un courrier électronique d'invitation à participer à l'étude contenant le lien vers le questionnaire était alors envoyé aux directions. Par cette procédure, 244 directions reçurent un mail d'invitation à l'étude. Parmi celles-ci, un minimum<sup>3</sup> de 96 furent représentées dans les 253 réponses complètes au questionnaire que nous avons été en mesure de récolter.

## Échantillon

Le taux de représentation des établissements de la population dans notre échantillon est donc de 20,3 %. Du côté des enseignants, le taux de réponse au sein des établissements visés est lui aussi faible, bien qu'il soit difficile d'établir la population de référence par établissement. Notre étude est toutefois loin d'être la seule à connaître un taux de réponse très bas, et ce particulièrement dans les contextes d'évaluation externe à faibles enjeux (Amrein-Beardsley, Berliner & Rideau, 2010; Dierendonck & Fagnant, 2010; Maier, 2009). On peut toutefois considérer que ce faible taux de réponse est moins problématique dès lors que les répondants sont représentatifs de la population de référence, ce qui semble être largement le cas pour toute une série d'indicateurs socio-biographiques institutionnels (voir tableau 1), et qu'on constate de la variance dans les variables dépendantes (voir tableaux 2 et 3).

Sur les 253 enseignants ayant entièrement complété le questionnaire, la balance de genres est en faveur des femmes (70,4 %). L'âge des participants se situe entre 22 et 62 ans, avec une moyenne de 39 ans. La plupart des enseignants de notre échantillon sont des personnels nommés à temps plein (71,1 %), le reste étant composé d'enseignants nommés à temps partiel

(8,3 %) et d'enseignants temporaires (20,6 %). L'expérience professionnelle des enseignants se distribue quant à elle de la manière suivante : 44,7 % des enseignants déclarent moins de 10 années d'expérience dans l'enseignement, 24,1 % se situent entre 11 et 20 années d'expérience, et 31,2 % dépassent les 20 années d'expérience. La majorité des enseignants exercent dans un établissement du réseau libre confessionnel catholique (58,5 %). Pour le reste, 25,7 % des enseignants travaillent dans un établissement du réseau officiel, 14,6 % dans un établissement du réseau officiel subventionné et 1,2 % dans un établissement du réseau libre non confessionnel. Cette répartition semble bien refléter le poids des différents réseaux de l'ensemble de la Communauté française de Belgique dans l'enseignement secondaire.

La dénomination exacte de l'établissement au sein duquel l'enseignant exerce était présentée comme une question non obligatoire. Ainsi, seuls 210 enseignants ont accepté de fournir cette information, ce qui nous a permis de distinguer 96 établissements différents représentés dans notre échantillon. Pour ceux-ci, nous avons été en mesure de repérer les indices ISE<sup>4</sup> qui leur étaient attribués et de constater que 64 des enseignants de l'échantillon exercent dans un établissement classé entre 1 et 5 dans l'indice; 39 dans un établissement classé entre 6 et 10; 43 dans un établissement classé entre 11 et 15; et 63 dans une école classée entre 16 et 20. La grande majorité des enseignants de l'échantillon travaillent dans la filière générale de l'enseignement secondaire (84,6 %), ce qui est parfaitement cohérent avec le fait que le niveau d'enseignement visé constitue la dernière étape du tronc commun. Le tableau 1 résume les caractéristiques principales de notre échantillon, ainsi que le test de sa représentativité sur les variables présentées.

En ce qui concerne les matières enseignées, 36,8 % des enseignants sont chargés du cours de français, 33,6 % du cours de mathématiques, 22,5 % du cours de langues modernes et 7,1 % du cours de sciences. Les quatre matières testées dans l'évaluation externe sont donc représentées dans l'échantillon, avec une dominante pour les deux matières principales.

4 Indice socio-économique. Cet indicateur officiel opère une catégorisation des établissements en fonction du niveau socio-économique du public accueilli par ordre croissant sur une échelle de 20 points.

3 Certains enseignants ont refusé d'indiquer leur établissement d'appartenance.



**Tableau 1. Caractéristiques de l'échantillon d'enseignants et test de sa représentativité**

	Population de référence	Échantillon (n=253)	Représentativité
Sexe	F : 63 ; H : 37	F : 70,4 ; H : 29,6	$\chi^2_{(1)} = 2,102, ns$
Âge moyen	M= 38,4	M= 39,4	$t_{252} = 1,450, ns$
Statut professionnel	Nommés temps plein : 72,6 Nommés temps partiel : 5,7 Temporaires : 21,7	Nommés temps plein : 71,1 Nommés temps partiel : 8,3 Temporaires : 20,6	$\chi^2_{(2)} = 0,986, ns$
Ancienneté professionnelle	Moins de 10 ans : 52,5 Entre 11 et 20 ans : 21,8 Plus de 20 ans : 25,7	Moins de 10 ans : 44,7 Entre 11 et 20 ans : 24,1 Plus de 20 ans : 31,2	$\chi^2_{(2)} = 2,386, ns$
Réseau d'enseignement	Libre subventionné : 60,6 Officiel : 23,2 Officiel subventionné : 16,2	Libre subventionné : 58,5 Officiel : 25,7 Officiel subventionné : 14,6	$\chi^2_{(2)} = 0,469, ns$

## Variables

Cinq parties principales composaient notre questionnaire. Les quatre premières parties comprenaient respectivement une série d'échelles relatives 1) aux effets de l'évaluation externe sur les pratiques enseignantes; 2) à l'interprétation des résultats; 3) à l'attitude à l'égard de l'épreuve; 4) à l'influence sociale (voir en annexe pour la liste complète des items). Une dernière partie du questionnaire visait à récolter des indications personnelles, démographiques et contextuelles. L'ensemble des échelles étaient présentées sous la forme d'une affirmation associée à une échelle de Likert à 5 points où la valeur 1 était « pas du tout d'accord » et la valeur 5 « tout à fait d'accord ».

### Les pratiques enseignantes et l'évaluation externe

Un premier ensemble de variables était centré sur les changements de pratiques rapportés par les enseignants comme conséquences de l'évaluation externe. La structure des échelles de mesure a été conçue en fonction d'une triple dimension : 1) les changements relatifs au *teaching to the test*, comprenant la distinction entre l'utilisation instrumentale de l'épreuve à des fins internes d'une part et les changements opérés sur les contenus enseignés d'autre part; 2) le sentiment relatif à l'intensité des changements liés aux pratiques pédagogiques, considéré ici comme un indicateur de l'innovation pédagogique; et 3) l'intention de changement (pédagogique). Une analyse en composantes principales<sup>5</sup> avec rotation Oblimin des 13 items, ne

retenant que les composantes disposant de valeurs propres supérieures à 1, a montré une structure à 4 facteurs qui correspond bien à la structure théorique supposée, expliquant 67 % de la variance totale. La première dimension décrit les changements dans les méthodes pédagogiques (3 items,  $\alpha = 0,78$ ); la seconde, l'utilisation des tests eux-mêmes à des fins instrumentales (3 items,  $\alpha = 0,78$ ); la troisième, l'intention au changement (4 items,  $\alpha = 0,78$ ); et la quatrième, l'alignement curriculaire (3 items,  $\alpha = 0,77$ )<sup>6</sup>.

