

DOCUMENTER LE JUGEMENT DES ENSEIGNANTS EN LANGUE SECONDE AU 3^E CYCLE DU PRIMAIRE AU QUÉBEC DANS LE CADRE DU BILAN DES APPRENTISSAGES DE FIN DE CYCLE

MICHELINE-JOANNE DURAND *Université de Montréal*

RÉSUMÉ. Cette étude vise à décrire les pratiques évaluatives d'enseignantes en langue seconde en fin de scolarité à l'école primaire au Québec. Elle a pour but de cerner la prise en compte des résultats à l'épreuve ministérielle et des résultats obtenus tout au cours du cycle lors du jugement porté au bilan des apprentissages. Deux études de cas serviront à documenter leur jugement professionnel face à la tension lors de la confrontation de résultats divergents.

DOCUMENTING SECOND LANGUAGE TEACHER JUDGMENT IN ELEMENTARY CYCLE THREE WITHIN THE FRAMEWORK OF THE END-OF-CYCLE REPORT

ABSTRACT. The aim of this study is to describe second language teachers' evaluation practices at the completion of elementary school in Quebec. The study compares results obtained from the Ministry evaluation process with results obtained throughout the cycle. Two case studies serve to document teachers' professional judgment when confronted with differing results.

INTRODUCTION

À la fin de l'année scolaire 2007, le ministère de l'Éducation, des Loisirs et du Sport (MELS) du Québec a introduit une épreuve obligatoire en anglais, langue seconde, à tous les élèves qui terminaient leur scolarité à l'ordre d'enseignement primaire. Cette épreuve s'adressait tant aux élèves des groupes réguliers (qui bénéficient d'une ou deux périodes d'enseignement de 30 min à 50 minutes par semaine) qu'à une clientèle d'enfants auxquels différents projets d'enrichissement du programme d'anglais langue seconde (PEPALS) sont proposés. Notre étude porte sur la façon dont deux enseignantes de ces classes ont fondé leur jugement pour établir le bilan de fin de cycle de chaque élève lorsque les résultats de l'épreuve ministérielle diffèrent des résultats obtenus tout au long de l'année scolaire. La problématique est celle de la confrontation du jugement professionnel que doit poser l'enseignant lorsque ses évaluations en cours d'année ne correspondent pas aux résultats de l'épreuve externe. Cette situation est nouvelle : jusqu'ici, le jugement des enseignants ne reposait que sur leurs propres évaluations. Notre étude décrit l'instrumentation qui sous-

tend les pratiques évaluatives au quotidien et l'épreuve obligatoire du MELS. Nous montrons comment nous avons établi les écarts dans les résultats entre les évaluations internes et externes. Par la suite, nous présentons les entretiens des enseignants exposant comment ils collectent l'information, l'interprètent et portent leur jugement ainsi que la prise en compte de l'évaluation externe (l'épreuve du MELS) dans du bilan des apprentissages. Deux études de cas montrent comment des enseignants s'y prennent pour porter leur jugement et la justification lorsque les résultats aux différentes épreuves divergent. Nous terminons en posant les limites de cette recherche.

PROBLÉMATIQUE

On s'entend pour considérer que l'évaluation comporte plusieurs étapes dont la collecte de données et le jugement. On sait aussi qu'une grande partie des travaux du domaine a porté sur la collecte des données simples, produisant des mesures précises et valides de telle sorte que le jugement se porte à partir d'un cumul de mesures. Le problème se pose quand le jugement ne repose pas, et plus, ne peut pas reposer, sur le cumul et la moyenne (additions, pourcentage) de résultats homogènes provenant de sources différentes. Analogiquement, on peut dire qu'au tribunal, le jugement doit s'appuyer sur des témoignages fidèles et valides. Or, c'est comme si l'essentiel des préoccupations avait jusqu'ici porté sur la critique du témoignage et peu sur le processus du jugement. Comment les enseignants s'y prennent-ils alors pour porter leur jugement lorsque celui-ci ne se fait pas à partir d'un cumul de mesures ? Est-ce que cette façon diffère que l'on soit en cours ou en fin de cycle ?

La littérature scientifique examinée comporte surtout des travaux qui portent sur l'accord inter juge ou qui examinent les biais qui contaminent le jugement. D'autres recherches récentes se sont intéressées au jugement professionnel de l'enseignant. L'accord inter juge est défini comme la conformité de deux ou plusieurs jugements qui se rapportent au même objet, à une même épreuve. Les études des biais liés au jugement se retrouvent chez plusieurs auteurs (Bressoux et Pansu, 2003; Jorro, 2000; Louis, 1999) qui traitent de la façon dont les enseignants notent les travaux des élèves. Le jugement professionnel de l'enseignant s'exerce à toutes les étapes de la démarche d'évaluation et s'appuie sur une collecte bien documentée composée d'informations pertinentes et suffisantes (Allal et Mottier Lopez, 2007, 2009; Lafortune et Allal, 2008; Lavault, 2008; Legendre, 2001; MEQ, 2003). Par contre, l'accord intra juge lorsque l'enseignant confronte des résultats tirés d'épreuves internes différents de ceux d'une épreuve externe est peu documenté. Nous allons aborder cette question de la conformité du jugement de l'enseignant à partir d'épreuves internes réalisées tout au long du cycle d'apprentissage avec une épreuve externe réalisée en fin de cycle.

Le système éducatif québécois associe la certification, reliée à la sanction des études, uniquement à l'ordre secondaire. À l'ordre d'enseignement primaire, on parle de reconnaissance des acquis. Celle-ci prend la forme d'un bilan des apprentissages à la fin de deux années de scolarité, soit à la fin de chacun des trois cycles du primaire. Pour réaliser ce bilan, l'enseignant utilise les nombreuses observations consignées tout au long de l'année ainsi que des épreuves externes, lorsque celles-ci sont disponibles.

Le bilan des apprentissages

Le bilan des apprentissages n'est pas un bilan sommatif consistant en un cumul des résultats des évaluations formelles effectuées au cours de l'année (Legendre, 2001). Il indique le niveau de développement de la compétence atteint par l'élève. Pour ce faire, l'enseignant exerce un jugement, c'est-à-dire qu'il établit un profil du niveau de développement atteint par l'élève pour chacune des compétences. Ce profil est beaucoup plus instructif qu'une moyenne parce qu'il permet de cerner précisément l'état de développement des compétences de l'élève et de déterminer ses forces et ses besoins afin de lui offrir les mesures de soutien ou d'enrichissement appropriées à la poursuite de ses apprentissages (Laurier, Tousignant, & Morissette, 2005). Il sert également à faire le point avec l'élève et avec ses parents, tout en donnant des informations pertinentes aux enseignants du cycle suivant (Durand & Chouinard, 2006). Pour communiquer ce profil, une échelle des niveaux de compétence (MEQ 2002b) est construite. Dans la présente recherche, l'échelle illustrée au tableau 1 a été utilisée pour porter le jugement global et pour communiquer le résultat du bilan des apprentissages pour chacune des trois compétences en anglais, langue seconde soit :

C1 : To interact orally in English.

C2 : To reinvest understanding of oral and written texts.

C3 : To write texts.

