

Évaluation des apprentissages en classe et évaluation à grande échelle : quels sont les effets des épreuves externes sur les pratiques évaluatives des enseignants ?

Gonzague Yerly

Volume 40, Number 1, 2017

URI: <https://id.erudit.org/iderudit/1041003ar>

DOI: <https://doi.org/10.7202/1041003ar>

[See table of contents](#)

Publisher(s)

ADMEE-Canada - Université Laval

ISSN

0823-3993 (print)

2368-2000 (digital)

[Explore this journal](#)

Cite this article

Yerly, G. (2017). Évaluation des apprentissages en classe et évaluation à grande échelle : quels sont les effets des épreuves externes sur les pratiques évaluatives des enseignants ? *Mesure et évaluation en éducation*, 40(1), 33–60. <https://doi.org/10.7202/1041003ar>

Article abstract

Teachers must nowadays deal with large-scale external testing in most educational systems. Nevertheless, the impact of the uses of external testing by the teachers are still not well known. The aim of this study is to analyse specifically the impact of external testing on teacher's assessment practices. A qualitative comprehensive analysis was carried out on the experiences of 16 primary teachers in the canton of Fribourg in Switzerland (semi-directive interviews). Results show that teachers may use external tests' form and contents to adjust their assessment practices. In contrast, they seldom use their results, and the observed adjustments are sometimes made in an instrumental way, to optimize pupils' achievement to the tests. In the studied context, moderate impact is due to different postures of teachers about external testing and/or to their perceptions of specific characteristics of local devices.

Évaluation des apprentissages en classe et évaluation à grande échelle : quels sont les effets des épreuves externes sur les pratiques évaluatives des enseignants ?

Gonzague Yerly

Université de Fribourg

MOTS CLÉS : épreuves externes, effets sur les pratiques évaluatives, enseignants du primaire, accountability, Suisse

Dans la plupart des systèmes éducatifs, les enseignants font face aujourd'hui à l'expansion des évaluations à grande échelle. Néanmoins, les effets de ces dispositifs sur les pratiques enseignantes sont encore peu connus. L'objectif de cet article est de proposer une analyse spécifique des effets des épreuves externes sur les pratiques d'évaluation des apprentissages. L'analyse qualitative compréhensive porte sur l'expérience de 16 enseignants du primaire confrontés aux épreuves externes du canton de Fribourg en Suisse (entretiens semi-dirigés). Les résultats montrent que ceux-ci se servent des contenus et de la forme des épreuves externes pour ajuster certaines démarches du processus d'évaluation sommative. Par contre, ils ne mobilisent que très peu les résultats, et les ajustements observés sont parfois réalisés de manière plutôt instrumentale afin d'optimiser la réussite des élèves aux épreuves. Dans le contexte étudié, ces effets plutôt modérés s'expliquent par différentes postures des enseignants quant à l'évaluation externe et/ou par leurs perceptions de certaines caractéristiques des dispositifs locaux.

KEY WORDS: external testing, impact on assessment practices, primary teachers, soft accountability, Switzerland

Teachers must nowadays deal with large-scale external testing in most educational systems. Nevertheless, the impact of the uses of external testing by the teachers are still not well known. The aim of this study is to analyse specifically the impact of external testing on teacher's assessment practices. A qualitative comprehensive analysis was carried out on the experiences of 16 primary teachers in the canton of Fribourg in Switzerland (semi-directive interviews). Results show that teachers may use external tests' form and contents to adjust their assessment practices. In contrast, they seldom use their results, and the observed adjustments

are sometimes made in an instrumental way, to optimize pupils' achievement to the tests. In the studied context, moderate impact is due to different postures of teachers about external testing and/or to their perceptions of specific characteristics of local devices.

PALAVRAS-CHAVE: provas externas, efeitos sobre as práticas avaliativas, professores do ensino primário, responsabilização débil, Suíça

Na maioria dos sistemas educativos, os professores enfrentam hoje em dia a expansão das avaliações em larga escala. No entanto, os efeitos destes dispositivos nas práticas de ensino ainda são pouco conhecidos. O objetivo deste artigo é propor uma análise específica dos efeitos das provas externas sobre as práticas de avaliação das aprendizagens. A análise qualitativa compreensiva centra-se na experiência de 16 professores primários com provas externas no cantão de Friburgo na Suíça (entrevistas semiestruturadas). Os resultados mostram que os professores primários se servem do conteúdo e da forma das provas externas para ajustar certos procedimentos do processo de avaliação sumativa. Em contrapartida, praticamente não mobilizam os resultados, e os ajustamentos observados são por vezes realizados de maneira meramente instrumental para maximizar o sucesso dos alunos nas provas. No contexto estudado, estes efeitos bastante moderados explicam-se devido a diferentes posturas dos professores sobre a avaliação externa e às suas percepções de certas características dos dispositivos locais.

Note de l'auteur : La correspondance liée à cet article peut être envoyée à : [gonzague.yerly@unifr.ch].

Nos remerciements vont à Cécile Mathou (Université de Montréal) et aux professeurs Philippe Genoud (Université de Fribourg) et Dany Laveault (Université d'Ottawa) pour leur relecture et leurs précieux conseils. Nous remercions également les évaluateurs de la revue pour leurs commentaires enrichissants.

Introduction

Les épreuves externes sont des dispositifs d'évaluation des apprentissages des élèves à grande échelle qui sont conçus par une entité extérieure à la classe, au palier local, national ou transnational. Elles ont une longue histoire dans les systèmes éducatifs européens et nord-américains (Eurydice, 2009). Au départ, ces examens – dont les caractéristiques présentent une grande diversité (p. ex., contenus, forme, conception, administration) – avaient pour but d'uniformiser le passage, la certification, voire la sélection des élèves. Depuis la fin des années 2000, les épreuves externes connaissent une expansion. Si leurs fonctions pédagogiques précédemment exposées n'ont pas disparu, sont venues s'ajouter celles d'un outil de politique éducative (Mottier Lopez, 2009). En effet, l'évaluation externe ne sert aujourd'hui plus uniquement à évaluer les apprentissages des élèves, mais aussi à contrôler et à orienter le système et ses acteurs (Maroy, 2013).

Dans la plupart des systèmes éducatifs, les enseignants font face aujourd'hui à de nouveaux modes de régulation institutionnelle et à leurs outils standardisés. Ils voient leur profession se transformer en partie : ils jouissent d'une autonomie de moyens pour atteindre les objectifs fixés, mais doivent en contrepartie rendre des comptes quant aux résultats des élèves (Demailly, 2001). Ces politiques nationales ou locales d'*accountability*¹ se distinguent quant à leurs modes de régulation et à leurs outils. Dans la littérature, elles sont généralement réparties en deux catégories en fonction de leurs enjeux pour les différents acteurs (Carnoy & Loeb, 2002; Mons, 2009). En ce qui concerne le modèle d'*accountability* « dure » (*strong accountability*), les résultats des épreuves externes s'accompagnent de conséquences plutôt importantes, notamment pour les enseignants (*high-stakes tests*). On mise alors sur des incitations extérieures pour produire de meilleurs résultats, par exemple des primes ou des licenciements. Ce modèle est plutôt le fait des systèmes scolaires anglo-saxons (p. ex., États-Unis, Angleterre). Dans le modèle d'*accountability* « douce » (*soft accountability*), les résultats des épreuves externes impliquent peu de conséquences pour les enseignants (*low-stakes tests*). Ils servent surtout à leur

fournir des informations pour ajuster leurs actions. On mise alors sur l'initiative des enseignants pour procéder à des changements ou ajustements, sans incitations extérieures. Ce modèle est principalement européen continental (p. ex., Allemagne, Autriche, Belgique, France, Suisse).

Ainsi, il est attendu par les administrations scolaires que les épreuves externes soient un moyen de réguler les pratiques des enseignants sur le terrain, de les harmoniser, voire de les « aligner ». Ces dispositifs sont devenus un des principaux moyens pour communiquer les « bonnes pratiques » aux enseignantes. Néanmoins, si les effets « théoriques » de l'évaluation externe sur les pratiques enseignantes sont bien connus, il semble que ce ne soit pas encore tout à fait le cas des effets « réels », c'est-à-dire ceux observés sur le terrain. De plus, l'impact des épreuves externes est remis en question : dans les classes, il serait plutôt limité, surtout lorsque les épreuves comportent de faibles enjeux (Mons, 2009).