### L'interprétation des résultats

La deuxième partie du questionnaire était centrée sur la façon dont les enseignants font sens des résultats issus de l'épreuve externe. Un premier item de cette section déterminait le pourcentage moyen de réussite des élèves à la dernière passation de l'épreuve. Celui-ci était complété par une série de trois échelles destinées à mesurer 1) la dissonance, positive ou négative, générée par les résultats de l'épreuve, 2) la validité perçue de ces derniers pour mesurer les performances des élèves, mais aussi des enseignants, et 3) l'attribution causale interne, en termes de responsabilité personnelle perçue à l'égard des résultats. Une dernière échelle tentait de

étude. La rotation de type Oblimin est justifiée par la supposition de possibles corrélations entre les facteurs.

<sup>6</sup> Un item mesurant l'alignement curriculaire sature plus fortement sur l'échelle de changements pédagogiques (voir coefficients de saturation en annexe), et un autre présente un *cross-loading* avec la même échelle. Pour des raisons de fiabilité de la mesure de l'échelle d'alignement curriculaire, nous avons décidé d'inclure ces deux items dans cette échelle, mais cela peut avoir pour effet de surestimer la corrélation entre les échelles de changements pédagogiques et d'alignement curriculaire.

<sup>5</sup> Une analyse de type exploratoire a ici été réalisée, car la plupart des échelles étaient des créations originales développées pour cette



mesurer l'importance accordée aux résultats de l'épreuve externe de manière générale. Une analyse en composantes principales avec rotation Oblimin sur les 17 items composant ces différentes échelles a permis la distinction de 5 facteurs expliquant 66 % de la variance totale. La première dimension mesurait très clairement l'importance que les enseignants accordent aux résultats issus de l'épreuve externe (5 items,  $\alpha=0,84$ ), tandis que la deuxième regroupait les indicateurs relatifs à la dissonance négative (2 items,  $\alpha=0,79$ ). La troisième dimension (4 items,  $\alpha=0,81$ ) regroupait pour sa part les items relatifs à la validité perçue des résultats comme instrument de mesure des compétences des élèves et ceux relatifs à la dissonance positive (c'est-à-dire des résultats plus élevés par rapport aux attentes). De façon assez intéressante, la configuration de ces deux dernières dimensions semble indiquer que le rejet de la validité des résultats comme mesure des compétences des élèves est essentiellement lié à la dissonance positive. La quatrième dimension détermine la part de responsabilité personnelle reconnue par les enseignants dans la production des résultats, c'est-à-dire l'attribution causale interne. La fiabilité des 2 items qui composaient cette échelle est cependant insatisfaisante (2 items,  $\alpha=0,55$ ). Enfin, la cinquième dimension distinguait nettement les items destinés à déterminer la validité perçue des résultats comme mesure des compétences de l'enseignant (4 items,  $\alpha=0,68$ ).

### Attitude relative à l'épreuve externe

Un troisième ensemble d'échelles mesurait la façon dont les enseignants perçoivent l'épreuve externe en distinguant l'attitude par rapport à celle-ci et l'importance qu'ils lui accordent dans leur travail quotidien. En effet, de nombreux auteurs (Eagly & Chaiken, 1998) insistent sur le fait que l'influence de l'attitude proprement dite sur le comportement dépend également de l'importance accordée à l'objet de cette attitude.

L'analyse par composantes principales avec rotation Oblimin réalisée sur les 10 items indique 2 composantes expliquant 67 % de la variance totale. Conformément aux attentes, la première composante

regroupe les items relatifs à l'importance de l'épreuve (5 items,  $\alpha=0,85$ ), tandis que la deuxième composante regroupe l'ensemble des items mesurant l'attitude de l'enseignant à l'égard des épreuves externes (5 items,  $\alpha=0,89$ ). Un seul item présente une saturation croisée.

Le tableau 2 présente les corrélations entre les dimensions relatives à l'attitude d'une part, et à l'interprétation des résultats d'autre part, et permet au lecteur de se faire une première idée de l'état des relations qui sont au cœur de notre seconde question de recherche.

### L'influence sociale

Pour tenter de mesurer le poids de l'influence sociale, deux dernières échelles visaient à déterminer la perception que l'enseignant a de l'importance accordée par ses collègues à l'épreuve externe (2 items,  $\alpha=0,60$ ), ou par sa direction (6 items,  $\alpha=0,85$ ). L'analyse en composantes principales avec rotation Oblimin confirme les deux dimensions tout en expliquant 61 % de la variance totale.

### Données personnelles, contextuelles et démographiques

Enfin, les enseignants étaient interrogés quant à leur âge et leur sexe, mais aussi quant à leur parcours de formation et professionnel : la formation initiale, l'ancienneté, le statut professionnel, la filière d'enseignement, la/les matières enseignées. Il était également demandé dans cette section, à titre facultatif, d'indiquer le nom de l'établissement au sein duquel l'enseignant exerçait son métier. Des analyses de covariance ont permis de montrer qu'aucune des données personnelles, contextuelles et démographiques n'est liée aux principales variables dépendantes : qu'il s'agisse de l'intention au changement, du sentiment de changement ou des changements effectifs opérés.

## Analyses

Les données récoltées ont été soumises à une analyse descriptive, puis à deux types d'analyse inférentielle. Pour répondre à notre première question de recherche, nous avons comparé entre eux les niveaux de change-

7 Eu égard à la centralité de cette variable dans la théorie du feedback de performance, nous la décrivons et l'utilisons comme prédicteur dans les analyses de régression multiple. Nous l'écartons par contre pour les analyses en pistes causales, étant donné sa trop faible fiabilité. Notons aussi que sa faible fiabilité tient avant tout au petit nombre d'items (2), qui sont par ailleurs très largement corrélés à la composante ( $r > 0,80$ ).

8 Bien que cette échelle ait une consistance interne faible, même acceptable, particulièrement au regard du faible nombre d'items qui la compose (2), nous la conservons pour les analyses descriptives et inférentielles, car elle joue un rôle central dans la modélisation théorique liée à la théorie du comportement planifié.

ment de pratiques rapportés par les enseignants. Pour répondre à la seconde, des régressions multiples<sup>9</sup> par bloc de variables ont été menées pour prédire les intentions de changement, le changement pédagogique, l'utilisation stratégique de l'épreuve à des fins internes et l'alignement curriculaire, puis des analyses en pistes causales pour modéliser les « chemins » par lesquels l'évaluation externe des élèves semble reliée aux changements de pratiques rapportés par les enseignants<sup>10</sup>.

## Analyses descriptives

### Les pratiques enseignantes et l'évaluation externe

Quelles pratiques les enseignants déclarent-ils modifier en lien avec l'évaluation externe ? Les changements les plus fréquents sont les changements de type instrumental (Moyenne [M] = 3,48 ; Médiane [Md] = 3,67 ; Écart-type [DS] = 0,89) suivi par l'alignement curriculaire (M = 3,34 ; Md = 3,33 ; DS = 0,93), puis par l'intention de changement (M = 3,01 ; Md = 3 ; DS = 0,91), et enfin le sentiment général de changement, lui, sous le point central de l'échelle (M = 2,41 ; Md = 2,33 ; DS = 0,98).