TABLEAU 1. *Cotes et échelons utilisés pour le bilan des apprentissages, à la fin du cycle*

Cotes	Évaluation des compétences au bilan des apprentissages	Échelons
N	Ne répond pas aux attentes de fin de cycle	3
P	Répond en partie aux attentes de fin de cycle	4
R	Répond aux attentes de fin de cycle	5

Pour choisir la cote la plus appropriée à la situation observée, les enseignants se réfèrent aux attentes de fin du troisième cycle décrites dans le Programme de formation de l'école québécoise (PFÉQ), (MEQ, 2002a). Ces attentes sont des niveaux de compétence que devraient atteindre les élèves à la sortie du

cycle. Au troisième cycle, cette attente correspond à l'échelon 5 du document ministériel *Échelles de niveau de compétences* pour l'enseignement primaire (MEQ, 2002b). Par exemple, pour la compétence *Écrire des textes*, les élèves du 3^e cycle qui répondent partiellement aux attentes (P) se situent à l'échelon 4 et ceux qui répondent totalement aux attentes (R) à l'échelon 5. On situe un élève qui ne répond pas aux attentes du 3^e cycle (N) à partir des échelons établis pour le 2^e cycle. Depuis 2002 (soit de l'avènement du nouveau programme par compétences) les enseignants portaient leur jugement au bilan des apprentissages de fin de cycle à partir de ces côtes et selon des tâches variées dont le choix dépendait de leur propre responsabilité.

Épreuve ministérielle

Pour des fins de régulation du système et afin de cerner le niveau d'appropriation du PFEQ, le ministère de l'Éducation des Loisirs et du Sport (MELS) fournit des épreuves nationales obligatoires à la fin du 3^e cycle du primaire pour le français et les mathématiques, et ce, depuis 2004. L'enseignant ne peut cependant pas tenir compte de cette seule épreuve pour porter un jugement au bilan des apprentissages. Or, pour évaluer une compétence, il faut plusieurs tâches. Toutefois, depuis l'avènement du nouveau programme par compétences en anglais langue seconde, aucune instrumentation n'avait été proposée aux enseignants.

En juin 2007, les enseignants ont pu expérimenter un modèle national d'évaluation, à l'état de prototype. Dans cette recherche, l'épreuve externe, *What I Really Want and Need From Adult Is...*, a permis d'évaluer deux compétences : la compétence *Interagir oralement en anglais* (C1) et la compétence *Écrire des textes* (C3). Cette épreuve externe cible le langage prescrit dans la section «Savoirs essentiels » du PFEQ des 2^e et 3^e cycles en anglais langue seconde. On y propose huit tâches interreliées qui se déroulent parfois en équipe de deux, parfois en plus grand groupe et finalement de façon individuelle pour la production finale. Celle-ci consiste à donner l'occasion aux jeunes d'exprimer leurs souhaits et leurs besoins à un adulte sous la forme d'une lettre personnelle. Des rubriques d'autoévaluation sont incluses dans le Cahier de l'élève, ce qui lui permet de s'autoréguler pendant l'apprentissage et donc de mieux démontrer ses compétences. Une échelle descriptive comprenant trois critères et quatre échelons (A, B, C, D) permet à l'enseignant de situer le niveau de compétence de l'élève. Ces résultats (par exemple A, B, A) sont par la suite convertis en une note globale, toujours exprimée sous la forme d'une lettre (par exemple, A).

Ce sont les résultats obtenus à cette épreuve (note globale A, B, C ou D), ceux obtenus au cours des deux années du cycle (note au bulletin 1,2 ,3 et 4 traduit par A, B. C ou D) et le jugement porté par les enseignantes au bilan de fin de cycle (R, P ou N) qui forment la base des données de cette étude. Dans la

prochaine section, nous situons le jugement au regard de sa construction, de la collecte de données et des décisions qui en découlent.

CADRE THÉORIQUE

Construction du jugement

Allal et Mottier Lopez (2007) définissent le jugement comme un acte cognitif individuel qui s'élabore et s'intègre dans un réseau de communication au sein d'un contexte social et institutionnel. Il implique une mise en relation entre les exigences du milieu, les valeurs personnelles et les stratégies pédagogiques de l'enseignant. Pour leur part, Lafortune et Allal (2007), suite à deux études réalisées à Genève et au Québec, caractérisent le jugement, entre autres, par le fait de réunir et de confronter des sources diverses, de les combiner à travers un raisonnement interprétatif et de s'appuyer sur des procédures d'évaluation que l'enseignant élabore et applique à l'ensemble de ses élèves. Legendre (2001) ajoute que le jugement professionnel doit être instrumenté, évolutif, fondé sur l'autonomie et la responsabilité, éthique et conforme à un ensemble de valeurs. Dans un contexte d'évaluation de compétences, articulé autour de tâches complexes, le jugement devrait permettre d'expliquer plus en profondeur les performances des élèves, dans le but de discriminer leur position parmi les différents niveaux de performance (Scallon, 2004; Wiggins, 1998). Comme l'affirme Legendre (2001) « le jugement professionnel de l'enseignant prend appui sur « la compétence à observer en situation » (p.18). Scallon (2004) ajoute que : « Mener une évaluation consiste donc à dégager un jugement d'ensemble, qui n'est pas la simple somme de ses composantes, tout en tenant compte de plusieurs habiletés et d'un ensemble de compétences. » (p. 20). Parce que porter un jugement, c'est avoir un regard global et évolutif sur la capacité de l'élève et sur sa progression, ce qui nécessite une appréciation instrumentée et bien documentée.

Jugement et la collecte de données

Dans une visée d'évaluation formative, l'enseignant s'assure de recueillir tout au long du cycle, des données pertinentes (la collecte) et d'être en mesure de leur donner un sens (l'interprétation) afin de juger du développement des compétences sur la base des observations notées. Pour recueillir ces données, l'enseignant place l'élève dans des tâches suffisamment nombreuses et diversifiées afin de lui permettre de montrer ses compétences, car la compétence ne s'évalue pas à partir d'une seule situation. Il détermine d'abord des cibles et les différencie au besoin, il élabore des outils adéquats, il réunit et consigne l'information en quantité suffisante et l'interprète à la lumière de références pertinentes. Ces mêmes conditions menant à un jugement éclairé prévalent que l'on se trouve en cours de cycle ou en fin de cycle (Durand & Chouinard, 2006). Cependant, à la fin du cycle, les tâches sont généralement conçues par

le gouvernement ou par l'administration d'un territoire scolaire pour être appliquées à une très grande population d'élèves. Elles servent à réguler le système et à faire le point sur le développement des compétences des élèves à la fin d'un cycle. Ce sont des tâches complexes qui invitent les élèves à résoudre un problème lié principalement à une seule compétence. Le problème tend à circonscrire l'ensemble des composantes de la compétence ainsi que les préoccupations de la majorité des élèves d'un territoire donné. De ce fait, sa structure est plus rigide, la durée est prédéfinie et les adaptations sont limitées et balisées. De plus, des outils d'évaluation formels sont fournis et des indications sont données pour l'interprétation des résultats, comme des exemples-types. Moins contextualisées et signifiantes que celles qui ont été élaborées en salle de classe parce que s'adressant à une plus large clientèle, ces épreuves n'en demeurent pas moins des situations complexes qui visent l'évaluation d'une compétence (Durand & Chouinard, 2006). Harlen (2007) explique qu'une épreuve externe de ce genre doit être un outil fidèle, puisqu'on en attend qu'il donne des résultats nationaux similaires, peu importe l'évaluateur qui l'utilise et l'école dans laquelle l'épreuve est administrée. L'effort doit donc être mis sur la fidélité de l'interprétation des traces de l'élève recueillies.