Cet article propose d'analyser les effets des épreuves externes spécifiquement sur les pratiques évaluatives des enseignants dans un contexte d'accountability douce. Il présente tout d'abord un cadre théorique et empirique, qui débouche sur les objectifs et les questions de recherche. Il précise ensuite la méthodologie adoptée, avant que les résultats de l'analyse soient présentés, puis discutés au regard de la littérature.

Cadre théorique et empirique

Effets des épreuves externes sur les pratiques enseignantes

Les études empiriques traitant des effets des épreuves externes sur les pratiques des enseignants sont surtout anglo-saxonnes, voire essentiellement nord-américaines (accountability dure). Elles restent peu nombreuses en Europe continentale (accountability douce). Il s'agit d'études de cas ou d'enquêtes par questionnaire et/ou par entretiens auprès d'enseignants de toutes années d'études et dans divers contextes. Certains travaux de synthèse (Au, 2007 ; Mons, 2009 ; Rozenwajn & Dumay, 2015 ; Yerly, 2014) documentent le fait que, de manière générale et indépendamment de la variété des dispositifs, les épreuves externes entraînent surtout les enseignants à aligner les contenus enseignés et évalués en classe sur les contenus des examens, mais peu à développer leurs méthodes d'enseignement. Surtout lorsqu'elles comportent des conséquences importantes pour les

élèves et/ou les enseignants, les épreuves externes poussent les enseignants vers des pratiques centrées sur la réussite des tests, par exemple le rétrécissement du curriculum et l'enseignement pour l'examen (*teaching to the test*).

Les effets des épreuves externes sur les pratiques enseignantes seraient surtout faibles dans les contextes d'accountability douce (Rozenwajn & Dumay, 2015). Dans ces contextes où les enjeux des épreuves sont moins importants, les enseignants montrent généralement une certaine acceptation des dispositifs (Cattonar, Dumay, Mangez & Maroy, 2010; Dierendonck, 2008; Yerly, 2014), mais auraient de la difficulté et seraient peu enclins à utiliser les épreuves externes pour modifier leurs pratiques. L'impact limité des épreuves externes sur les pratiques enseignantes dans les contextes d'accountability douce est constaté par des études empiriques, notamment en France, en Belgique francophone et en Suisse romande (Lafontaine, Soussi & Nidegger, 2009; Yerly, 2017). Il existerait donc, dans ces systèmes éducatifs, un «découplage» important entre les changements attendus par les autorités et ceux effectifs dans les classes (Coburn, 2004). Dans le cas d'une accountability dure, ce découplage serait moins marqué (Spillane, Reiser & Reimer, 2002; Diamond, 2012), mais les stratégies pour augmenter les résultats y sont plus importantes (Amrein & Berliner, 2002; Kempf, 2016).

Effets des épreuves externes sur les pratiques d'évaluation des apprentissages

La relation, voire les tensions entre épreuves externes et pratiques d'évaluation en classe est l'objet de nombreuses contributions théoriques. Nous relevons notamment l'important apport de l'ouvrage collectif dirigé par Mottier Lopez et Crahay (2009) et celui de l'analyse de Laveault, Dionne, Lafontaine, Tessaro et Allal (2014). Certaines études qualitatives traitant des effets des épreuves externes sur les pratiques des enseignants du primaire au sens large montrent que, dans différents contextes d'accountability douce (Belgique francophone, France et Suisse romande), les épreuves externes peuvent avoir des effets non seulement sur les contenus évalués, mais aussi sur les méthodes d'évaluation (voir Tableau 1). Ce constat est partagé par Tessaro (2015) dans son analyse spécifique à l'évaluation des apprentissages.

Tableau 1
Effets des épreuves externes sur les pratiques évaluatives

	Contexte	Méthodologie	Effets sur les contenus d'évaluation	Effets sur les méthodes d'évaluation
Cattonar et al. (2010)	Belgique francophone Épreuves régionales	Entretiens 28 enseignants	Évaluation d'autres processus cognitifs (notamment les compétences)	Imitation des conditions de passation (voire un <i>coaching</i>) Modification de la façon de poser les questions
Dutercq et Lanéelle (2013)	France Épreuves nationales	Entretiens 20 enseignants	Alignement sur les contenus attendus	Ajustement du degré d'exigence Diversification des exercices
Longchamp (2010)	Canton de Vaud (Suisse) Épreuves cantonales	Entretiens 20 enseignants	Alignement sur les contenus attendus	Clarification des consignes Ajustement des barèmes
Tessaro (2015)	Canton de Genève (Suisse) Épreuves cantonales	Groupe de discussion (<i>focus group</i>) 28 enseignants	Alignement sur les contenus attendus	Usage d'un tableau de spécifications Ajustement des barèmes
Yerly (2014) (partie qualitative)	Canton de Fribourg (Suisse) Épreuves cantonales	Entretiens 8 enseignants	Alignement sur les contenus attendus Évaluation d'autres processus cognitifs (notamment les compétences)	Imitation des conditions de passation Modification de la façon de poser les questions

Les travaux quantitatifs qui permettent de constater l'ampleur des effets des épreuves externes sur les pratiques évaluatives des enseignants sont plus rares et généralement non spécifiques à l'évaluation externe. Il s'agit tout d'abord de deux sondages à large échelle portant sur les pra-

tiques évaluatives des enseignants en France et en Suisse romande. Ces études par questionnaire tendent à démontrer un usage limité (faible à modéré) des épreuves externes par les enseignants.

En France, dans l'étude du ministère de l'Éducation nationale (MEN, 1994; questionnaire auprès de 3 106 enseignants) sur les pratiques évaluatives des enseignants au collège (élèves de 11 à 15 ans), il est possible d'observer une utilisation très modérée des épreuves externes, même pour les enseignants des disciplines évaluées. En effet, seuls 13% des enseignants de lettres et 9% des enseignants de mathématiques déclarent les utiliser « toujours » ou « souvent ». Ils sont même 41% (lettres) et 48% (mathématiques) à ne « jamais » les utiliser.

En Suisse romande, une étude similaire de Soussi, Ducrey, Ferrez, Nidegger et Viry (2006) menée dans le canton de Genève sur les pratiques évaluatives des enseignants (questionnaire auprès de 718 enseignants du primaire et de 450 du secondaire) montre que les enseignants des deux ordres acceptent modérément l'idée que l'évaluation externe soit « génératrice d'idées nouvelles » et qu'elle permette « une certaine régulation des pratiques ».

Encore en Suisse, dans le canton de Fribourg, la partie quantitative de notre thèse de doctorat (Yerly, 2014), qui traite des effets de l'évaluation externe sur les pratiques des enseignants du primaire (questionnaire auprès de 305 enseignants), montre que, même si les effets des épreuves externes dans ce contexte touchent avant tout les pratiques d'évaluation sommative des enseignants, ces effets sur les pratiques évaluatives restent modérés. Un tiers des enseignants ne se sent pas ou peu touché, un tiers se sent modérément touché et un tiers se sent plutôt ou fortement touché. Un intérêt particulier de cette recherche est d'expliquer la variation de ces effets spécifiques. Les analyses de régression montrent qu'elle ne dépend ni du profil des enseignants (variables socioprofessionnelles et croyances envers l'enseignement et l'évaluation) ni du type d'évaluation externe (épreuve de référence par rapport aux épreuves d'orientation), mais de la perception du dispositif par les enseignants (enjeux et utilité perçus).

Toutefois, ces études ne permettent pas d'analyser plus finement comment et pour quelles raisons les enseignants utilisent l'évaluation externe pour leurs pratiques d'évaluation des apprentissages. C'est l'apport de cet article.