### L'interprétation des résultats

En matière d'interprétation des résultats (voir tableau 2), soulignons d'abord que les enseignants ont plutôt tendance à considérer que les résultats ne sont pas une mesure appropriée des compétences de leurs élèves (M = 2,69 ; Md = 2,75 ; DS = 0,99) et encore moins des leurs (M = 1,98 ; Md = 2 ; DS = 0,74). Ils ont également tendance à faiblement attribuer les performances de leurs élèves à l'interne (M = 2,05 ; Md = 2 ; DS = 0,8). En

dépît de ces constats, la plupart des enseignants considèrent que les résultats sont un élément relativement important dans leur travail quotidien (M = 3,17 ; Md = 3,4 ; DS = 0,96).

Il convient ici de mentionner que la moyenne des scores des élèves rapportés par les enseignants était plutôt positive, avec une grande majorité des enseignants déclarant des scores moyens de leurs élèves au-dessus du seuil de réussite (M = 64,19 % ; DS = 25,63). De manière cohérente, la plupart des enseignants rapportent peu de dissonance négative entre les résultats espérés et obtenus (M = 2,51 ; Md = 2,5 ; DS = 1,11). La variabilité sur ces deux échelles (de score et de dissonance) est toutefois importante.

### Attitude relative à l'épreuve externe

L'analyse de l'échelle destinée à mesurer l'attitude de l'enseignant à l'égard de l'évaluation externe permet de constater (voir tableau 2) que les enseignants adoptent une position plutôt favorable aux épreuves standardisées (M = 3,39 ; Md = 3,6 ; DS = 1). Pour autant, l'attitude plutôt favorable des enseignants ne se reflète pas dans l'importance qu'ils accordent à l'utilisation des épreuves externes. L'échelle destinée à mesurer cette dimension indique une moyenne légèrement inférieure au point central de l'échelle (M = 2,41 ; Md = 2,4 ; DS = 0,88), avec un nombre relativement important d'enseignants accordant une importance tout à fait négligeable aux épreuves.

### L'influence sociale

Une majorité d'enseignants estiment enfin que leur direction accorde peu d'importance aux épreuves externes (M = 2,56 ; Md = 2,5 ; DS = 0,86). Ils sont en règle générale plus nombreux à considérer que leurs collègues accordent une certaine importance aux épreuves, bien que les valeurs de cette échelle se centrent sur le point central (M = 2,97 ; Md = 3 ; DS = 0,89).

## Analyses inférentielles

### Les pratiques enseignantes

Venons-en aux résultats liés à notre première question de recherche. Une série de tests t pour échantillons pairs indique que toutes les paires de pratiques considérées se différencient significativement : utilisation instrumentale et alignement du curriculaire ( $t_{252} = 2,193$ ,  $p < 0,05$ ), alignement du curriculaire et changement pédagogique ( $t_{252} = 16,389$ ,  $p < 0,01$ ), et changement pédagogique et utilisation instrumentale ( $t_{252} = 15,273$ ,

9 La nature multi-niveau des données enseignantes a été ignorée dans les analyses de prédiction pour trois raisons. Premièrement, le nombre d'enseignants par établissement est le plus souvent très faible, avec pour risque de sous-pondérer la place de ces données dans les analyses par la procédure de *shrinkage*. Deuxièmement, la variable établissement connaît un taux de réponses manquantes plus important ( $n = 210$ ). Troisièmement, une analyse de variance simple avec l'établissement pour effet aléatoire menée sur nos données indique que pour les principales variables dépendantes de notre étude, la part de variance inter-établissements ne dépasse jamais 3 %.

10 À la différence d'analyses par équations structurelles complètes, les analyses en pistes causales testent uniquement des modèles de prédiction, et non pas des modèles de mesure. Tester les deux simultanément exige un échantillon de taille importante, ce qui n'est pas le cas dans la présente étude. Les analyses descriptives et de régression multiple ont été réalisées dans SPSS 23, et celles en pistes causales dans Mplus7. De manière à augmenter la fiabilité des analyses de médiation en Mplus, un *bootstrap* ( $= 5000$ ) a également été appliqué. Cette technique permet de produire les estimations des modèles sur de multiples sous-échantillons tirés au hasard à partir de l'échantillon principal, et ainsi de mieux estimer les risques d'erreur dans les tests d'hypothèse.

Tableau 2. Indications descriptives et corrélations entre les échelles de changement de pratiques

	Moyenne	Écart-type	Changement pédagogique	Utilisation instrumentale	Alignement curriculaire	Intention de changement
Changement pédagogique	2,41	0,98	1			
Utilisation instrumentale	3,48	0,89	0,172**	1		
Alignement curriculaire	3,34	0,93	0,361**	0,274	1	
Intention de changement	3,01	0,91	0,267**	0,144*	0,305**	1

Note : \* : < 0,05 ; \*\* : < 0,01.

Tableau 3. Indications descriptives et corrélations entre les échelles d'attitude et d'interprétation des résultats

	Moyenne	Écart-type	Attitude générale	Importance évaluation	Importance résultats	Dissonance négative	Validité : Élèves	Attribution interne	Validité : Enseignant
Attitude générale	3,39	1	1						
Importance évaluation	2,41	0,88	0,499**	1					
Importance résultats	3,17	0,96	0,444**	0,620**	1				
Dissonance négative	2,51	1,11	0,007	0,001	0,042	1			
Validité : élèves	2,67	0,99	0,417**	0,570**	0,353**	0,169**	1		
Attribution interne	2,05	0,8	0,176*	0,249**	0,200**	0,063	0,202**	1	
Validité : enseignant	1,98	0,74	0,243**	0,342**	0,204**	0,149*	0,180**	0,256**	1

Note : \* : < 0,05 ; \*\* : < 0,01.

$p < 0,01$ ). Ainsi, pour les enseignants de notre échantillon, et conformément à la littérature, le changement dans les méthodes pédagogiques est le type de changement le moins souvent rapporté par les enseignants. Les changements curriculaire sont plus courants, et plus encore les changements instrumentaux.

### L'interprétation des résultats et les changements de pratiques

Qu'en est-il de la contribution des résultats des élèves à ces changements, mais aussi à l'intention de changer? Les analyses de régressions multiples montrent que seuls les changements de nature instrumentale sont liés de manière significative aux scores à

hauteur de 6 % de variance expliquée. Plus précisément, l'utilisation instrumentale de l'épreuve est associée avec le pourcentage moyen de réussite des élèves à l'épreuve de l'année précédente ( $\beta = -0,275, p < 0,01$ ). Plus les élèves d'un enseignant donné ont de mauvais résultats, plus ce dernier tend à utiliser le matériel des évaluations externes pour ses propres évaluations ou comme bases d'exercices pour ses élèves. Dit autrement, les scores des élèves ne sont pas du tout liés aux changements de méthodes pédagogiques ou à l'alignement du curriculum.

Ensuite, nous avons cherché à prédire les changements en fonction des dimensions relatives à l'interprétation des résultats, en rentrant simultanément

comme prédicteurs la perception de l'importance des résultats, de la dissonance négative, de la validité des résultats (pour mesurer les compétences des enseignants et des élèves) et de l'attribution causale interne des résultats des élèves. Pour la prédiction des changements de nature instrumentale, ces variables ont été considérées en plus de l'influence du pourcentage moyen de réussite des élèves. Il en ressort que la centralité des résultats pour les enseignants est associée de manière significative à l'intention de changement ( $\beta=0,308, p < 0,01$ ), aux changements pédagogiques ( $\beta=0,185, p < 0,01$ ), à l'alignement curriculaire ( $\beta=0,276, p < 0,01$ ), mais aussi aux changements instrumentaux ( $\beta=0,174, p < 0,01$ ). La perception de la validité des résultats par rapport aux élèves est liée seulement à l'intention de changement ( $\beta=0,218, p < 0,01$ ) et aux changements instrumentaux ( $\beta=0,223, p < 0,01$ ), tandis que l'attribution causale interne l'est uniquement avec l'intention de changement ( $\beta=0,158, p < 0,01$ ). La dissonance négative et la validité perçue des résultats à l'égard des enseignants ne sont liées ni à l'intention ni aux changements effectifs.