Jugement et décisions

Le jugement, au centre de la démarche évaluative, amène l'enseignant à réguler ses actions pédagogiques afin de favoriser l'apprentissage chez ses élèves. Le jugement est « cette étape centrale de la démarche d'évaluation pédagogique qui consiste à donner un avis sur la progression ou l'état de réalisation des apprentissages, à la lumière de diverses informations recueillies » (Legendre, 2005, p. 816). Il revêt aussi un caractère définitif lorsqu'il est porté au terme d'un cycle d'apprentissage, car des décisions administratives seront prises en lien avec le classement des élèves dans les différents parcours offerts par l'école secondaire. Lunz, Sthal, et Wright (1994) expliquent que l'un des buts de l'évaluation d'une performance devrait être la reproductibilité de la décision (*decision reproductibility*, p. 914), cette dernière reposant sur la comparabilité des résultats à partir de laquelle une décision est portée. À l'instar de ces propos, Jorro (2000) prétend que nous possédons un héritage de la pensée juste : l'enseignant veut s'assurer à tout prix de l'objectivité de ses évaluations. Selon elle, cette ligne de pensée est mythique et piège l'évaluateur. Gerard (2002) abonde en ce sens lorsqu'il mentionne que « ce qu'il est nécessaire d'éviter n'est pas la subjectivité de l'évaluation, mais bien l'arbitraire » (p. 1).

Dès lors, comment l'enseignant s'y prend-il lorsque les résultats de l'élève sont différents à l'épreuve ministérielle (en fin de cycle) des résultats consignés tout au long de l'année (en cours de cycle) ? Sur quoi se base-t-il pour porter son jugement final et attribuer une notation au bilan de fin de cycle ? Sur le cumul des résultats ? Sur le niveau actuel de maîtrise de la compétence ? Sur une combinaison d'observations ?

BUT DE LA RECHERCHE

Ce projet de recherche a permis de documenter le jugement des enseignants et de dresser leur portrait d'enseignant-évaluateur.

Pour ce faire, nous avons examiné de façon spécifique :

- 1) les écarts entre les résultats des élèves à l'épreuve ministérielle et ceux obtenus en contexte d'évaluation pendant le cycle, au bulletin ;
- 2) les pratiques évaluatives des enseignants en regard de l'instrumentation utilisée et du jugement porté au bulletin en cours de cycle, ou ;
- 3) les pratiques évaluatives des enseignants en regard du jugement porté au bilan des apprentissages, ainsi que ;
- 4) les régulations pédagogiques prévues suite à cette première passation de l'épreuve externe.

Les choix méthodologiques qui suivent identifient les procédures utilisées pour collecter et analyser ces différentes données.

MÉTHODOLOGIE

Cette recherche exploratoire a permis dans un premier temps de comparer les résultats des élèves au bulletin à ceux de ce premier examen ministériel. Il a été possible ensuite de générer des constats sur l'instrumentation utilisée par les enseignants, sur la prise en compte de ces différents résultats lors du jugement porté au bilan des apprentissages et sur les régulations envisagées. Nous avons opté pour une méthodologie d'étude de cas recourant à l'entrevue semi-structurée ainsi qu'au questionnaire écrit, croisant ainsi des données suscitées et provoquées (Van der Maren, 2003). L'évaluation formative et le bilan des apprentissages étant des processus dynamiques, l'étude de cas favorise la prise en compte de multiples facteurs et devrait plus facilement permettre de repérer les tensions qui se présentent lorsque l'enseignant doit poser son jugement (Lamoureux, 2006). Nous décrivons maintenant les principaux paramètres à la base de la méthodologie utilisée dans ce projet de recherche, soient les participants, la collecte et l'analyse des données

Les participants

Deux enseignantes ont accepté de participer de façon volontaire à cette étude. Alice¹ est titulaire au primaire en anglais, langue seconde. Elle enseigne en sixième année dans un projet d'enrichissement du programme d'anglais langue seconde (PEPALS), modèle intensif, où les élèves reçoivent un enseignement pendant trois mois consécutifs (24 périodes par semaine). Ses élèves (échantillon 1) proviennent d'une école située dans un milieu socio-économique moyennement faible et certains d'entre eux présentent des difficultés particulières relevant de la psychopathologie, des troubles d'apprentissage ou des troubles

de comportement. Elle participe à des groupes de recherche et de développement qui ont élaboré, expérimenté et validé des situations d'apprentissage et d'évaluation, des grilles d'observation et des outils pour le bilan des apprentissages.

La deuxième enseignante, Bettie, spécialiste en anglais, accumule plusieurs années d'expérience. Elle enseigne en cinquième et en sixième année dans un projet d'enrichissement du programme d'anglais langue seconde (PEPALS) modèle enrichi, où les élèves reçoivent huit périodes de 54 minutes par semaine, échelonnées sur toute l'année. Ses élèves (échantillons 2 et 3) reçoivent son enseignement pendant deux ans et proviennent d'une école située dans un milieu socio-économique plus aisé et peu d'entre eux se retrouvent en grande difficulté d'apprentissage.

À la fin de l'année, les élèves de ces trois échantillons devraient se retrouver en général au cinquième échelon pour la compétence *To write texts*, ce qui correspond aux attentes terminales de la fin du 3^e cycle et à la cote R (répond aux attentes) inscrites au bilan des apprentissages.

Instrumentation

Pour répondre au premier objectif de la recherche, qui est d'identifier les écarts dans les résultats des élèves, les données recueillies sont : 1) les résultats au bulletin des élèves à chacune des étapes du cycle qui se traduisent par des cotes allant de A à D, 2) les résultats à l'examen ministériel qui se traduisent par des cotes allant de A à D et 3) des résultats sous forme de cotes (N, P ou R) se retrouvant au bilan de fin de cycle (voir Figure 1) pour témoigner de leur jugement global. La comparaison entre les différents résultats se fait en suivant la table de conversion suivante :

TABLEAU 2. *Cotes des données recueillies en cours et à la fin du cycle*

	Bulletin Exigences PEPALS	Examen ministériel Exigences du pro- gramme régulier	Bilan Exigences du pro- gramme régulier
Très réussi	1 traduit en A	A	R
Réussi	2 B	B	
Partiellement réussi	3 C	C	P
Non réussi	4 D	D	N

Afin d'examiner l'instrumentation (objectifs 2 et 3) utilisée par les enseignantes lors de leur collecte et de leur interprétation des résultats des élèves, une deuxième catégorie de données a été collectée au moyen d'une entrevue semi-structurée. Celle-ci porte sur les étapes de la démarche d'évaluation et

a été réalisée à l'aide de questions concernant des aspects précis de l'exercice du jugement à chacune de ces étapes. L'entrevue a porté sur leurs pratiques évaluatives en général et dans un deuxième temps sur la manière dont elles expliquent l'écart entre les épreuves internes (résultats au bulletin) et externes (résultats à l'examen du ministère) observé chez certains élèves. Les thèmes abordés lors de l'entrevue et les cas particuliers d'élèves ont été communiqués aux participantes une semaine avant l'entrevue pour qu'elles puissent se référer au dossier de traces (portfolio) concernant ces élèves. L'entrevue a également permis de questionner les enseignants sur les régulations pédagogiques (objectif 4) apportées à leurs pratiques évaluatives suite à la passation de l'épreuve externe.