Objectifs et questions de recherche

Le présent article entend contribuer à la compréhension des effets de l'évaluation externe sur les pratiques enseignantes dans un contexte d'accountability douce. L'intérêt principal de l'article est de proposer une analyse spécifique et approfondie des effets des épreuves externes sur les pratiques d'évaluation des apprentissages en classe. Pour nous, l'évaluation des apprentissages des élèves fait partie intégrante du processus d'enseignement/apprentissage. Toutefois, la littérature présentée montre que l'évaluation des apprentissages est une activité particulièrement touchée par l'évaluation externe, tant en ce qui concerne les contenus évalués que les méthodes adoptées. Pour des raisons analytiques, nous isolons donc l'évaluation des autres pratiques didactiques qui peuvent être également touchées par les épreuves externes (p. ex., planification de l'enseignement, enseignement de la matière aux élèves). Nos objectifs sont de rendre compte des modifications qu'entraînent les épreuves externes sur les pratiques évaluatives des enseignants, de les documenter et surtout – et c'est là l'apport original de notre analyse – d'en comprendre le processus : comment et pour quelles raisons les enseignants procèdent-ils ou non à ces ajustements de pratiques évaluatives ?

Contexte de la recherche

Les épreuves externes connaissent depuis une décennie une forte expansion en Suisse aux paliers fédéral, régional et cantonal (CSRE, 2014)². En effet, depuis l'accord intercantonal HarmoS (CDIP, 2007) visant à harmoniser la scolarité obligatoire, les cantons sont, parmi d'autres mesures, tenus de participer au pilotage du système. Les autorités cantonales, surtout celles des cantons francophones, ont fait le choix de développer l'usage d'épreuves externes communément appelées épreuves cantonales. Ces épreuves ont une histoire antérieure à cette réforme ; elles avaient alors comme finalité le passage des élèves, voire leur sélection, surtout à la fin du primaire. Si, pour certaines épreuves, cette fonction originelle n'a pas disparu, elles servent aujourd'hui également au pilotage des systèmes cantonaux (Ntamakiliro & Tessaro, 2010). Le tableau 2 permet d'observer leur répartition dans les cantons romands au moment de notre étude (2014) et leur forte expansion depuis 2004.

Tableau 2
Épreuves cantonales en Suisse romande – année scolaire 2013-2014

Année d'études	3 ^e	4 ^e	5 ^e	6 ^e	7 ^e	8 ^e	9 ^e	10 ^e	11 ^e
Canton									
Fribourg		0		0		X			0
Genève		X		X		X	X	X	X
Jura				0		X		0	
Neuchâtel	0	0	0	0	0	X		0 X	
Valais		0	0	0 X	0	0 X		X	0 X
Vaud		X		X		X		X	X

Note. X = avec conséquences sur le cursus de l'élève; 0 = sans conséquences (ou limitées) sur le cursus de l'élève; en gras = épreuves externes inscrites déjà avant 2004.

Le tableau a été conçu d'après diverses sources d'information: CSRE (2014) et IRDP (2015).

La scolarité primaire du canton de Fribourg a été sélectionnée pour notre étude, soit de la 3^e à la 8^e année de scolarité selon la terminologie HarmoS³. Elle est organisée en cycles d'apprentissage, où les enseignants ont la charge de toutes les disciplines pour un même groupe-classe pendant deux ans. Ce contexte présente l'intérêt de compter deux types d'épreuves externes : épreuves de référence (fin du cycle 3-4^{es} années ; fin du cycle 5-6^{es} années) et épreuves d'orientation (fin du cycle 7-8^{es} années). Ces épreuves servent au pilotage du système, dans un modèle d'accountability douce. Il est attendu par l'administration que ces outils soient un moyen d'informer les inspecteurs, les parents et surtout les enseignants quant aux apprentissages effectués par les élèves à la fin de chaque cycle. Ils doivent à terme permettre l'alignement des contenus et l'harmonisation des pratiques⁴. L'administration compte alors sur l'initiative individuelle des enseignants pour modifier leurs pratiques, sans prévoir ni suivi spécifique individuel ou collectif des résultats (aucune présentation, formation ou documentation pédagogique), ni sanction ou conséquences externes. Ces épreuves sont toutes de type sommatif. Elles font le bilan des apprentissages des élèves à la fin d'un cycle de deux ans. Elles testent l'acquisition d'une sélection de connaissances et de compétences inscrites dans le

programme commun des cantons francophones, le plan d'études romand (PER). Bien que ces examens soient standardisés et à grande échelle, il s'agit de questions à développement ou d'items à réponse construite. Chaque épreuve est conçue par un groupe d'enseignants supervisé par le responsable pédagogique de la discipline. La passation est assurée par l'enseignant dans sa classe. Outre ces éléments partagés, deux types d'épreuves peuvent être considérés comme divergeant principalement quant à l'importance de leurs résultats pour les élèves et les enseignants : les épreuves de référence et les épreuves d'orientation.

Épreuves de référence

Ce type d'épreuve a été systématisé depuis le début des années 2000. Une seule discipline (mathématiques, français ou allemand) est évaluée à la fin des cycles de deux ans, en 4^e et en 6^e année, en juin. La passation est parfois facultative. L'enseignant corrige lui-même les épreuves selon un canevas de correction. Le résultat des épreuves est noté par l'enseignant selon un barème conçu a priori par l'administration (interprétation critériée). L'enseignant se sert librement de ce résultat pour l'évaluation finale de l'élève dans le bulletin. Les résultats n'ont donc pas de conséquences ou ont des conséquences très limitées sur le cursus de l'élève. L'enseignant a accès uniquement aux résultats de sa classe. Aucun *feedback* spécifique n'est prévu. Le retour des résultats à l'attention de l'administration est d'ailleurs le plus souvent facultatif. Les résultats n'ont pas non plus de conséquences directes sur l'enseignant, ne faisant pas l'objet d'un traitement spécifique par les autorités.

Épreuves d'orientation

Ce type d'épreuve est ancré depuis longtemps dans le paysage éducatif fribourgeois, soit dès les années 1980. Quatre disciplines (mathématiques, français, allemand et environnement) sont évaluées chaque année en 8^e et dernière année de la scolarité primaire, en mars-avril. La correction de l'épreuve se fait en majeure partie par des enseignants d'autres classes ; les enseignants corrigent une petite partie pour leurs propres élèves. Le résultat des épreuves est formulé sous forme de points et fixé selon un barème conçu a priori par l'administration (interprétation critériée). L'administration fixe ensuite une note après comparaison des résultats cantonaux (interprétation normative). Pour chaque discipline, l'enseignant reçoit une comparaison de la moyenne de sa classe à la moyenne cantonale.

Les résultats de ces épreuves jouent un rôle important dans le cursus de l'élève, déterminant en partie son orientation dans les trois filières d'études du secondaire⁵. Ils jouent aussi un rôle important pour la validation des pratiques des enseignants. Les résultats, présentés aux inspecteurs, aux directeurs d'école secondaire et aux parents, sont comparés à l'évaluation de l'enseignant pour ses élèves. Les résultats confirment ou non l'orientation proposée par l'enseignant sur la base du bulletin du premier semestre de 8^e année. Aucune sanction n'est cependant prévue pour les enseignants, même en cas d'écarts répétés.

Méthodologie

Pour mener à bien nos objectifs, nous avons fait le choix d'une approche qualitative et compréhensive. Notre but est de décrire un phénomène social ou humain complexe, sans toutefois suggérer la généralisation des résultats (Albarello, 2003 ; Miles & Huberman, 2010). L'échantillon a été formé afin d'interroger différents profils d'enseignants du canton de Fribourg. Toutefois, ces profils ne sont pas exhaustifs et ne sont donc pas totalement représentatifs de toute la population.