Au total, ces diverses facettes du processus d'interprétation des résultats sont associées avec 28 % de la variance de l'intention de changement, et respectivement, 9, 15 et 8 % des changements pédagogiques, instrumentaux et de l'alignement curriculaire. Notons aussi qu'après contrôle par les dimensions cognitives liées à l'interprétation des résultats, l'effet du pourcentage moyen de réussite cesse d'être significatif dans l'explication des effets instrumentaux.

### L'attitude à l'égard de l'évaluation externe et les changements de pratiques

Dans la suite des analyses de régression, nous avons mis en relation l'attitude à l'égard de l'évaluation externe et les changements de pratiques. Ici, les résultats des régressions multiples soulignent que l'attitude générale à l'égard de l'évaluation externe prédit de manière significative l'intention de changement ( $\beta=0,142, p < 0,05$ ) ainsi que l'alignement curriculaire ( $\beta=0,210, p < 0,01$ ). L'importance perçue de l'évaluation est quant à elle associée significativement au changement pédagogique ( $\beta=0,205, p < 0,01$ ), à l'utilisation instrumentale de l'épreuve ( $\beta=0,303, p < 0,01$ ) et à l'intention de changement ( $\beta=0,456, p < 0,01$ ). Les deux échelles expliquent respectivement 29, 9, 12 et 4 % de la variance de l'intention de changement, du changement pédagogique, de l'utilisation instrumentale et de l'alignement curriculaire.

### L'influence sociale

Un dernier ensemble de régressions multiples a tenté de mettre en évidence le poids de l'influence sociale. Les résultats montrent que l'influence sociale exercée par les collègues autour de l'évaluation externe est significativement liée à l'intention de changement ( $\beta=0,425, p < 0,01$ ), au changement pédagogique ( $\beta=0,223, p < 0,01$ ), à l'utilisation instrumentale de l'épreuve ( $\beta=0,317, p < 0,01$ ) et à l'alignement curriculaire ( $\beta=0,258, p < 0,01$ ). L'influence sociale de la direction est par contre uniquement associée à l'intention de changement ( $\beta=0,272, p < 0,01$ ). Les deux échelles d'influence sociale sont liées respectivement à 32, 5, 10 et 8 % de la variance de l'intention de changement, du changement pédagogique, de l'utilisation instrumentale et de l'alignement curriculaire.

### Analyses en pistes causales

Nous terminons en présentant les résultats des analyses en pistes causales, dont l'intérêt majeur est de comparer le pouvoir prédictif de diverses modélisations du processus de changement des pratiques enseignantes liées à l'évaluation externe. Quatre modèles principaux ont été comparés. Les deux premiers modèles sont directement dérivés de la théorie du feedback de performance (Bickman & Riemer, 2011 ; Kluger & DeNisi, 1996) : le premier teste à quel point les changements rapportés par les enseignants dans leurs pratiques sont liés aux résultats des élèves tandis que le second ajoute aux résultats eux-mêmes les facteurs liés à leur interprétation (centralité des résultats pour l'enseignant, et reconnaissance de la validité des résultats pour mesurer les compétences des élèves ou celles des enseignants). Le troisième teste dans quelle mesure la théorie du comportement planifié (Ajzen, 1991) et le rôle central qu'elle donne aux intentions comportementales, aux attitudes et à l'influence sociale permettent d'expliquer les principaux changements rapportés par les enseignants. Le dernier modèle teste, quant à lui, si et comment la combinaison des deux perspectives théoriques (théories du feedback de performance et du comportement planifié) est pertinente pour rendre compte des modifications des pratiques enseignantes en lien avec l'évaluation externe.

Dans le premier modèle, nous avons testé si le changement pédagogique, l'utilisation instrumentale de l'épreuve et l'alignement curriculaire étaient prédits par le pourcentage moyen de réussite des élèves à la dernière passation de l'épreuve et par la dissonance

négative. Pour rappel, la dissonance négative renvoie à l'idée que les enseignants pourraient être surpris par les mauvais résultats de leurs élèves, et du coup entreprendre des démarches en vue de réduire l'écart entre les résultats espérés et les résultats effectifs des élèves. Les résultats indiquent que la dissonance négative, mécanisme supposé central dans la théorie du feedback de performance, ne prédit aucun des changements rapportés par les enseignants. Seul le pourcentage moyen de réussite des élèves à la dernière passation de l'épreuve prédit de manière significative l'utilisation instrumentale de l'épreuve ( $\beta = -0,271$ ,  $p < 0,01$ ). Le *fit* du modèle n'est pas satisfaisant ( $CFI = 0,318$ ;  $RMSEA = 0,179$ ;  $SRMR = 0,074$ ), venant par là confirmer que la régulation du comportement par la seule prise en compte des résultats des élèves et une éventuelle dissonance négative explique mal les transformations opérées par les enseignants dans leurs pratiques. Le modèle explique 5 % de la variance de l'utilisation instrumentale de l'épreuve, et 0 % pour le changement pédagogique et l'alignement curriculaire.

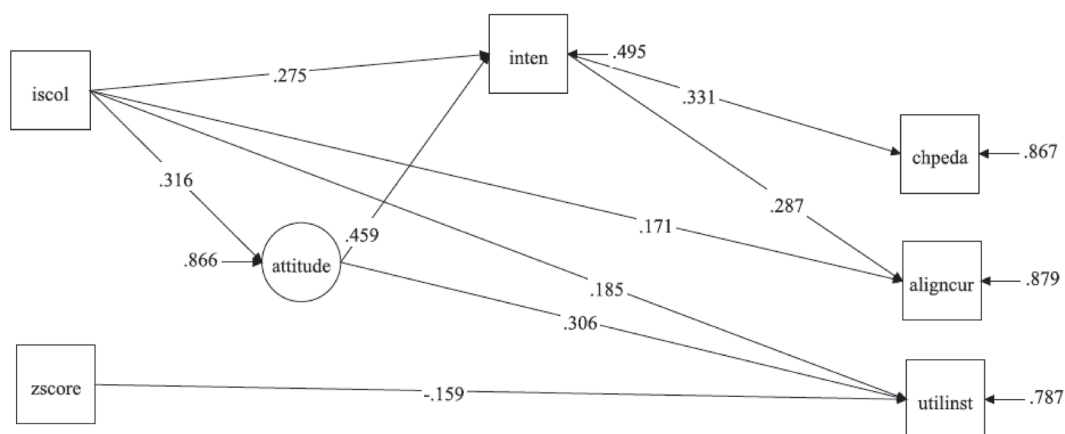
Dans le second modèle, au pourcentage moyen de réussite des élèves à la dernière passation de l'épreuve (prédisant l'utilisation instrumentale de l'épreuve) ont été ajoutés trois prédicteurs décrivant l'interprétation et le sens donné par les enseignants aux résultats (la centralité des résultats pour l'enseignant et la reconnaissance de la validité des résultats pour mesurer les compétences des élèves ou celles des enseignants). Les analyses montrent que la centralité perçue des résultats par les enseignants prédit de manière significative le changement pédagogique ( $\beta = 0,197$ ,  $p < 0,01$ ), l'alignement curriculaire ( $\beta = 0,274$ ,  $p < 0,01$ ) et l'utilisation instrumentale de l'épreuve ( $\beta = 0,178$ ,  $p < 0,01$ ), tandis que la reconnaissance de la validité des résultats pour mesurer les compétences des élèves prédit significativement le changement pédagogique ( $\beta = 0,149$ ,  $p < 0,05$ ) et l'utilisation instrumentale de l'épreuve ( $\beta = 0,227$ ,  $p < 0,01$ ). Ce modèle explique respectivement 14, 8 et 8 % de la variance de l'utilisation instrumentale de l'épreuve, du changement pédagogique et de l'alignement curriculaire. Les indices de *fit* du modèle s'améliorent mais n'atteignent pas les limites acceptées par la littérature ( $CFI = 0,880$ ;  $RMSEA = 0,088$ ;  $SRMR = 0,037$ ).