Suite à l'entrevue et à une première analyse de ces données, des précisions ont été demandées par écrit sur différents contextes de jugement selon chaque cas d'élève. Celui-ci a été envoyé par courriel aux participantes deux semaines après l'entrevue. Elles ont eu une semaine pour le retourner aux chercheuses. Cette manière de procéder a permis d'enrichir les propos tout en assurant une certaine fidélité aux données.

Analyse des résultats

La première catégorie de données a été traitée à l'aide du logiciel SPSS de la manière suivante. Afin d'observer les différences entre, d'une part, les résultats obtenus en cours de cycle et en cours d'année (6^e) et, d'autre part, les résultats obtenus lors de l'épreuve externe ministérielle, nous avons compilé tous les résultats obtenus par les élèves. Pour chaque groupe, des tableaux synthèses sont produits (tableaux 3, 5 et 7) dans le but de distinguer les écarts et les similarités entre les résultats. Cette approche a permis d'isoler les résultats d'élèves montrant des écarts majeurs (de deux échelons entre, par exemple, A : réussi très bien et C : partiellement réussi, B : réussi et D : non réussi, etc.) entre les résultats aux bulletins et à l'épreuve externe. Ces cas ont été utilisés dans les entrevues avec les enseignantes afin de mieux documenter leur processus de jugement lorsqu'il y a tension entre les deux types d'évaluation.

L'analyse de la deuxième catégorie de données a été réalisée avec le logiciel Atlas-ti et a permis de cerner, à partir des discours recueillis lors des entrevues et des questions écrites, l'instrumentation utilisée pour la collecte et l'interprétation des données en cours de cycle et la manière dont les enseignantes ont porté leur jugement au bilan de fin de cycle et plus particulièrement dans le cas des élèves ciblés. L'analyse a aussi porté sur leur appréciation de l'instrumentation utilisée et sur les pistes de régulation que les enseignantes aimeraient adopter dans leur pratique future suite à la passation de ce premier examen et à l'illustration des résultats des élèves.

PRÉSENTATION DES RÉSULTATS

Les analyses illustrent le processus de prise de décision des enseignantes en général et lors de cas ciblé, plus litigieux.

Les sources

Sachant que le jugement repose sur une collecte de données suffisantes et pertinentes (MEQ, 2003), et qu'il s'exerce dès le choix de l'objet d'apprentissage (Lafortune et Allal, 2008), la première question posée était à cet effet : « Qu'est-ce que vous utilisez pour évaluer la compétence *To write texts* en anglais, langue seconde, en cours de cycle, et à la fin du cycle ? »

• *En cours de cycle*

Alice ne s'inspire pas des manuels scolaires; elle utilise et construit des situations d'apprentissage élaborées ayant le même format que l'épreuve externe ministérielle. Elle fait participer les élèves à leur évaluation et collige l'ensemble dans un portfolio annoté. L'enseignante dit différencier son enseignement en apportant des adaptations ou des modifications dans la réalisation des tâches afin que ces élèves puissent réaliser le défi proposé. Elle mentionne également qu'elle modélise les stratégies cognitives et métacognitives propres à l'apprentissage d'une langue seconde dans son enseignement.

Pour sa part, Bettie dit colliger, dans un portfolio, des informations venant tant du processus que du produit final. Elle utilise du matériel didactique et des situations d'apprentissage et d'évaluation proposées par la commission scolaire, en les adaptant au besoin. Elle privilégie les situations qui donnent du sens et qui permettent le transfert des apprentissages. Elle insiste également sur l'enseignement des stratégies d'apprentissage. Elle mentionne prendre des notes régulièrement et utiliser des grilles à l'occasion pour compléter ses observations.

• *À la fin du cycle.*

Notons qu'Alice et Bettie utilisaient pour la première fois l'épreuve externe fournie par le MELS. Pour sa part, Alice a aimé l'épreuve externe parce que ses élèves étaient bien préparés : ils avaient déjà travaillé la thématique proposée (les métiers), ils avaient vu le vocabulaire utilisé et ils étaient capables d'exprimer un souhait et un besoin (intention d'écriture familière). Du point de vue de l'enseignante, les élèves avaient donc les habiletés pour réussir l'épreuve. De plus, le défi présenté était réaliste pour eux car ils profitent d'un projet d'enrichissement du programme d'anglais langue seconde (PEPALS). Il y a donc correspondance entre ce qu'elle fait pendant l'année et le format de l'épreuve externe.

Quant à elle, Bettie a trouvé que l'épreuve externe du MELS était très longue et qu'elle permettait d'évaluer seulement deux des trois compétences au

programme. Par contre, ses élèves participent à un projet d'enrichissement du programme d'anglais langue seconde (PEPALS), ce qui a été propice au déroulement de l'épreuve. Elle considère, de ce fait, qu'ils étaient mieux préparés à cette épreuve, qui était somme toute assez difficile selon elle et qui aurait posé des problèmes à des élèves de groupe régulier.

L'interprétation des données

Selon Laveault (2008), l'interprétation de l'enseignant devrait s'appuyer sur un référentiel commun permettant de situer le niveau de compétence de l'élève. Notre deuxième question concernait cette pratique : « Quelle instrumentation utilisez-vous pour interpréter les données colligées en cours et à la fin du cycle ? »

• En cours de cycle

Pour interpréter les informations, Alice utilise des grilles qui décrivent des comportements observables lui permettant d'identifier facilement le niveau d'aisance de chaque élève, et ce pour chacune des compétences. Ces grilles étant mises à la disposition des élèves en début de travail, l'enseignante s'assure ainsi qu'ils utilisent des stratégies d'apprentissage enseignées et qu'ils connaissent bien les attentes face à une tâche spécifique. Globalement, elle considère qu'elle est bien outillée pour situer ses élèves et qu'elle a suffisamment d'information pour porter son jugement.

Bettie utilise également des grilles élaborées par la commission scolaire pour évaluer la participation et les efforts des élèves, ainsi que pour identifier les stratégies mobilisées par ceux-ci. L'utilisation de ces grilles lui simplifie la tâche. Elle utilise aussi les grilles proposées par les maisons d'édition, mais elle les trouve moins efficaces car elle doit identifier correctement chacune des compétences et transformer les résultats en cotes, ce qui ne représente pas, d'après elle, la performance observée de ses élèves dans d'autres contextes plus complexes. En effet, certains élèves peuvent très bien réussir un exercice extrait des manuels, mais sont incapables de résoudre une situation plus complexe.