L'enquête a été menée auprès de 16 enseignants représentant toutes les années d'études du cursus primaire fribourgeois, de la 3^e à la 8^e année. Ils sont issus de 12 établissements de la partie francophone du canton de Fribourg. Un échantillon (voir Tableau 3) a été formé en variant le profil des enseignants (expérience professionnelle, genre, établissement, type d'épreuves externes qu'ils doivent faire passer tous les deux ans). Ces enseignants travaillent dans des établissements différents quant à leur localisation (ville ou campagne) et à leur taille : petit (au maximum un groupe-classe par année d'études), moyen (un à deux groupes-classes par année d'études) et grand (au moins deux groupes-classes par année d'études). Ces établissements sont situés dans quatre arrondissements (sur les huit existants), qui sont chacun pilotés par un inspecteur. Le rôle de directeur d'établissement n'existe pas au moment de l'enquête. Les inspecteurs des quatre arrondissements concernés par notre étude ont été interrogés afin qu'ils décrivent leur contexte et leur usage de l'évaluation externe. Pour ce dernier point, aucune différence significative ne peut être observée. L'usage est semblable à ce que nous décrivons précédemment pour les

deux types d'épreuves. Quatre conseillers pédagogiques qui sont superviseurs de la conception des épreuves des disciplines évaluées ont également été interrogés.

Les données recueillies sont issues de projets de recherche menés sur les effets de l'évaluation externe sur les pratiques des enseignants en général. Les entrevues avec huit enseignants de 7-8^{es} années ont été menées dans le cadre de notre thèse de doctorat (Yerly, 2014). Sept enseignants de 3-6^{es} années et un enseignant de 7-8^{es} années (Telma) ont été interrogés afin de compléter ces données. La grille de questions était la même pour toutes les entrevues. Les questions portaient sur le profil des enseignants, sur leurs perceptions des épreuves externes et sur les effets de celles-ci sur leurs pratiques didactiques. Pour ce dernier thème, trois axes de pratiques étaient étudiés : planification des contenus, enseignement en classe et évaluation des apprentissages. C'est ce dernier axe qui est exploité dans cet article. Les enseignants ont répondu individuellement à ces questions lors d'un entretien semi-directif d'environ une heure.

Notre démarche d'analyse est «inductive» (Deslauriers, 1997). Elle s'appuie sur l'analyse approfondie d'un certain nombre de cas pour arriver à la formulation d'explications causales quant à la question étudiée. Nos constats émergent donc des données recueillies, sans hypothèses a priori. Les entretiens ont été intégralement retranscrits, puis traités au moyen du logiciel Nvivo 10 pour mener une analyse de contenu sur le discours de chaque enseignant (intra-site), puis sur l'ensemble des données (inter-site; Miles & Huberman, 2010). Les résultats sont tout d'abord présentés, illustrés de quelques verbatims sélectionnés dans le discours des enseignants. Par la suite, les principaux constats sont confrontés et approfondis grâce au cadre théorique et empirique de l'article.

Tableau 3
Description de l'échantillon

Prénom fictif	Cycle	Expérience (années)	Genre	Arrondissement	Établissement	
Épreuve de référence (N=7)						
Isabelle	3-4 ^{es}	8	F	C	ville	grand
Emma	3-4 ^{es}	14	F	A	ville	moyen
Stéphanie	3-4 ^{es}	35	F	B	campagne	petit
Thérèse	3-4 ^{es}	40	F	D	campagne	petit
Rosalie	5-6 ^{es}	8	F	A	ville	petit
Germaine	5-6 ^{es}	22	F	C	campagne	moyen
Samuel	5-6 ^{es}	19	H	C	campagne	moyen
Épreuve d'orientation (N=9)						
Patrick	7-8 ^{es}	2	H	C	campagne	petit
Vincent	7-8 ^{es}	3	H	B	ville	moyen
Telma	7-8 ^{es}	6	F	D	ville	moyen
Aurélie	7-8 ^{es}	11	F	C	ville	grand
Chantal	7-8 ^{es}	13	F	A	ville	grand
Romain	7-8 ^{es}	14	H	B	campagne	petit
Julien	7-8 ^{es}	29	H	A	ville	grand
Ghislain	7-8 ^{es}	31	H	D	campagne	moyen
Claude	7-8 ^{es}	33	H	B	ville	moyen

Présentation des résultats

Tous les enseignants interrogés relèvent un certain impact de l'évaluation externe sur leurs pratiques d'évaluation des apprentissages des élèves. Plus précisément, ces effets portent sur leurs pratiques d'évaluation sommative. Les fonctions formative et diagnostique, pourtant mises en pratique par les enseignants interrogés, ne sont pas touchées. Selon la

définition d'Allal (2008), l'évaluation sommative permet d'établir un bilan qui certifie le degré d'acquisition de l'élève quant à une sélection d'objectifs ou de compétences après une séquence d'enseignement/apprentissage ou d'une étape d'un cursus de formation. Toute évaluation sommative suit une série de «démarches» (Allal, 2008) : 1) définir l'objet ou les objets à évaluer, 2) récolter les informations portant sur l'objet ou les objets auprès des élèves, avec ou sans outils, 3) interpréter les informations retenues, et 4) prendre des décisions et communiquer les appréciations aux élèves, aux parents, à l'administration, etc. Nous proposons d'analyser les effets des épreuves externes de manière spécifique sur ces quatre démarches. Pour chacune d'elles, nous tentons de répondre à nos objectifs en faisant d'abord une brève description de la façon dont les enseignants réalisent des ajustements dans leurs pratiques (*comment?*) et, surtout, en identifiant les raisons qui entraînent ou non ces ajustements (*pour quelles raisons?*).

Définition des objets d'évaluation sommative

Comment?

Les enseignants interrogés se servent des épreuves externes afin d'identifier les principales connaissances et compétences attendues des élèves aux différentes étapes de la scolarité (dans ce contexte, à la fin des cycles de deux ans). Les tâches ou exercices inscrits dans les épreuves sont pour eux une référence très concrète pour identifier les objets qui doivent être évalués au cours et à la fin d'un cycle. Ils vont alors y relever quelles sont les notions et/ou compétences qui sont à traiter en priorité.

Pour quelles raisons?

Pour ces enseignants, les épreuves traduisent de manière plus opérationnelle les attentes institutionnelles concernant les contenus et les compétences à évaluer (et à enseigner) :

Le PER, on peut l'interpréter comme on veut. Les épreuves, ça donne quand même une idée de ce qui est attendu. Je trouve qu'on est vraiment poussés. Par exemple, en français, dans l'expression et la lecture, comprendre des textes et écrire des textes. C'est ce qu'ils évaluent le plus, mais ce n'est pas forcément ce qu'on évaluait le plus avant. Ça a quand même changé ma pratique. [...] On fait maintenant d'office une évaluation de structuration et une évaluation de lecture-expression (Germaine, épreuve de référence).

En effet, dans le contexte étudié, le PER vient tout juste d'être implanté au moment de la recherche et est le plus souvent jugé par les enseignants comme complexe à opérationnaliser. L'épreuve externe devient alors pour certains, notamment pour les enseignants les moins expérimentés, un référent plus simple à suivre :

C'est vrai qu'on a des objectifs, on a le PER, mais c'est un peu vaste. [...] L'épreuve te permet de voir si tu es bien dans la ligne, si tu as bien touché ce qui est demandé à la fin du cycle. Ça te permet de voir si tu as tout traité (Isabelle, épreuve de référence).

Dans le cas de l'épreuve d'orientation de 8^e année, qui comprend des conséquences pour les élèves et les enseignants, l'évaluation externe peut même devenir la référence principale (ou unique) en ce qui concerne les objets à évaluer. Cela peut amener un certain rétrécissement du curriculum sur certains contenus ou compétences qui apparaissent régulièrement dans les épreuves :

Je pense qu'au début je me détachais plus de la PPO [nom de l'épreuve externe d'orientation]. En y réfléchissant, je pense que j'étais un peu naïf, mais pas dans un sens péjoratif. Il y avait des choses que je ne savais pas et je prenais plus de liberté. J'allais plus volontiers à gauche et à droite, et les élèves n'étaient pas forcément perdants. Avec la répétition des épreuves, je me méfie de tout en me disant qu'une année, ça va coïncider. Par la répétition des épreuves, ça me bloque un peu sur certains points du programme (Romain, épreuve d'orientation).