Le troisième modèle visait à mettre à l'épreuve la théorie du comportement planifié dans sa capacité à prédire les changements rapportés par les enseignants. Ceci signifie que l'intention de changement, l'attitude des enseignants envers l'évaluation externe et l'influence sociale des collègues et de la direction ont été

introduites comme prédicteurs des trois types de changement de pratiques, mais aussi que l'intention de changement a été posée comme médiateur de l'effet de l'influence sociale et de l'attitude sur les changements rapportés. Cette analyse de médiation permet de questionner dans quelle mesure l'environnement social et l'attitude des enseignants produisent un effet sur l'intention de changement qui elle-même, ensuite, expliquerait les changements de pratiques. Les résultats indiquent un *fit* légèrement meilleur de ce modèle ( $CFI = 0,963$ ;  $RMSEA = 0,083$ ;  $SRMR = 0,036$ ) par rapport au modèle 2. L'attitude prédit significativement les trois changements, de manière directe pour l'utilisation instrumentale de l'épreuve ( $\beta = 0,318$ ,  $p < 0,01$ ) et indirecte, à travers une médiation par l'intention de changement, pour le changement pédagogique ( $\beta = 0,192$ ,  $p < 0,01$ ) et l'alignement du curriculum ( $\beta = 0,224$ ,  $p < 0,01$ ). L'influence sociale des collègues uniquement (et non celle de la direction) prédit les changements de pratiques, soit de manière directe pour l'utilisation instrumentale de l'épreuve ( $\beta = 0,214$ ,  $p < 0,01$ ), soit de manière indirecte pour le changement pédagogique ( $\beta = 0,151$ ,  $p < 0,01$ ) et pour l'alignement curriculaire ( $\beta = 0,086$ ,  $p < 0,05$ ). Le *fit* du modèle est légèrement amélioré dès lors qu'on retire les effets directs l'influence sociale de la direction sur les changements des pratiques, et l'effet de l'intention comportementale sur l'utilisation instrumentale de l'épreuve ( $CFI = 0,965$ ;  $RMSEA = 0,066$ ;  $SRMR = 0,044$ ). Ce dernier modèle explique respectivement 50, 19, 13 et 12 % de la variance de l'intention de changement, de l'utilisation instrumentale de l'épreuve, du changement pédagogique et de l'alignement curriculaire.

Le dernier modèle visait à intégrer des prédicteurs liés aux résultats et à leur interprétation (dérivés de la théorie du feedback de performance), et à la théorie du comportement planifié. Ceci signifie qu'en plus de l'influence sociale des collègues, de l'attitude envers l'évaluation externe et de l'intention de changement, étaient considérés comme prédicteurs de l'utilisation instrumentale de l'épreuve, le pourcentage de réussite à l'épreuve précédente, la centralité des résultats et la reconnaissance de la validité de l'épreuve externe pour mesurer les compétences des élèves; comme prédicteurs du changement pédagogique, la centralité des résultats et la reconnaissance de la validité de l'épreuve externe pour mesurer les compétences des élèves; et comme prédicteur de l'alignement du curriculum, la centralité des résultats. Le *fit* de ce quatrième modèle n'est pas meilleur ( $CFI = 0,651$ ;  $RMSEA = 0,185$ ;  $SRMR = 0,129$ ). Ce mauvais *fit* s'explique par le fait que

Figure 1. Représentation schématique (Mplus7) des effets observés



Note : les coefficients sont des coefficients bêta, c'est-à-dire des coefficients standardisés. iscol : influence sociale des collègues; zscore : scores des élèves à la dernière épreuve externe; attitude : attitude à l'égard de l'évaluation externe; inten : intention de changement; chpedit : changement pédagogique; aligncur : alignement curriculaire; utilinst : utilisation instrumentale.

les variables liées à l'interprétation des résultats, à l'exception de leur centralité pour prédire l'alignement curriculaire, et aux résultats bruts pour prédire l'utilisation instrumentale de l'épreuve, n'ajoutent aucun pouvoir prédictif au modèle.

Le meilleur modèle est au final celui qui, en plus des variables d'influence sociale, d'attitude et d'intention comportementale, considère uniquement et également que les résultats bruts conditionnent l'utilisation instrumentale de l'épreuve (CFI = 0,972; RMSEA = 0,048; SRMR = 0,040). Ce modèle explique 50 % de la variance de l'intention de changement, 21 % de celle de l'utilisation instrumentale de l'épreuve, 13 % de celle du changement pédagogique, et 12 % de l'alignement curriculaire. Ce modèle est représenté dans la figure 1.

## Discussion

Notre objectif dans cet article était double : d'abord identifier les types de changement dans les pratiques enseignantes suscités par une évaluation externe à faibles enjeux, mais aussi de mieux comprendre les processus cognitifs en jeu dans les liens qui se tissent entre ce type de dispositifs et les pratiques enseignantes. En mobilisant les observations établies dans la littérature empirique internationale, nous nous sommes intéressés, d'une part, à l'intensité des chan-

gements relatifs à la pédagogie de manière générale; et d'autre part, plus spécifiquement, aux changements opérés par les enseignants sur les contenus enseignés au travers de l'utilisation instrumentale de l'épreuve comme source d'inspiration à des fins internes et de l'alignement curriculaire sur les exigences de l'épreuve. Les analyses que nous avons effectuées indiquent tout d'abord que, si les enseignants belges francophones rapportent avoir opéré des modifications sur les contenus enseignés à une certaine intensité, ils sont moins portés à considérer que les épreuves externes ont provoqué des changements majeurs dans leur façon d'enseigner. Ces résultats confirment les observations établies dans la littérature empirique selon lesquelles l'impact des dispositifs d'évaluation externe sur le changement pédagogique, et *a fortiori* sur l'innovation pédagogique, est relativement faible (Au, 2007; Diamond, 2007; Dierendonck & Fagnant, 2010); ainsi que les observations selon lesquelles les dispositifs d'évaluation externe à faibles enjeux ne seraient pas exempts d'une tendance à l'alignement de l'enseignement sur les exigences de l'épreuve (Dutercq & Lanéelle, 2013; Longchamp & Gilliéron, 2009). Ces données permettent également de considérer que, si les enseignants rapportent procéder à des changements en matière de contenus enseignés, ils ne considèrent pas ces modifications comme des changements déterminants dans leur façon d'enseigner.