• À la fin du cycle

Alice utilise une instrumentation provenant des situations d'apprentissage et d'évaluation de la commission scolaire et celle fournie par le MELS dans l'épreuve externe. Sa bonne compréhension des instruments lui facilite la tâche lorsqu'elle doit situer le niveau de compétence de ses élèves. Le document ministériel *Échelles de niveaux de compétence* et les attentes de fin de cycle (incluses dans le PFEQ) sont des références qui l'aident à porter son jugement : établir le portrait de chacun de ses élèves. Toutefois les attentes de fin de cycle sont parfois difficiles à utiliser parce qu'elles correspondent au profil d'un élève ayant un parcours régulier et non pas de celui qui participe à un projet

d'enrichissement (PEPALS). Les attentes de fin de cycle, échelon 5, vont en deçà des performances développées par ses élèves (effet plafond).

Bettie s'est servie des échelles descriptives et des grilles d'autoévaluation fournies dans l'épreuve externe du MELS pour interpréter les résultats de ses élèves. Elle a vérifié le carnet de l'élève pour avoir une bonne idée de ce qu'il avait appris. L'évaluation de la compétence en oral était plus difficile lorsqu'elle devait suivre les élèves qui se déplaçaient pour questionner leurs pairs. Elle aussi observe un effet plafond.

Le jugement des enseignantes

Pour cette troisième question, nous nous sommes intéressés à la façon dont chaque enseignante dit porter son jugement. L'enseignant peut adopter différentes postures (Jorro, 2000), en cumuler deux de façon séquentielle, selon qu'il se trouve dans une visée d'apprentissage ou d'évaluation certificative. Il peut aussi combiner deux approches (qualitative et quantitative) pour décider de la notation (Harlem, 2007). Nous leur avons alors demandé : « De quelle façon procédez-vous pour porter votre jugement au bilan des apprentissages? »

• Alice

Alice mentionne qu'elle porte son jugement sur la progression faite par l'élève à la fin du cycle. Elle tient davantage compte de la performance sans négliger les efforts de l'élève. Elle dit qu'elle ne pénalise pas un élève qui a été absent ou qui est dérangeant. Toutefois, elle ne tient pas compte seulement de l'épreuve ministérielle pour faire le bilan des apprentissages.

... s'il avait fallu que je calcule seulement le prototype d'épreuve peut-être que ça ne reflétait pas nécessairement ce que l'élève peut donner...

Pour réaliser le bilan de fin de cycle, elle dit qu'il est plus révélateur d'utiliser toutes les traces colligées dans les divers contextes à l'aide de diverses instrumentations. Il est alors important d'appuyer son jugement sur diverses sources d'information.

... je trouvais que c'était plus révélateur de me concentrer sur les traces que j'avais ramassées durant toute l'année parce qu'à la fin du cycle, c'est un jugement global qui est demandé...

Deux facteurs influencent la qualité du jugement porté par cette enseignante: (1) le fait des suivre les élèves pendant les deux années du cycle aurait permis de les connaître davantage et de réaliser un bilan mieux documenté que pendant les trois mois d'enseignement intensif où elle leur enseigne, d'autant plus qu'il y a un écart important dans les résultats de ses élèves entre les deux années qui se suivent; (2) bénéficier d'un projet d'enrichissement du programme d'anglais langue seconde (PEPALS), modèle intensif lui permet cependant de réaliser plus d'observations plus précises et de mieux documenter la réussite des élèves.

- *Bettie*

Bettie se sert du document ministériel *Échelles des niveaux de compétences*, (MEQ, 2002b) des attentes de fin de cycle provenant du programme (MEQ, 2002a), des observations de la progression pendant le cycle et de ce que l'élève a accompli à la dernière étape de l'année scolaire pour établir le bilan des apprentissages. Elle prétend bien connaître ses élèves, car elle leur enseigne pendant deux ans. Qu'elle se réfère ou non aux échelons, elle se dit capable de situer l'élève puisqu'elle sait ce qu'il est capable de faire. Son expérience (une vingtaine d'années) lui permet de facilement situer les élèves qui sont constants et ceux qui éprouvent beaucoup de difficultés, peu importe le temps d'enseignement. Cette enseignante se fait une «image dans sa tête» : elle prend une photo de l'élève, elle fait son portrait en l'observant et en intégrant différentes facettes à la fois.

... je sais ce qu'il est capable de faire, ce qu'il n'est pas capable de faire et ce n'est pas l'examen du MELS qui me donne ma réponse finale...

De plus, grâce au portfolio, elle regarde la progression de l'élève de la cinquième à la sixième année. Elle tient compte non seulement des traces des compétences, mais aussi de l'attitude de l'élève pour porter son jugement. Elle ajoute qu'au bout de deux ans, elle a créé un lien affectif avec les élèves plus faibles et que ceux-ci ont plus d'occasions de lui montrer ce qu'ils sont capables de faire. En contraste, dans le contexte de l'enseignement régulier (sans enrichissement), elle se sent incapable de porter son jugement sur le développement des compétences d'un élève ayant des troubles de comportement si elle ne le voit qu'une seule période par semaine.

L'analyse du discours permet de constater que les enseignantes se servent principalement de traces écrites (SAE, grilles d'observation, autoévaluation de l'élève) et que, parfois, d'autres facteurs entrent en jeu dans l'exercice du jugement professionnel tel l'effort et l'attitude. Ces points sont également soulevés dans les recherches de Lafortune et Allal (2008) ainsi que celles de Arter et Chappuis (2007). Qu'en est-il maintenant lorsqu'il y a écart entre le rendement observé en cours de cycle et le résultat à l'examen ministériel ? Nous avons demandé aux enseignantes si leur façon de porter leur jugement différait dans ces cas.

Présentation des résultats de chaque échantillon et illustration de quelques cas

L'analyse des résultats d'élèves présentant des écarts importants entre les épreuves internes et externes nous apporte des informations intéressantes quant au jugement final porté à la fin du cycle pour chacune des enseignantes.

- *Les élèves du premier échantillon de la classe d'anglais intensif d'Alice*

Rappelons qu'Alice ne donne qu'un cours intensif de trois mois en 6^e année, alors que le cycle se déroule sur 2 ans, la cinquième et la sixième année, mais

que c'est elle qui réalise le bilan de fin de cycle. Rappelons aussi que ce bilan n'est pas une sommation de résultats d'épreuves, mais situe l'élève par rapport au développement de sa compétence. Cela étant, on constate dans le tableau 3, que certains élèves ont obtenu un résultat moyen de C dans leur bulletin scolaire autant en 5^e (23%) qu'en 6^e année (59,3%), mais que seulement 11% ont eu la mention P (partiellement réussi) au bilan. Or, un élève obtenant un C dans son bulletin devrait être situé à P au bilan.

Alice l'explique par le fait que ses élèves font partie d'un groupe d'anglais enrichi et sont évalués selon des critères plus exigeants au bulletin alors que le profil (attente de fin de cycle, échelon 5) est formulé pour des élèves ayant un parcours régulier. Malgré l'obtention d'un C lors du cycle, la plupart se situe à l'échelon 5 qui représente l'attente de fin de cycle. Ce qui est illustré ici est la différence dans le niveau d'exigence entre les deux référentiels. En cours de cycle, l'enseignante va au-delà des attentes prescrits afin d'être conforme aux visées d'un programme d'enrichissement tandis qu'au bilan de fin de cycle, il n'y a pas d'échelles prévues dans le référentiel pour juger de la progression des groupes PEPALS. Le tableau 3 met en évidence les élèves qui obtiennent, par défaut, des cotes plus élevées au bilan.