Enfin, si les enseignants interrogés alignent les objets qu'ils évaluent dans leurs classes sur les contenus des épreuves, c'est aussi car ils acceptent l'évaluation externe comme un « standard » qui, d'une part, garantit une certaine équité entre les classes et, d'autre part, facilite leur travail. Les attentes institutionnelles en matière de contenus sont claires et les élèves ont travaillé les mêmes contenus, même s'ils proviennent de classes différentes (p. ex., au début d'une année scolaire, lors de l'arrivée d'un nouvel élève en cours d'année) :

Je remarque qu'il y a énormément d'enseignants qui ne collaborent pas du tout, qui n'ont pas du tout les mêmes exigences. Nous recevons toujours des élèves qui viennent de deux classes différentes et nous remarquons beaucoup de différences, soit sur le plan de la tolérance des enseignants par rapport à des objectifs ou ce genre de choses. Donc, [l'évaluation externe] est un moyen d'égaliser et de se retrouver face à un critère qui est égal pour tous les élèves. Parce que les évaluations sont différentes, la façon d'enseigner est différente.

Je pense que ceux qui collaborent moins, c'est leur choix, mais avoir un critère commun, c'est une bonne chose. Ça nous évite d'avoir trop de différences dans les années d'études d'après. Nous avons remarqué assez souvent ce genre de différences (Aurélié, épreuve d'orientation).

Conceptions et usages d'outils d'évaluation sommative

Comment ?

Pour tous les enseignants interrogés, les épreuves externes ont un impact sur la conception ou l'usage d'outils d'évaluation sommative. C'est d'ailleurs un effet attendu par les autorités, les méthodes d'évaluation n'étant pas au moment de l'enquête inscrites ni dans le plan d'études, ni dans des directives. Les épreuves externes en fin d'année ou celles des années précédentes⁶ constituent alors un modèle pour la conception des instruments de collecte de l'information. Les enseignants interrogés sont amenés à modifier leurs tests sommatifs, prioritairement quant à leur questionnement et aux tâches demandées aux élèves, mais aussi quant aux conditions de passation (surtout le temps à la disposition des élèves) :

Par exemple, en allemand, ils mettent une image et un texte, et l'élève doit colorier sur les images. En fait, c'est peut-être des choses que je faisais, mais pas de la même façon. Enfin, c'est surtout sur le plan de la structure de l'exercice. Des fois, il y a des structures qui sont nouvelles ou d'autres approches. Ça donne des idées de tâches. Moi, je trouve que, dans mes évaluations, au bout d'un moment, je fais toujours la même chose. Ils arrivent à faire des exercices qui font apparaître d'autres savoir-faire que ceux que je fais dans mes évaluations. Ou d'autres façons de chercher. [...] Des fois, la façon dont eux me disent de poser les points peut m'aider quand je crée un exercice. D'essayer de repérer vraiment quels sont les éléments que je vais évaluer dans cet exercice (Rosalie, épreuve de référence).

Les enseignants utilisent les épreuves des années précédentes de deux manières différentes. Soit les enseignants s'en servent en tout ou en partie pour remplacer leurs outils d'évaluation sommative, soit ils s'en servent comme une source d'inspiration pour créer de nouveaux outils ou modifier les outils préexistants. Les enseignants, surtout dans le cas de l'épreuve de référence, voient même les épreuves comme un outil qui leur est « offert » par l'administration :

Mais je pense que c'est un outil qu'ils nous donnent. D'un côté, c'est bien sympa, je trouve. Ça me fait une évaluation de moins à préparer. Celle de maths, je l'ai trouvée affreusement longue, mais j'étais bien contente de l'avoir.

Sinon, je refais une évaluation en piquant des exercices et en cherchant d'autres trucs. Créer une nouvelle évaluation, ça me prend un temps monstre. Là, elle est faite. C'est un outil offert (Stéphanie, épreuve de référence).

Pour quelles raisons ?

Si la plupart des enseignants interrogés se servent des anciennes épreuves externes pour modifier leurs propres outils d'évaluation sommatrice, c'est surtout car ils relèvent leur qualité pédagogique. Ces épreuves sont un bilan pertinent des cycles d'apprentissage de deux ans. Elles parviennent à couvrir le programme des disciplines évaluées. L'intérêt repose également sur le fait que les épreuves touchent différentes compétences (p. ex., expression écrite, compréhension orale). Les enseignants relèvent aussi la qualité du matériel produit (p. ex., mise en page, illustrations). Enfin, les enseignants apprécient le fait que les épreuves correspondent aux pratiques du terrain et leur accordent une certaine expertise. Ils relèvent que les examens ont été construits par des enseignants encadrés par un conseiller pédagogique et sur une longue période :

Non seulement elles correspondent totalement avec ce qu'on voit en 7^e et 8^e, mais elles évaluent encore certaines compétences générales. Et ça, je trouve ça très bien. Je les trouve excellentes. [...] C'est vachement varié. Les questions ne sont pas des questions fermées ; ce ne sont pas des questions trop banales. Je dirais que c'est quand même assez des questions de réflexion qui demandent à l'élève d'aller chercher des choses au fond de lui-même, des choses qu'il a apprises. Moi, je les trouve franchement de qualité (Vincent, épreuve d'orientation).

Par contre, certains enseignants sont peu convaincus par le modèle évaluatif des épreuves. Dans notre échantillon, ils sont pourtant peu nombreux. Il s'agit d'enseignants très expérimentés : Claude (33 années) et Ghislain (31 années). Ils s'opposent surtout à une conception plus globale de l'évaluation des apprentissages et à l'évaluation de compétences au détriment des connaissances. Ces enseignants sont d'ailleurs ceux qui s'opposent également le plus à l'idée même d'évaluation externe. Ils considèrent que ce dispositif est avant tout destiné à contrôler les enseignants, et non à servir d'outil de développement professionnel.

La pression ressentie pour la réussite de leurs élèves, surtout pour l'épreuve d'orientation, peut également amener les enseignants, même les plus réticents, à adopter des types de questions ou d'exercices semblables à ceux des épreuves afin de préparer leurs élèves à la réussite des examens :

Non, je n'adapte pas mon évaluation à celle de la PPO [nom donné à l'épreuve externe d'orientation]. Mais je sais que je le devrais. [...] Sur le débat connaissances-compétences, je n'ai pas été très attentif à ça. Mais, si je regarde mes évaluations, c'est surtout question-réponse. Par exemple, quelle est la formule de l'eau? H₂O. [...] Aujourd'hui, c'est dommage, on n'ose plus parler de connaissances. On évalue seulement la débrouillardise, trouver la réponse à partir de documents, etc. [...] Mais je dois – je dois! – les préparer à avoir des évaluations longues. Par exemple, cette année, on a des épreuves de français qui durent 120 minutes. Alors, il faut que [les élèves] aient joué un match de 90 minutes avant le championnat. [...] Tu ne peux pas ne pas le faire, tu dois être déontologique. [...] Je suis très honnête, je fais le PP compatible (Claude, épreuve d'orientation).

Tableau 4
*Synthèse des ajustements observés sur les objets
et outils d'évaluation par discipline*

Disciplines	Changements/ajustements observés : <i>objets et outils</i>
Langue 1 (français)	<ul style="list-style-type: none"> - <i>Priorité mise sur la production et la compréhension de textes</i> - <i>Accentuation sur la compréhension de l'oral</i> - <i>Compétences du français évaluées de manière intégrée autour d'un genre textuel</i> - <i>Reprise des grilles de correction des épreuves pour les productions écrites</i>
Langue 2 (allemand)	<ul style="list-style-type: none"> - <i>Toutes les compétences langagières sont évaluées (compréhension/écriture; orale/écrite)</i> - <i>Tâches portant sur des situations de communication</i>
Mathématiques	<ul style="list-style-type: none"> - <i>Priorité mise sur la résolution de problèmes</i> - <i>Différents « thèmes » sont évalués de manière intégrée</i>
Environnement	<ul style="list-style-type: none"> - <i>Priorité mise sur l'« approche par compétences » ou les « savoir-faire »</i> - <i>Utilisation de documentation annexe par les élèves (graphiques, photographies, Internet, etc.)</i>

Le tableau 4 propose des exemples d'ajustement des objets et d'outils d'évaluation souvent cités par les enseignants pour les différentes disciplines évaluées dans les épreuves.