En deuxième lieu, c'est la relation plus spécifique entre les résultats aux épreuves et les changements dans les pratiques enseignantes qui a retenu notre attention. Cet intérêt naît de l'importance accordée aux résultats dans les discours normatifs à l'égard des épreuves externes (Hanushek & Raymond, 2002; Klieme, 2004; Thélot, 2002), qui en font un élément déterminant dans la régulation des pratiques enseignantes. Appuyant les observations réalisées dans d'autres études (Maier, 2009; Schildkamp & Kuiper, 2010; Vanhoof, Verhaeghe & Van Petegem, 2012), nos analyses indiquent que les résultats n'entretiennent généralement aucune relation avec les changements dans les pratiques enseignantes. Seule l'utilisation instrumentale des épreuves paraît significativement liée aux résultats de l'épreuve externe, mais de manière assez faible.

Les résultats des élèves aux épreuves externes constitueraient donc une source d'information peu déterminante pour les enseignants pour orienter leurs pratiques. Partant de ce constat, l'enjeu de l'étude était d'approfondir la compréhension de cette apparente incapacité des résultats à réguler les pratiques enseignantes en analysant le rôle des mécanismes cognitifs intervenant dans l'interprétation des résultats.

De manière générale, force est de constater que les mécanismes cognitifs que nous avons mesurés ici ne rendent compte que faiblement des modifications dans les pratiques rapportées par les enseignants. Qui plus est, et de manière cohérente avec le faible impact des résultats eux-mêmes, les mécanismes les plus déterminants ne sont pas ceux directement liés à l'interprétation des résultats obtenus lors de la dernière passation d'épreuve. En effet, les analyses en pistes causales indiquent que c'est l'attitude envers l'évaluation externe et l'importance que les enseignants lui accordent qui jouent un rôle central dans le changement pédagogique et dans l'utilisation instrumentale de l'épreuve à des fins internes. Pour cette dernière catégorie de pratiques en outre, cette influence s'ajoute à celle de l'influence sociale des collègues. En somme, plus les enseignants accordent de l'importance aux épreuves et estiment que leurs collègues y accordent également de la valeur, plus ils ont tendance à rapporter avoir opéré des changements dans leur façon d'enseigner en réponse aux épreuves externes et à mobiliser ces épreuves afin de concevoir du matériel pédagogique. Pour le reste, ni les variations relatives à la légitimité accordée aux résultats comme feedback, ni celles liées à la dissonance négative ne

constituent des facteurs déterminants dans la régulation des pratiques enseignantes.

Comment le comprendre? Une première piste est à chercher du côté des statistiques descriptives. La plupart des enseignants de notre échantillon ressentent en effet peu de dissonance cognitive que nous avons qualifiée de négative. Ils semblent bien plus portés à ressentir une dissonance positive elle-même liée à la faible légitimité accordée aux résultats à la fois comme feedback relatif aux compétences des élèves et comme feedback personnel pour l'enseignant. Cette dissonance positive nous porte d'ailleurs à considérer que le rejet de la validité des résultats comme feedback ne constitue pas un mécanisme de défense destiné à la préservation de l'image de soi, puisqu'il est peu probable que des résultats supérieurs aux attentes constituent une menace pour celle-ci. En outre, cette faible dissonance négative combinée au rejet de la légitimité des résultats comme feedback de performances s'accompagnait d'une faible attribution interne à l'égard des résultats. Dans ce contexte où les résultats de l'évaluation externe ne sont pas considérés par les enseignants comme une source d'information pertinente, le processus réflexif attendu des dispositifs d'évaluation externe ne semble pas se produire. Les épreuves externes ne provoqueraient donc pas l'« effet miroir » espéré (Thélot, 2002).

Nos résultats indiquent finalement que la prédiction de l'intention comportementale est bien meilleure que celle des comportements effectifs. Reste cependant à savoir si cet écart entre l'intention comportementale et les comportements effectifs est conjoncturel ou si, ce qui semble davantage probable compte tenu de la complexité des comportements en question (Sheeran, 2002), cet écart se maintiendra. Même dans ce dernier cas cependant, cette observation permet de soutenir l'idée selon laquelle l'influence relativement faible des épreuves externes sur les pratiques enseignantes ne dépend pas d'un manque de volonté de leur part, mais probablement d'obstacles compromettant la mise en œuvre de nouvelles routines comportementales (tels que le manque de temps ou les injonctions contradictoires).

## Conclusions, limites et perspectives de recherche

Les résultats de notre recherche indiquent que la faible utilisation des résultats issus des épreuves externes en



Belgique francophone est due au fait que les enseignants belges francophones ne les perçoivent pas comme un feedback de performances, ce qui limite leur engagement dans un processus réflexif une fois confrontés aux résultats.

Les observations établies sur la base de cette étude doivent cependant être replacées dans leur contexte. Ainsi, le niveau relativement élevé des scores moyens au CE1D pourrait être considéré comme une des principales raisons de la dissonance positive prédominante dans les comptes rendus des enseignants. Nous ne savons donc pas si, confrontés à des résultats plus défavorables, ces derniers s'engageraient davantage dans un processus réflexif aboutissant bel et bien à des modifications comportementales.

D'autres limites sont également à considérer pour apprécier la portée des résultats. Nous avons déjà souligné à cet égard le problème du faible taux de réponse et ses menaces pour la validité interne et externe. Toutefois, la variance observée dans les variables dépendantes (qui limite le risque d'une surreprésentation d'enseignants très attachés ou au contraire très remontés contre l'évaluation externe) ainsi que la représentativité de l'échantillon par rapport à une série d'indicateurs personnels et institutionnels nous laissent penser que ces menaces sont limitées. Une deuxième difficulté plus spécifiquement liée à la méthodologie utilisée fait référence au fait que le questionnaire repose sur des comptes rendus de la part des enseignants. La relativité des changements auto-rapportés fait alors obstacle à la possibilité d'établir une représentation précise de l'étendue de l'impact des épreuves externes sur les pratiques enseignantes. Rappelons néanmoins que l'objectif de l'étude était aussi de saisir la nature de la dyna-

mique engagée par les enseignants sur la base des évaluations externes et de leurs résultats, et que, de ce point de vue, se reposer sur des mesures auto-rapportées était pertinent. Troisièmement, les tailles d'effet sont modestes, à l'exception des effets sur l'intention de changement. Pour le dire autrement, les explications du changement des pratiques enseignantes en termes d'interprétation des résultats, d'attitude à l'égard de l'évaluation externe et d'influence sociale sont loin d'épuiser l'explication des différences entre les enseignants dans leur rapport avec l'évaluation externe et son utilisation. Ce résultat invite donc les chercheurs à élaborer de nouvelles explications, moins liées aux résultats et à leur interprétation, qui complètent ou approfondissent les principales conclusions mises au jour dans cette étude. Nous suggérons notamment qu'une meilleure conceptualisation des effets de contexte et d'influence sociale, dans la ligne des travaux sur les communautés professionnelles locales (Ronfeldt, Farmer, McQueen *et al.*, 2015), soit élaborée. Une telle conceptualisation devrait permettre de mieux saisir comment les enseignants s'emparent du dispositif d'évaluation externe (et pas uniquement des résultats) pour interroger leurs pratiques d'enseignement. Finalement, les relations entre l'intention au changement et le changement effectif dans les pratiques enseignantes méritent elles aussi d'être approfondies en procédant à des études longitudinales.