TABLEAU 3. Comparatif des scores obtenus et de la décision prise au bilan par Alice

Cote	% des résultats obtenus par la moyenne du cycle ^a	% des résultats obtenus en 6 ^e année	% des résultats obtenus à l'épreuve	Résultats au bilan des apprentissages à la fin du cycle		
				Répond	Partiellement	Ne répond pas
A	38,5%	3,7%	18,2%	88,9%	11,1%	0%
B	38,5%	33,3%	45,5%			
C	23%	59,3%	22,7%			
D	0%	3,7%	13,6%			

^a5^e et 6^e années avec des enseignants différents.

D'autre part, il est à noter qu'un des trois élèves ayant obtenu un D lors de l'épreuve externe avait obtenu une moyenne de B durant le cycle (cas Andréanne). D'après l'enseignante, (voir tableau 4), cette élève aurait toujours fait bonne figure pendant le cycle, bien qu'elle ait obtenu un faible score lors de l'épreuve externe. Cette dernière justifie ainsi le fait qu'elle ne se fie pas uniquement à l'épreuve externe pour porter son jugement, car un problème survenu dans le milieu familial ou à l'école le jour de la passation de l'examen, s'est avéré néfaste pour l'élève. Cette façon de pratiquer son jugement, c'est-à-dire de relever la cote de l'épreuve externe ou tout simplement de ne pas en tenir compte, se confirme avec les autres cas.

TABLEAU 4. Illustration des cas d'élèves du groupe d'Alice ayant obtenu un D lors de l'épreuve ou Ne répondant pas au bilan des apprentissages

Liste d'élèves	Résultats du 3 ^e cycle pour la compétence <i>Write Texts</i>			Résultats à la tâche 7 de l'épreuve externe		Bilan des apprentissages
	5 ^e année	6 ^e année	Moyenne-3 ^e cycle	Task 7 : <i>writing my letter</i>	Overall <i>Performance Profiles</i>	N Ne répond pas P Répond partiellement R Répond
Andréanne	c-a-a	c	B	d-c-d	D	R
Alex	c-b-d	d	C	d-d-d	D	P
Annabelle	d-b-c	c	C	d-d-d	D	R

• *Les élèves du deuxième échantillon de la classe de Bettie*

Pour les élèves de ce groupe, et selon le tableau 5, les résultats obtenus lors de l'épreuve externe sont similaires aux résultats obtenus au bilan des apprentissages de fin de cycle. Ainsi, 66,7% des élèves ont obtenu un A ou un B en moyenne du cycle, comparativement à 64 % des élèves qui ont reçu la mention R (répond) au bilan. Les cotes A et B correspondant à la cote R (répond aux attentes) au bilan des apprentissages. On peut donc supposer que les exigences de Bettie en cours de cycle sont les mêmes que celles des documents ministériels.

TABLEAU 5. Comparatif des scores obtenus et de la décision prise au bilan par Bettie pour le premier groupe

Cote	% des résultats obtenus à la moyenne du cycle [*]	% des résultats obtenus en 6 ^e année	% des résultats obtenus à l'épreuve [#]	Résultats au bilan des apprentissages à la fin du cycle		
				Répond	Partiellement	Ne répond pas
A	12,5%	3,7%	8%	64%	32%	4%
B	54,2%	33,3%	48%			
C	33,3%	59,3%	36%			
D	0%	3,7%	8%			

^{*} 5^e et 6^e année avec le même enseignant.
[#] Deux élèves parmi les 27 n'ont pas fait l'épreuve ministérielle.

Il est intéressant ici de se pencher sur le cas de Bruno (voir tableau 6) dont les résultats auraient dû, logiquement, être semblables à ceux de Béatrice et Benoît. Cet élève a obtenu un D en 5^e année et a progressé vers un C en 6^e année dans son bulletin scolaire.

TABLEAU 6. Illustration des cas d'élèves du premier groupe de Bettie ayant obtenu un D lors de l'épreuve ou Ne répond pas au bilan des apprentissages

Liste d'élèves	Résultats du 3 ^e cycle pour la compétence <i>Write Texts</i>			Résultats à la tâche 7 de l'épreuve externe		Bilan des apprentissages
	5 ^e année	6 ^e année	Moyenne-3 ^e cycle	Task 7 : <i>writing my letter</i>	Overall Performance Profiles	N Ne répond pas P Répond partiellement R Répond
	Étapes 1-2	Étapes 3-4				
Bruno	c-c	d-d	C	b-b-c	B	N
Béatrice	c-c	d-c	C	d-d-d	D	P
Benoît	c-c	c-c	C	D	D	P

Il a obtenu un B lors de l'épreuve externe, ce qui devrait lui donner la mention P (répond partiellement aux attentes) au bilan des apprentissages. Par contre, il a obtenu la mention N (Ne répond pas aux attentes de fin de cycle). Bettie explique que cet élève a de grandes difficultés d'apprentissage. Il a bien progressé pendant le cycle (d'où l'attribution d'un C pour le cycle), mais en obtenant quand même de faibles résultats tout au long de l'année. Selon l'enseignante, cet élève ne possédait pas les acquis nécessaires pour réussir son 3^e cycle. Pour qu'il ne se sente pas exclu du groupe, elle lui a fait passer l'épreuve externe en même temps que les autres élèves de la classe, mais avec l'aide d'une accompagnatrice, sachant très bien qu'il ne pouvait la réussir et qu'elle n'allait pas en tenir compte dans son jugement final. L'enseignante a donc documenté son jugement à partir des caractéristiques personnelles de l'élève.

• *Les élèves du troisième échantillon de la seconde classe de Bettie*

Pour les élèves de ce dernier échantillon et selon le tableau 7, les résultats obtenus lors de l'épreuve externe et en 6^e année sont similaires aux résultats obtenus au bilan des apprentissages de fin de cycle. Ainsi, 72% des élèves ont obtenu un A ou un B en 6^e année et 64% à l'examen ministériel, comparativement à 68 % des élèves qui ont reçu la mention R (répond) au bilan. Les cotes A et B correspondant à la cote R (répond aux attentes) au bilan des apprentissages. On peut donc supposer, que les exigences de Bettie en sixième année sont sensiblement les mêmes que celles des documents ministériels.

TABLEAU 7. Comparatif des scores de l'élève et de la décision prise au bilan par Bettie pour le deuxième groupe

Cote	% des résultats obtenus à la moyenne du cycle *	% des résultats obtenus en 6 ^e année	% des résultats obtenus à l'épreuve	Résultats au bilan des apprentissages à la fin du cycle		
				Répond	Partiellement	Ne répond pas
A	4,2%	8%	28%	68%	32%	0%
B	75%	64%	36%			
C	20,8%	28%	28%			
D	0%	0%	8%			

* 5^e et 6^e année avec le même enseignant

Deux particularités sont observables dans ce groupe (voir tableau 7). Un des élèves (cas Charles-Antoine) a obtenu un B en 6^e année, un D lors de l'épreuve externe et un P (Répond partiellement aux attentes) au bilan des apprentissages. L'enseignante justifie son jugement par le fait que ce dernier progresse bien et possède plusieurs connaissances et capacités en écriture. Cependant, il a beaucoup de difficulté à les mettre en application dans une tâche complexe qui demande un certain niveau de maîtrise de la compétence, d'où son faible résultat à l'épreuve ministérielle (résultat : D). Ici, l'enseignante a tenu compte de la performance de l'élève à l'épreuve externe pour porter son jugement au bilan des apprentissages et ainsi rabaisser le résultat obtenu dans le bulletin scolaire.