Interprétation des données d'évaluation sommative

Comment ?

La plupart des enseignants interrogés disent se servir des résultats des épreuves pour valider leur niveau d'exigence, mais de manière très générale. Les barèmes fixés par l'administration ainsi que les comparaisons entre la classe et une norme cantonale (pour les épreuves d'orientation) leur permettent de se rassurer ainsi que de confirmer globalement l'évaluation des apprentissages d'un élève et de leur groupe-classe. Toutefois, l'analyse du discours des enseignants montre que, lorsque ce travail est réalisé, il est très superficiel.

Pour quelles raisons ?

Tout d'abord, le faible usage des résultats des épreuves par les enseignants peut être expliqué par le manque d'occasions et/ou de temps à la disposition des enseignants. Aucun travail individuel ou collectif concernant ces données n'est exigé ou encouragé par les autorités. Les enseignants ne prennent que rarement du temps spécifique pour analyser les résultats des élèves aux épreuves, ni de manière individuelle, ni de manière collective. Ainsi, les enseignants ne se rencontrent pas spontanément pour discuter des résultats des épreuves ; ils en parlent « sur le pas de porte » de manière très générale. Ils se questionnent de manière très globale sur les résultats de leur classe en général. Le faible usage des résultats s'explique également par le fait que les enseignants leur donnent peu de valeur et d'utilité. Ainsi, pour les enseignants, les résultats pourraient avoir une incidence sur leurs pratiques uniquement en cas d'écart important observé, et ce, pour de nombreux élèves. Néanmoins, même dans ce cas apparemment très rare, les enseignants disent ne pas forcément remettre en cause leur niveau d'exigence. Les écarts individuels sont justifiés par la fatigue ou le stress des élèves ou par le degré de difficulté du test. Parfois même, les écarts importants sont justifiés par une trop grande standardisation des épreuves ou un niveau d'exigence trop élevé ou trop faible de l'épreuve :

Un élève peut arriver [à l'épreuve] avec une migraine, n'a pas dormi de la nuit, a des problèmes familiaux. Par contre, quand j'évalue les élèves sur un mois, deux mois, je les connais finalement très bien et je sais qui est capable de quoi. Je trouve plus valable que sur deux ans. Évaluer les compétences

d'un élève après un mois, elles sont réelles, elles sont là. Un élève qui a oublié quelque chose après six mois, il a juste oublié, mais ça ne veut pas dire qu'il n'a pas compris (Vincent, épreuve d'orientation).

Pour les épreuves de référence, certains enseignants interrogés vont même refaire une évaluation des mêmes contenus, voire modifier le barème officiel. Aussi, les résultats sont peu mobilisés par les enseignants, car les épreuves ont lieu en fin d'année et en fin de cycle, juste avant que les élèves ne changent de classe. Les enseignants ne pourront alors plus avoir d'influence sur les apprentissages des élèves concernés. Les praticiens interrogés considèrent souvent que les résultats des épreuves ne sont pas destinés à leur propre usage, mais plutôt aux élèves et aux parents (pour se situer, pour leur orientation) et à l'administration (pour le contrôle du système). Ce sentiment est d'ailleurs renforcé par le fait que les finalités des épreuves ne sont pas communiquées aux enseignants et qu'aucune rétroaction n'est proposée aux enseignants sur la base des résultats, ni individuellement, ni collectivement. En fin de compte, les enseignants font davantage confiance à leurs propres résultats d'évaluation et à leur connaissance des élèves :

Une fois l'épreuve passée, c'est basta, fini. Tu fermes l'enveloppe et c'est F-I-N-I. On ne reparle plus de cette période. Et évidemment, par moments, tu te dis : l'année prochaine peut-être que 28,5 points, ce sera 28. Et tu passes l'été et tu ne changes rien. [...] Par exemple, je sais que 28,5, ce sera toujours [la note] 6 chez moi. Je ne changerai pas, parce que je crois en des trucs et c'est vrai qu'au fil des années, finalement, je n'étais pas si faux (Claude, épreuve d'orientation).

Prise de décisions et communication des appréciations à la suite de l'évaluation sommative

Comment ?

Les enseignants se servent très peu des épreuves de référence pour prendre des décisions sur le passage des élèves dans l'année d'études ou le cycle suivant, par exemple. Ils leur donnent généralement le même poids qu'une évaluation sommative faite en classe. Le résultat des épreuves est le plus souvent « dilué » dans la moyenne des évaluations sommatives du semestre (de deux à cinq évaluations par discipline et par semestre). Comme nous l'avons détaillé et expliqué auparavant, même un écart important ne remettrait pas en cause leur jugement sur les apprentissages

des élèves. Aussi, les épreuves sont transmises aux parents dans le dossier d'évaluations avec le même statut que les évaluations sommatives de l'année.

Pour quelles raisons ?

Les épreuves de 8^e année, de par leur fonction d'orientation pour les élèves, jouent automatiquement un rôle de communication de l'évaluation vers l'extérieur de la classe. Un document destiné à l'enseignant, à l'élève, aux parents, à l'inspecteur et au directeur de l'école secondaire présente la comparaison des résultats de l'élève aux épreuves externes et ses notes du bulletin pour chaque discipline évaluée. Les résultats externes servent alors à confirmer ou non l'orientation de l'élève dans les filières du secondaire. L'enseignant ayant déjà transmis ses notes avant les épreuves, les résultats de cette dernière n'entrent donc pas formellement dans sa prise de décisions quant à l'orientation de l'élève, mais plutôt dans la validation de son évaluation. Toutefois, dans ce cas également, même un écart important entre les résultats ne semble pas entraîner les enseignants à remettre en cause leur jugement. Par contre, les résultats des deux types d'épreuves peuvent être utilisés parfois par les enseignants pour justifier leur évaluation en cas de besoin, par exemple pour appuyer des décisions remises en cause par des parents (p. ex., redoublement, orientation). Ils sont alors considérés comme une preuve « officielle » :

Ces élèves-là, qui ont vraiment des difficultés et qui devraient refaire [l'année], pour qui il faudrait une autre orientation. Si les parents ne sont pas d'accord, tu peux leur dire : ça, c'est l'évaluation officielle. Ça veut dire que ce n'est pas que mon enseignement qui fait que l'enfant ne réussit pas (Stéphanie, épreuve de référence).

Discussion des résultats

Selon la littérature, les enseignants se serviraient peu des épreuves externes, surtout dans une configuration d'accountability douce (Rozenwajn & Dumay, 2015). Pourtant, des enquêtes empiriques menées dans différents contextes identifient certains effets de l'évaluation externe non seulement sur les contenus évalués, mais aussi sur les méthodes d'évaluation utilisées en classe (Cattonar et al., 2010; Dutercq & Lanéelle, 2013; Longchamp, 2010; Tessaro, 2015; Yerly, 2014). Un intérêt de notre analyse est d'identifier les « démarches » (Allal, 2008) du processus évaluatif

qui sont touchées. Tous les enseignants interrogés se servent des épreuves externes comme référent pour la définition des objets à évaluer et pour la conception de leurs propres outils de collecte de l'information. Par contre, les démarches évaluatives liées aux résultats des épreuves (interprétation, prise de décisions et communication) sont très peu touchées. Dans le contexte étudié, les épreuves externes semblent donc avoir des effets importants sur les contenus et les méthodes d'évaluation. Ainsi, notre étude montre que les enseignants, lorsqu'il s'agit de l'évaluation des apprentissages, se servent des contenus et de la forme des épreuves externes, mais qu'ils mobilisent très peu ou pas les résultats des élèves.