Esteban Rozenwajn

Université catholique de Louvain, GIRSEF  
esteban.rozenwajn@uclouvain.be

Xavier Dumay

Université catholique de Louvain, GIRSEF  
xavier.dumay@uclouvain.be

## Bibliographie

- AJZEN I. (1991). «The theory of planned behavior». *Organizational Behavior and Human Decision Processes*, n° 50, p.179-211.
- AJZEN I. & FISHBEIN M. (2005). «The influence of attitudes on behavior». In D. Albarracín, B. T. Johnson & M. P. Zanna (dir.), *The handbook of attitudes*. Mahwah (Etats-Unis) : Erlbaum, p.173-222.
- AMREIN-BEARDSLEY A., BERLINER D. C. & RIDEAU S. (2010). «Cheating in the first, second and third degree: Educators' responses to high-stakes testing». *Educacion Policy Analysis Archives*, n° 18(14), p.1-36.
- AU W. (2007). «High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis». *Educational Researcher*, n° 36(5), p.258-267.
- BICKMAN L. & RIEMER M. (2011). «Using program theory to link social psychology and program evaluation». In M. M. Melvin, D. I. Stewart & B. Campbell (dir.), *Social Psychology and Evaluation*. New York : The Guilford Press, p.104-138.
- CIALDINI R. B. & TROST M. R. (1998). «Social influence : social norms, conformity, and compliance». In D. T. Gilbert, S. T. Fiske & L. Gardner (dir.), *Handbook of Social Psychology*. New York : McGraw Hill, 4<sup>e</sup> éd., p.151-192.
- DIAMOND J. B. (2007). «Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction». *Sociology of Education*, n° 80(4), p.285-313.

- DIERENDONCK C. (2008). *Comment les évaluations externes des acquis des élèves sont-elles perçues par les enseignants du primaire dans les cantons de Neuchâtel, Vaud et Fribourg ? Enquête exploratoire*. Neuchâtel : IRDP.
- DIERENDONCK C. & FAGNANT A. (2010). « Quelques réflexions autour des épreuves d'évaluation développées dans le cadre de l'approche par compétences ». *Le Bulletin de l'ADMEE-EUROPE*, n°1.
- DRAELANTS H., DUPRIEZ V. & MAROY C. (2011). *Le système scolaire*. Bruxelles : CRISP.
- DUPRIEZ V. & MONS N. (dir.) (2011). « Les politiques d'accountability. Des changements institutionnels aux transformations locales ». *Éducation comparée*, n°5, p.7-154.
- DUPRIEZ V. (2015). *Peut-on réformer l'école ?* Bruxelles : De Boeck.
- DUTERCQ Y. & LANÉELLE X. (2013). « La dispute autour des évaluations des élèves dans l'enseignement français du premier degré ». *Sociologie*, n°4, p.43-62.
- EAGLY A. H. & CHAIKEN S. (1998). « Attitude structure and function ». In D. T. Gilbert, S. T. Fiske & G. Lindzey (dir.), *Handbook of Social Psychology*, 4<sup>th</sup> ed. New York : McGraw Hill, p.269-322.
- HANUSHEK E. A. & RAYMOND M. E. (2002). « Sorting out accountability systems ». In W. M. Evers & H. J. Walberg (dir.), *School accountability*. Stanford : Hoover Institution Press, p.75-104.
- HELLRUNG K. & HARTIG J. (2013). « Understanding and using feedback. A review of empirical studies concerning feedback from external evaluations to teachers ». *Educational Research Review*, n°9, p.174-190.
- JORDAN A. H. & AUDIA P. G. (2012). « Self-Enhancement and Learning from Performance Feedback ». *Academy of Management Review*, n°37(2), p.211-231.
- KLIEME E. (2004). *Le développement de standards nationaux de formation : une expertise*. Bonn : BMBF.
- KLUGER A. N. & DENISI A. (1996). « The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory ». *Psychological Bulletin*, n°119(2), p.254-284.
- LAFONTAINE D., SOUSSI A. & NIDEGGER C. (2009). « Évaluations internationales et/ou épreuves nationales : tensions et changements de pratiques ». In M. L. Lopez & M. Crahay (dir.), *Évaluations en tension*. Bruxelles : De Boeck, p.61-80.
- LONGCHAMP A.-L. & GILLIÉRON P. G. (2009). *Épreuves cantonales de référence de mathématiques en 6<sup>e</sup> année : le point de vue des enseignants sur l'utilisation du dispositif*. Lausanne : URSP.
- MAIER U. (2009). « Towards state mandated testing in Germany: how do teachers assess the pedagogical relevance of performance feedback information? » *Assessment in Education: Principles, Policy & Practice*, n°16(2), p.205-226.
- MAIER U. (2010). « Accountability policies and teachers' acceptance and usage of school performance feedback : A comparative study ». *School Effectiveness and School Improvement*, n°21(2), p.145-165.
- MONS N. (2009). *Les effets théoriques et réels de l'évaluation standardisée*. Paris : Eurydice.
- PEDULLA J. J., ABRAMS L. M., MADAUS G. F., RUSSELL M. K., RAMOS M. A. & MIAO J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: findings from a national survey on teachers*. En ligne : <https://eric.ed.gov/?id=ED481836> (consulté le 2 mars 2015).
- PHELPS R. (2005). *Defending standardized testing*. New Jersey : Lawrence Erlbaum Associates.
- RONFELDT M., OWENS FARMER S., MCQUEEN, K. & GRISSOM J. A. (2015). « Teacher collaboration in instructional teams and student achievement ». *American Educational Research Journal*, n°52, p.475-514.
- ROZENWAJN E. (à paraître). *Low-stakes testing and teachers' practices: what about test scores?*
- SCHILDKAMP K. & KUIPER W. (2010). « Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors ». *Teaching and Teacher Education*, n°26(3), p.482-496.
- SHEERAN P. (2002). « Intention-Behavior Relations: A Conceptual and Empirical Review ». *European Review of Social Psychology*, n°12(1), p.1-36.
- SMITH L. M. (1991). « Meanings of Test Preparation ». *American Educational Research Journal*, n°28(3), p.521-542.
- THÉLOT C. (2002). « Évaluer l'école ». *Études*, n°10(397), p.323-334.
- VALLI L. & BUESE D. (2007). « The changing roles of teachers in an era of high-stakes accountability ». *American Educational Research Journal*, n°44(3), p.519-558.
- VANHOOF J., VERHAEGHE G. & VAN PETEGEM P. (2012). « Flemish primary teachers' use of school performance feedback and the relationship with school characteristics ». *Educational Research*, n°54(4), p.431-449.
- VERDIÈRE J. (2013). « Les enseignants du collège en France : Faire face à un travail plus formalisé rendu plus visible ». In V. Dupriez & R. Malet (dir.), *L'évaluation dans les systèmes scolaires*. Bruxelles : De Boeck, p.77-90.