Pour le cas de Caroline (voir tableau 8), l'enseignante explique l'obtention d'un D lors de l'épreuve externe et de la mention P (Répond partiellement aux attentes) au bilan par les difficultés de comportement de cette élève. L'enseignante a constaté, lors des deux années où elle a enseigné à Caroline, que cette dernière était capable de réussir les tâches demandées et qu'elle était motivée lorsqu'elles étaient réalisées individuellement. L'épreuve externe ministérielle comportait plusieurs tâches d'équipe, ce qui a nettement nuit à la performance de cette élève, étant donné son problème de socialisation. Bettie n'a donc pas tenu compte de l'épreuve externe pour porter son jugement en se fiant uniquement sur les nombreuses observations faites en classe.

CONSTATS ET LIMITES

Dans cette dernière partie, nous présenterons une synthèse des analyses en les commentant puis nous indiquerons les limites de cette étude.

Synthèse des analyses

Les différentes analyses nous permettent de constater que :

TABLEAU 8. Illustration des cas d'élèves du deuxième groupe de Bettie ayant obtenu un D lors de l'épreuve ou ayant Ne répondant pas au bilan des apprentissages

Liste d'élèves	Résultats du 3 ^e cycle pour la compétence <i>Write Texts</i>			Résultats à la tâche 7 de l'épreuve externe		Bilan des apprentissages
	5 ^e année	6 ^e année	Moyenne-3 ^e cycle	<i>Task 7 : writing my letter</i>	<i>Overall Performance Profiles</i>	N Ne répond pas P Répond partiellement R Répond
	Étapes 1-2	Étapes 3-4				
Charles-Antoine	c-b	b-c	B	d-d-c	D	P
Caroline	c-c	d-c	C	d-d-d	D	P

- 88,9% des élèves du premier échantillon du groupe d'Alice ont réussi leur 3^e cycle (Répond aux attentes du bilan des apprentissages), comparativement à 63,7% qui ont réussi l'épreuve (résultats obtenus A et B);
- les élèves du deuxième échantillon de la première classe de Bettie ont réussi leur 3^e cycle à 64% (Répond aux attentes du bilan des apprentissages) et ont réussi leur examen à 56% (résultats obtenus A et B);
- les élèves du troisième échantillon de la seconde classe de Bettie ont réussi leur 3^e cycle à 64% (Répond aux attentes du bilan des apprentissages) et ont réussi leur examen à 68% (résultats obtenus A et B).

L'écart entre le pourcentage de réussite au bilan de 3^e cycle et le pourcentage de réussite à l'épreuve externe est plus grand pour les élèves d'Alice (25%) que pour les élèves de Bettie (4% et 8%). Est-ce le fait que Bettie porte un jugement à partir d'un référentiel plus conforme aux normes ministérielles inscrites dans l'épreuve ? Est-ce le fait qu'Alice a peu tenu compte des résultats de l'épreuve externe pour porter son jugement au bilan ?

Quelques facteurs pourraient expliquer cet écart chez Alice:

- Les résultats obtenus indiquent moins de constance au bulletin scolaire en cinquième et en sixième année étant donné que des enseignants différents évaluaient les élèves ;
- Le fait que l'enseignante ne voit ses élèves que pendant une période de trois mois intensifs sur deux ans, comparativement à Bettie qui les a suivis une journée complète par semaine pendant deux ans ;
- La manière dont les documents ministériels ont été utilisés pour l'interprétation des données observées et consignées en se fiant uniquement aux attentes de fin de cycle, échelon 5 pour indiquer si l'élève répond (R) au bilan des apprentissages.

Toutefois, quel que soit leur groupe d'appartenance, il y a peu d'écart dans les résultats obtenus à l'épreuve externe. Est-ce dire que le fait de participer à un projet d'enrichissement du programme d'anglais langue seconde (PEPALS), dans un modèle intensif ou enrichi, rend ces élèves plus aptes que les élèves des groupes réguliers à réussir cette épreuve ? Les enseignants ont indiqué que leurs élèves étaient bien préparées, qu'ils avaient les ressources nécessaires pour bien réussir et c'est ce qui explique les bons résultats obtenus. L'effet plafond observé ne permet pas de juger efficacement la progression réelle des élèves participant à ce type de projet.

La régulation pédagogique

La passation de l'épreuve externe du MELS a amené Alice à réviser ses façons de faire pour la prochaine année scolaire en utilisant davantage d'autoévaluations et d'évaluations par les pairs, non seulement à la fin d'une tâche mais aussi en cours de réalisation. Elle considère que cette modalité, présente dans l'épreuve externe, aidera l'élève à mieux réussir.

Bettie prévoit mieux préparer ses élèves à ce type d'épreuve pour la prochaine année scolaire. Elle rendra ses élèves plus habiles dans la rédaction d'une lettre. Pour ce qui est de sa pratique quotidienne, elle constate qu'elle n'enseigne pas seulement des connaissances, mais elle les contextualise déjà dans des activités authentiques (aller au magasin pour faire des achats, par exemple) et propose des activités de transfert, ce qui correspond aux visées pédagogiques de l'épreuve ministérielle.

Cette première épreuve externe en anglais, langue seconde, aura donc suscité une certaine régulation des actions pédagogiques tout en confirmant les enseignantes dans l'utilisation qu'elles font des situations d'apprentissage et des grilles d'évaluation descriptives, « *rubrics* ». Elles désirent développer davantage ce type d'outils afin de porter un jugement plus documenté et mieux instrumenté.

Limites de la recherche

Dans une pratique de cumul de résultats et de moyennes, l'élève qui présente une inconstance dans ses résultats ne pose pas de problèmes à l'enseignant si celui-ci fait un calcul arithmétique selon une pondération définie et émet une note en pourcentage. Aujourd'hui, le bilan des apprentissages de fin de cycle exige que l'on porte un jugement sur le degré de maîtrise. Comme le démontre la recherche, plusieurs facteurs peuvent être considérés dans le contexte où le jugement est un processus continu. Des ajustements évaluatifs (Mottier Lopez et Allal, 2008) sont requis lorsque l'enseignant veut développer davantage son jugement afin de tenir compte des progrès réalisés par l'élève.

Les deux enseignantes nous ont permis d'illustrer le processus de jugement en cours et à la fin de cycle, toutefois, plusieurs questions restent à développer.

Est-ce que la contribution de collègues pourrait faciliter la prise de décision lorsqu'un écart se manifeste dans les résultats de certains étudiants ? Est-ce que cet écart serait moindre si le référentiel utilisé était bien documenté pour permettre une mesure plus conforme en cours et à la fin de cycle ? Un coefficient de contingence aurait probablement apporté plus de lumière sur le problème à savoir si cet écart est un écart dû à la grandeur de l'échantillon ou si l'une est plus près de la vérité que l'autre.