Toutefois, notre étude approfondie permet de nuancer ce constat. Nos analyses montrent, comme dans d'autres recherches, que les changements observés dans les classes peuvent être plus ou moins profonds (Coburn, 2004) et ne sont pas toujours ceux attendus (Au, 2007 ; Mons, 2009 ; Rozenwajn & Dumay, 2015 ; Yerly, 2014). Deux cas de figure sont observés auprès des enseignants interrogés, souvent même conjointement dans le discours d'un même individu. D'un côté, les enseignants peuvent procéder à des modifications de leurs pratiques évaluatives au contact de l'évaluation externe. Toutefois, il ne s'agit pas de l'intégration de pratiques nouvelles, mais d'un ajustement de pratiques préexistantes. D'un autre côté, les enseignants peuvent procéder à une imitation des pratiques évaluatives attendues sans abandonner leurs pratiques préexistantes. Il s'agit alors d'un usage instrumental des contenus et des méthodes évaluatives des épreuves dans le but spécifique de préparer les élèves à réussir les examens. Les enseignants pratiquent plutôt une sorte de *coaching* et de drill sur ces méthodes et sur certains contenus ou certaines compétences (p. ex., enseignement pour l'examen, rétrécissement du curriculum).

Selon ces premiers constats, les effets des dispositifs en question sur les pratiques évaluatives des enseignants semblent donc plutôt modérés, comme dans d'autres contextes étudiés (MEN, 1994 ; Soussi et al., 2006). L'apport original de notre analyse est d'identifier les raisons de ces effets modérés en documentant et en confirmant des éléments plus théoriques (Coburn, 2004 ; Laveault et al., 2014). En effet, une épreuve externe ne peut pas à elle seule être une solution magique pour changer ou, du moins, ajuster les pratiques évaluatives du terrain. Son implantation sur le terrain dépend de sa réception par les enseignants. Selon nos analyses, l'usage ou non des épreuves externes par les enseignants interrogés s'explique,

indépendamment ou non, par 1) différentes postures des enseignants quant à l'évaluation externe et 2) leurs perceptions de certaines caractéristiques des dispositifs locaux.

Différentes postures des enseignants quant à l'évaluation externe

Dans le contexte étudié, l'usage ou non des épreuves externes par les enseignants s'explique en partie par le positionnement général des enseignants quant à l'évaluation externe. Néanmoins, ces postures sont complexes. Elles sont même souvent en tension dans le discours d'un même enseignant. D'un côté, les épreuves peuvent être reconnues par les enseignants comme un moyen de développement professionnel. Les enseignants s'en servent alors comme des « balises » qui leur permettent d'ajuster les objets qu'ils évaluent dans leurs classes et les méthodes d'évaluation qu'ils utilisent. D'un autre côté, même si les enseignants interrogés ne s'opposent généralement pas fortement à ces épreuves, ils peuvent les considérer davantage comme un contrôle hiérarchique qu'un moyen de développer leurs pratiques. Dans ce cas, l'évaluation externe va surtout pousser les enseignants à un usage plutôt instrumental de l'évaluation externe, voire occulter totalement son rôle de « balises ». L'objectif de l'enseignant est alors d'optimiser les résultats de sa classe afin d'échapper aux pressions hiérarchiques. En outre, les enseignants interrogés déclarent généralement que les finalités des différentes épreuves locales ne leur sont pas communiquées : Qui est évalué ? Pour quelles raisons ? Le message institutionnel n'étant pas clair, la tension entre développement professionnel et contrôle hiérarchique est renforcée. Elle est même plus marquée pour l'épreuve d'orientation : ses finalités sont multiples (orientation des élèves, régulation des pratiques) et la crainte d'un contrôle hiérarchique est plus forte du fait de la comparaison des résultats des épreuves avec ceux du bulletin de l'élève.

Perceptions des enseignants de certaines caractéristiques des dispositifs locaux

Selon nos analyses, l'usage ou non des épreuves externes dans les classes s'explique également par les perceptions des enseignants quant à certaines caractéristiques des dispositifs. La plupart des ajustements réalisés par les enseignants interrogés s'expliquent en premier lieu par la pertinence pédagogique que les enseignants reconnaissent à la forme et aux contenus des examens, renforcée par leur proximité avec les pratiques du

terrain. À l'inverse, la majorité des enseignants interrogés mettent en doute la pertinence des résultats des épreuves à grande échelle. Pour eux, ils ne sont généralement pas un indicateur pertinent de l'apprentissage des élèves. Ils sont notamment marqués de nombreux biais (p. ex., état de l'élève durant la passation, effets du milieu). De ce fait, les résultats sont peu mobilisés par les enseignants pour l'évaluation des apprentissages des élèves et pour ajuster leurs pratiques évaluatives. En outre, dans le contexte étudié, aucun *feedback* n'est apporté à la suite des résultats des élèves, ni de manière individuelle (p. ex., analyse critériée et détaillée des résultats pour chaque classe), ni collective (p. ex., présentation/discussion des résultats au sein de l'établissement, formation continue). Ainsi, les enseignants, ni accompagnés ni encouragés, interprètent et utilisent peu les résultats des épreuves. Enfin, les conséquences que comportent les épreuves pour les élèves, surtout pour celles d'orientation, expliquent également l'usage instrumental de l'évaluation externe. Les enjeux, ressentis de manière variable, poussent les enseignants à des pratiques de préparation spécifique au test afin de faciliter la réussite de leurs élèves.

Conclusion

Notre étude interroge directement les enseignants afin de mieux comprendre les effets des épreuves externes sur leurs pratiques d'évaluation des apprentissages. Bien que ces données autorapportées soient inévitablement marquées par une certaine subjectivité et que l'étude porte sur un seul contexte, les résultats présentés apportent non seulement certaines confirmations, mais aussi de nouveaux éléments de compréhension.

Comme il a été avancé dans d'autres contributions (notamment Crahay, 2009 ; Laveault et al., 2014 ; Tessaro, 2015), notre analyse montre qu'une articulation entre évaluation externe et évaluation en classe n'est « pas impossible ». Les épreuves externes peuvent avoir, sous certaines conditions, des effets positifs sur les pratiques d'évaluation des apprentissages des élèves en classe. Toutefois, pour ce faire, il semble nécessaire que les enseignants perçoivent, dans ces dispositifs, l'opportunité de développer leurs pratiques, et non un contrôle hiérarchique de leurs compétences. Si les dispositifs sont conçus dans une logique de complémentarité entre évaluation externe et évaluation en classe, ils devraient permettre aux ensei-

gnants de développer leurs compétences et leurs propres outils en matière d'évaluation des apprentissages, sans toutefois provoquer ni de résistance ni de dépendance.

Finalement, chaque point développé dans cet article mériterait une étude plus approfondie. Il semble surtout important de comprendre le faible usage des résultats d'épreuves externes par les enseignants pour leurs pratiques d'enseignement et d'évaluation. Il serait notamment nécessaire de différencier davantage les raisons de ce découplage qui sont liées aux enseignants et/ou aux dispositifs d'évaluation externe. Pour ce faire, il semble pertinent de mener des études plus larges et notamment d'ouvrir les analyses à divers contextes éducatifs comprenant différentes configurations d'accountability où diffèrent les caractéristiques et les usages institutionnels d'épreuves externes.

Réception : 31 avril 2016

Version finale : 1 mars 2017

Acceptation : 10 mars 2017

NOTES

1. Dans la francophonie, les termes « obligation de résultats », « reddition de comptes » ou « responsabilisation » sont parfois préférés à celui d'« accountability ». Dans ce texte, nous préférons utiliser ce dernier terme, qui englobe toutes ces différentes significations.
2. Lors de la rédaction de cet article en 2016, des épreuves externes nationales (Suisse) et régionales (région linguistique francophone) étaient en cours d'élaboration.
3. La durée totale de la scolarité obligatoire (11 années) a été harmonisée entre les cantons suisses à la suite du concordat HarmoS. Néanmoins, la durée des cursus primaires et secondaires est encore différente selon les cantons.
4. Il est important de noter qu'au moment de l'enquête, aucun document officiel détaillant les finalités des épreuves cantonales fribourgeoises n'a été publié ou transmis aux enseignants. Les finalités décrites sont issues d'entretiens avec les responsables des épreuves externes (inspecteurs et conseillers pédagogiques).
5. À la fin de la scolarité primaire (8^e année), les élèves sont orientés dans des filières d'études : exigences de base, générale et prégymnasiale. Ces dernières vont déterminer leur parcours scolaire et professionnel (études postobligatoires, formation professionnelle, etc.). L'épreuve d'orientation fait partie de la Procédure cantonale de préorientation (PPO) des élèves du secondaire. L'orientation des élèves dépend d'indicateurs qui décrivent, premièrement, le degré d'acquisition des connaissances et des compétences de l'élève (les résultats de l'élève au premier semestre de la 8^e année, confirmés par ceux de l'épreuve externe) et, deuxièmement, son attitude envers les apprentissages scolaires (évaluée par l'enseignant, les parents et l'élève).
6. Les épreuves des années précédentes sont accessibles aux enseignants sur une plateforme Internet.