## Annexe. Saturations entre les items et les composantes

Tableau 4. Échelles de transformation dans les pratiques enseignantes

Items	Composantes			
	1	2	3	4
L'évaluation externe a provoqué des changements dans ma façon d'enseigner (Chpeda)	<b>0,857</b>	0,279	-0,405	0,260
En l'absence du CE1D, ma façon d'enseigner aurait été différente (Chpeda)	<b>0,866</b>	0,201	-0,254	0,194
L'évaluation externe m'a amené à opérer des changements au niveau de mes méthodes pédagogiques (Chpeda)	<b>0,664</b>	0,169	-0,573	-0,018
Cela m'arrive de reprendre des questions de l'épreuve externe pour la conception des exercices de classe (Utilinst)	0,198	<b>0,828</b>	-0,294	0,225
Cela m'arrive de reprendre des questions de l'épreuve externe pour la conception des examens internes (Utilinst)	0,224	<b>0,763</b>	-0,213	-0,138
Cela m'arrive de reprendre des épreuves externes à des fins internes (Utilinst)	0,272	<b>0,555</b>	-0,216	0,494
Je travaille davantage les matières présentées lors de l'épreuve externe (AlignCur)	0,571	0,235	-0,434	<b>0,739</b>
Je travaille davantage les matières de l'épreuve externe où mes élèves ont éprouvé des difficultés (AlignCur)	0,142	0,064	-0,294	<b>0,806</b>
L'épreuve externe influence les contenus que j'enseigne à mes élèves (AlignCur)	0,669	0,328	-0,428	<b>0,592</b>
Pour l'année qui vient, je compte davantage faire usage des épreuves externes dans mes cours (Inten)	0,334	0,352	<b>-0,864</b>	0,174
Pour l'année qui vient, je compte davantage travailler les matières de l'épreuve externe où mes élèves ont éprouvé des difficultés (Inten)	0,287	0,170	<b>-0,831</b>	0,393
Je ne tiendrais pas compte du contenu du CE1D pour planifier mes cours l'année prochaine (Inten)	0,129	0,516	<b>-0,551</b>	0,239
Pour l'année qui vient, je compte modifier mes méthodes pédagogiques à partir du CE1D (Inten)	0,483	0,222	<b>-0,778</b>	0,237

Note : Chpeda = Changements pédagogiques ; Utilinst = Utilisation instrumentale ; AlignCur = Alignement curriculaire ; Inten = Intention de changement.

Tableau 5. Échelles d'interprétation des résultats

Items	Composantes				
	1	2	3	4	5
L'amélioration des résultats de mes élèves au CE1D est un objectif important pour moi (CentR)	<b>0,850</b>	-0,079	0,030	-0,060	0,051
Réduire le nombre d'élèves en échec au CE1D est un objectif important pour moi (CentR)	<b>0,788</b>	-0,024	0,002	0,135	-0,096
Il est important pour moi que mes élèves obtiennent de meilleurs scores au CE1D que les années précédentes (CentR)	<b>0,907</b>	-0,172	-0,152	-0,091	-0,011
Les résultats du CE1D doivent guider le travail de l'enseignant (CentR)	<b>0,511</b>	0,171	0,341	0,035	0,090
Il est important pour moi d'orienter mon enseignement en fonction des résultats de mes élèves au CE1D (CentR)	<b>0,637</b>	0,206	0,207	0,093	0,069
Les résultats mesurent correctement les compétences d'un élève (VEI)	0,126	0,143	<b>0,726</b>	0,017	0,137
Les scores du CE1D reflètent le niveau de mes élèves (VEI)	0,055	0,077	<b>0,814</b>	-0,113	0,093
Les épreuves sont trop faciles (VEI)	-0,036	0,206	<b>-0,755</b>	-0,041	0,159
Les pourcentages d'élèves en réussite sont trop élevés (VEI)	0,062	0,351	<b>-0,740</b>	-0,016	0,106
Le nombre d'échecs était trop élevé (DissNeg)	0,060	<b>-0,812</b>	0,183	0,046	-0,047
Les scores de mes élèves étaient trop faibles (DissNeg)	0,077	<b>-0,875</b>	-0,031	0,008	0,106
Je me sens personnellement responsable des échecs de mes élèves (AttInt)	0,005	0,029	0,009	<b>0,811</b>	-0,004
J'aurais pu éviter les échecs de certains des élèves (AttInt)	-0,002	-0,060	-0,115	<b>0,815</b>	-0,015
Les résultats permettent d'évaluer la compétence de l'enseignant (VEns)	0,020	-0,033	0,045	0,008	<b>0,789</b>
Un enseignant compétent à peu d'élèves en échec aux épreuves externes (VEns)	-0,033	-0,023	-0,140	-0,118	<b>0,882</b>
Je considère les résultats de l'épreuve externe comme une évaluation de mon travail (VEns)	-0,061	-0,091	0,295	0,237	<b>0,447</b>
Les résultats de mes élèves constituent un reflet des efforts que je mets dans mon travail (VEns)	0,145	0,082	0,051	0,291	<b>0,494</b>

Note : CentR = Centralité résultats; VEI = Validité élèves; DissNeg = Dissonance négative; AttInt = Attribution interne; VEns = Validité enseignants.

**Tableau 6. Échelles de perception de l'épreuve**

Items	Composantes	
	1	2
Je suis satisfait du CE1D (Att)	<b>0,528</b>	-0,368
Une épreuve externe certificative était nécessaire à ce niveau (Att)	-0,056	<b>-0,897</b>
Il faudrait supprimer les épreuves externes (Att)	0,010	<b>-0,855</b>
Le CE1D est inutile (Att)	0,119	<b>-0,811</b>
Le fait qu'il y ait une épreuve externe certificative à ce niveau est une bonne chose (Att)	-0,023	<b>-0,895</b>
Le CE1D est important à la réalisation de mon travail (Cent)	<b>0,683</b>	-0,198
Je ne pourrais pas me passer du CE1D pour la réalisation de mon travail (Cent)	<b>0,848</b>	0,124
Le CE1D est un outil indispensable au travail enseignant (Cent)	<b>0,828</b>	-0,034
Les enseignants devraient davantage se baser sur le CE1D pour orienter leur travail (Cent)	<b>0,788</b>	-0,020
Je devrais davantage utiliser le CE1D dans mon travail (Cent)	<b>0,727</b>	0,047

Note : Att = Attitude générale ; Cent = Centralité de l'évaluation externe.

**Tableau 7. Échelles d'influence sociale**

Items	Composantes	
	1	2
En dehors des périodes de passation des épreuves externes, la direction se base-t-elle sur les résultats lorsqu'elle interagit avec les enseignants ? (isdir)	<b>0,773</b>	-0,055
La direction insiste-t-elle sur l'utilisation des informations transmises par l'évaluation externe ? (isdir)	<b>0,797</b>	0,042
La direction insiste-t-elle sur la nécessité d'améliorer les résultats ? (isdir)	<b>0,810</b>	0,078
En cas de résultats décevants, la direction fait-elle appel à des inspecteurs ou conseillers pédagogiques ? (isdir)	<b>0,721</b>	0,015
La direction réalise-t-elle des entretiens individuels au sujet des résultats ? (isdir)	<b>0,742</b>	-0,120
La direction fournit-elle des conseils pédagogiques en relation avec les résultats ? (isdir)	<b>0,647</b>	0,108
En dehors des périodes de passation, les épreuves externes sont-elles un sujet de discussion entre enseignants ? (iscol)	-0,062	<b>0,869</b>
Vos collègues insistent-ils sur la nécessité d'améliorer les résultats ? (iscol)	0,097	<b>0,804</b>

Note : isdir = influence sociale direction ; iscol = influence sociale collègues.