D'autres questions ont été soulevées en regard des biais évaluatifs. Même si notre recherche ne portait pas sur cet aspect, nous convenons que dans une visée où l'on ne se fie pas uniquement à un cumul de mesures réalisé de façon plus mécanique, le jugement professionnel de l'enseignant se doit d'être bien documenté. Ce n'est donc pas un problème de fidélité intra-juge ni inter-juge. Ce n'est pas un problème de fidélité du jugement, mais de validité !

De futures recherches portant sur la manière dont les enseignants s'y prennent pour porter leur jugement, spécifiquement lorsqu'il y a obligation de tenir compte de l'épreuve externe, présentent également des voies intéressantes d'investigation. Il serait aussi pertinent de vérifier si la façon de procéder est différente lorsqu'il n'y a pas d'épreuves externes imposées.

Cependant, s'il y a une indication qui sort clairement de ces deux études de cas, c'est que la pratique évaluative de ces enseignantes est « historique » au sens où leur jugement tient compte surtout de l'évolution de la progression des élèves et leurs relations pédagogiques à ceux-ci. Cette observation n'est pas neuve (Van der Maren, 1968), mais elle sied très bien aux visées actuelles de la Politique d'évaluation des apprentissages au Québec (MEQ, 2003).

CONCLUSION

Nous avons d'abord décrit la situation particulière au Québec, lorsque les enseignants remplissent un bilan des apprentissages à la fin de chaque cycle et où pour la première fois, en anglais, langue seconde, le Ministère a fourni, en 2007, une épreuve externe afin de donner un modèle aux enseignants du type des tâches dans lesquelles l'élève devrait être confronté pour développer les compétences inscrites dans le programme (PFÉQ). Nous avons par la suite relevé les recherches récentes portant sur l'importance du jugement professionnel dans la notation des élèves. Les analyses des études de cas ont permis de constater que les enseignantes interrogées utilisent des sources variées et qu'elles tiennent peu compte de l'épreuve externe pour porter leur jugement au bilan des apprentissages. Étant donné leur expérience, elles se fient davantage à la progression de l'élève et à l'état du développement de la compétence à la fin de l'année. Elles ne ressentent pas de malaise face à l'écart de certains résultats entre l'évaluation en cours d'année et l'épreuve externe. Dans une perspective d'évaluation tenant compte de l'histoire de leur relation pédagogique, elles

se sentent capables d'étayer leur jugement et de le documenter à l'aide des multiples observations recueillies dans différents contextes. Selon Cardinet (cité par Gerard, 2002), « évaluer, c'est avant tout donner du sens au résultat observé » (p. 2). C'est ce que les enseignantes de cette étude semblent bel et bien avoir cherché à faire en jugeant les élèves sur l'ensemble de ce qu'elles ont observé tout au long de leur enseignement plutôt que sur ce que le Ministère a voulu leur faire voir par les résultats de son épreuve. Comme au tribunal, un bon jugement se fonde sur la richesse et la variété des témoignages à la lumière de la jurisprudence.

NOTE

Les noms utilisés ont été modifiés pour respecter l'anonymat.

RÉFÉRENCES

Allal, L. et Mottier Lopez, L. (2007). *Régulation des apprentissages en situation scolaire et en formation*. Bruxelles: De Boeck.

Allal, L. et Mottier Lopez, L. (2009). Au Coeur du jugement professionnel en évaluation: des démarches de triangulation. *Les dossiers des sciences de l'éducation*, 22, 25-40.

Arter, J. A., et Chappuis, J. (2007). *Creating & recognizing quality rubrics*. Columbus, Ohio: UPER Saddle River

Durand, M.-J., et Chouinard, R. (2006). *L'évaluation des apprentissages, de la planification de la démarche à la communication des résultats*. Montréal: HMH.

Gerard, F.-M. (2002, Avril). L'indispensable subjectivité de l'évaluation, *Antipodes*, 156, 26-34

Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation* 33, 15-28

Jorro, A. (2000). *L'enseignant et l'évaluation: Des gestes évaluatifs en question*. Belgique: De Boeck.

Lafortune, L. et Allal, L. (2008). *Jugement professionnel en évaluation: pratiques d'enseignants primaires au Québec et à Genève*. Montréal: PUQ.

Lamoureux, A. (2006). *Recherche et méthodologie en sciences humaines*, 2^e éd. Montréal: Beauchemin, Chenelière Éducation.

Laurier, M. D., Tousignant, R., & Morissette, D. (2005). *Les principes de la mesure et de l'évaluation des apprentissages*, 3^e éd. Montréal: Gaëtan Morin.

Laveault, D. (2008). Le jugement professionnel: foyer de tensions et synergies nouvelles en évaluation scolaire. *Revue suisse des sciences de l'éducation*, 30(3), 483-500

Legendre, M.-F. (2001). Sens et portée de la notion de compétence dans le nouveau programme de formation. *Revue de l'AQEFLS*, 23(1), 12-30.

Legendre, R. (2005). *Dictionnaire actuel de l'éducation*, 3^e éd. Montréal: Guérin.

Lunz, M. E., Sthal, J. A., & Wright, B. D. (1994). Interjudge reliability and decision reproducibility. *Educational and Psychological Measurement*, 54 (4), 913-25

Ministère de l'Éducation du Québec (MEQ). (2002a). *Programme de formation de l'école québécoise, Éducation préscolaire – Enseignement primaire*. Québec: Gouvernement du Québec.

Ministère de l'Éducation du Québec (MEQ). (2002b). *Échelles des niveaux de compétence enseignement primaire*. Québec: Gouvernement du Québec.

- Ministère de l'Éducation du Québec (MEQ) (2003). Politique d'évaluation des apprentissages, Québec : Gouvernement du Québec
- Ministère de l'éducation, du loisir et des sports (MELS). (2007). Prototypage d'épreuve English as a second language, Fin du 3e cycle du primaire. 017660, 2007 <http://www.mels.gouv.qc.ca/DGFJ/de>
- Mottier Lopez, L., et Allal, L., (2008). Le jugement professionnel en évaluation: un acte cognitif et une pratique sociale située. *Revue suisse des sciences de l'éducation*, 30(3), 465-482
- Scallon, G. (2004). *L'évaluation des apprentissages dans une approche par compétences*. St-Laurent: Éditions du Renouveau Pédagogique.
- Van der Maren, J.-M. (1968) Analyse de la relation pédagogique en terme de négociation. *Revue de psychologie et des sciences de l'éducation*, 3(2), 183-198.
- Van der Maren J.-M., (2003) *La recherche appliquée en pédagogie*, 2^e éd. Belgique: De Boeck,
- Wiggins, G. (1998). *Educative assessment: Designing assessments to Inform and improve student performance*. San Fransisco: Jossey-Bass.

MICHELINE-JOANNE DURAND est professeure agrégée au Département d'administration et fondements de l'éducation, Université de Montréal. Spécialiste en curriculum et en évaluation des apprentissages, elle dirige des recherches sur l'évaluation en salle de classe et sur l'instrumentation dans le contexte de tâches complexes. La présente recherche a été rendue possible grâce à une subvention institutionnelle du CRSH.

MICHELINE-JOANNE DURAND is an associate professor in the Department of Administration and Educational Foundations at the University of Montreal. Specialized in curriculum and in evaluation of learning, she leads studies in the field of evaluation in the classroom context and the instrumentation in the context of complex tasks. The present study has been possible with institutional funding from the Social Sciences and Humanities Research Council of Canada.