RÉFÉRENCES

- Albarello, L. (2003). *Apprendre à chercher : l'acteur social et la recherche scientifique*. Bruxelles : De Boeck.
- Allal, L. (2008). Évaluation des apprentissages. Dans A. van Zanten (dir.), *Dictionnaire de l'éducation* (pp. 311-314). Paris : Presses universitaires de France.
- Amrein, A., & Berliner, D. (2002). *An analysis of some unintended and negative consequences of high-stakes testing*. East Lansing, MI : Great Lakes Center for Education Research and Practice.
- Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267. doi: 10.3102/0013189X07306523
- Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331. doi: 10.3102/01623737024004305
- Cattonar, B., Dumay, X., Mangez, C. et Maroy, C. (2010). *Évaluations externes dans l'enseignement primaire en Belgique francophone : réceptions et usages d'outils de régulation basés sur les connaissances*. Bruxelles : European Commission, Sixth Framework Program.

- Centre suisse de coordination pour la recherche en éducation (CSRE). (2014). *L'éducation en Suisse: rapport 2014*. Aarau, Suisse: CSRE.
- Coburn, C. E. (2004). Beyond decoupling: Rethinking the relationship between the institutional environment and the classroom. *Sociology of Education*, 77(3), 211-244. doi:10.1177/0038040704077
- Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP). (2007). *Accord intercantonal sur l'harmonisation de la scolarité obligatoire (concordat HarmoS)*. Berne: CDIP.
- Crahay, M. (2009). Articuler l'évaluation en classe et le pilotage des systèmes, est-ce possible? Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension: entre la régulation des apprentissages et le pilotage des systèmes* (pp. 233-251). Bruxelles: De Boeck.
- Demailly, L. (2001). *Évaluer les politiques éducatives: sens, enjeux, pratiques*. Bruxelles: De Boeck Supérieur.
- Deslauriers, J.-P. (1997). L'induction analytique. Dans J. Poupard, J.-P. Deslauriers, L.-H. Groulx, A. Laperrière, R. Mayer et A. Pires (dir.), *La recherche qualitative: enjeux épistémologiques et méthodologiques* (pp. 293-308). Montréal: Gaëtan Morin.
- Diamond, J. B. (2012). Accountability policy, school organization, and classroom practice: Partial recoupling and educational opportunity. *Education and Urban Society*, 44(2), 151-182. doi: 10.1177/0013124511431569
- 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, Suisse: IRDP.
- Dutercq, Y. et Lanéelle, X. (2013). La dispute autour des évaluations des élèves dans l'enseignement français du premier degré. *Sociologie*, 1(4). Repéré à <https://sociologie.revues.org/1439>
- Eurydice. (2009). *Les évaluations standardisées des élèves en Europe: objectifs, organisation et utilisation des résultats*. Bruxelles: Agence exécutive Éducation, audiovisuel et culture (EACEA P9 Eurydice).
- Institut de recherche et de documentation pédagogique (IRDP). (2015). *Pratiques cantonales concernant l'organisation d'épreuves de référence, d'examens, de tests ainsi que l'obtention de certificats ou diplômes*. Neuchâtel, Suisse: IRDP.
- Kempf, A. (2016). *The pedagogy of standardized testing: The radical impacts of educational standardization in the US and Canada*. New York: Palgrave Macmillan.
- Lafontaine, D., Soussi, A. et Nidegger, C. (2009). Évaluations internationales et/ou épreuves nationales: tensions et changement de pratiques. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension: entre la régulation des apprentissages et le pilotage des systèmes* (pp. 61-80). Bruxelles: De Boeck.
- Laveault, D., Dionne, É., Lafontaine, D., Tessaro, W. et Allal, L. (2014). L'évaluation des compétences en salle de classe: régulations entre politiques et pratiques. Dans C. Dierendonck, E. Loarer et B. Rey (dir.), *L'évaluation des compétences en milieu scolaire et en milieu professionnel* (pp. 85-96). Bruxelles: De Boeck.
- Longchamp, A.-L. (2010). *Épreuves cantonales de référence de mathématiques en 6^e année: point de vue d'enseignants sur l'utilisation du dispositif* (n° 143). Lausanne: Unité de recherche pour le pilotage des systèmes pédagogiques.

- Maroy, C. (2013). Politiques et outils de «l'école de la performance»: accountability, régulation par les résultats et pilotage. Dans C. Maroy (dir.), *L'école à l'épreuve de la performance: les politiques de régulation par les résultats* (pp. 14-31). Bruxelles: De Boeck.
- Miles, M. et Huberman, A. M. (2010). *Analyse des données qualitatives*. Bruxelles: De Boeck Université.
- Ministère de l'Éducation nationale (MEN). (1994). *L'opinion des enseignants sur l'évaluation dans l'enseignement primaire*. Note d'information. Paris: MEN, Direction de l'évaluation et de la prospective.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie – Recherches en éducation*, 169, 99-140. doi: 10.4000/rfp.1531
- Mottier Lopez, L. (2009). L'évaluation en éducation: des tensions aux controverses. Dans L. Mottier Lopez et M. Crahay (dir.), *Évaluations en tension: entre la régulation des apprentissages et le pilotage des systèmes* (pp. 7-25). Bruxelles: De Boeck.
- Mottier Lopez, L. et Crahay, M. (2009). *Évaluations en tension: entre la régulation des apprentissages et le pilotage des systèmes*. Bruxelles: De Boeck.
- Ntamakiliro, L. et Tessaro, W. (2010). Plus d'épreuves externes: vers une standardisation de l'évaluation scolaire? Dans P. Gilliéron Giroud et L. Ntamakiliro (dir.), *Réformer l'évaluation scolaire: mission impossible?* (pp. 165-185). Berne: Peter Lang.
- Rozenwajn, E. et Dumay, X. (2015). Les effets de l'évaluation externe sur les pratiques enseignantes: une revue de la littérature. *Revue française de pédagogie*, 189(4), 105-138. doi: 10.4000/rfp.4636
- Soussi, A., Ducrey, F., Ferrez, E., Nidegger, C. et Viry, G. (2006). Pratiques d'évaluation: ce qu'en disent les enseignants (à l'école obligatoire et dans l'enseignement postobligatoire général). *Note d'information du SRED*, n° 28, Genève: Département de l'instruction publique, Service de la recherche en éducation. Repéré à www.ge.ch/recherche-education/doc/publications/notesinfo/notes-sred-28.pdf
- Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387-431. doi: 10.3102/00346543072003387
- Tessaro, W. (2015). Les épreuves externes comme soutien au développement professionnel des enseignants dans leurs pratiques d'évaluation certificative: les apports d'une recherche collaborative. *Évaluer – Journal international de recherche en éducation et formation*, 1(2), 49-65. Repéré à <http://le-jref.org/index.php?id=102&file=1>
- Yerly, G. (2014). *Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignants: analyse du regard des enseignants du primaire* (Thèse de doctorat). Université de Fribourg, Fribourg, Suisse.
- Yerly, G. (2017). Quel impact des politiques de responsabilisation douce sur les pratiques enseignantes? Dans Y. Dutercq et C. Maroy (dir.) *Le travail et le professionnalisme enseignant face aux politiques de responsabilisation* (pp. 121-141). Bruxelles: De Boeck Supérieur.