

Les conceptions des futurs enseignants du secondaire en matière d'évaluation: entre un accord de principe et une vision limitée de l'évaluation formative

Annick Fagnant

Christelle Goffin

UNIVERSITÉ DE LIÈGE

MOTS CLÉS: évaluation formative, évaluation certificative, évaluation normative, conceptions des futurs enseignants, vécu en tant qu'apprenant

S'appuyant sur un questionnaire soumis à l'entrée en formation initiale, l'étude cherche à cerner la façon dont les futurs enseignants du secondaire perçoivent les liens entre évaluation et apprentissage. Elle s'intéresse aux différentes fonctions de l'évaluation, aux différentes formes d'évaluation formative que les futurs enseignants ont rencontrées dans leur parcours scolaire et à la façon dont ils peuvent imaginer des modalités d'évaluation variées. Alors que les résultats des questionnaires fermés montrent que les participants soutiennent une vision élargie de l'évaluation formative, les questions ouvertes témoignent quant à elles d'une vision étroite dans laquelle le versant régulation n'est pas suffisamment présent et où les élèves semblent peu impliqués. Les expériences vécues par les futurs enseignants au cours de leur scolarité semblent limitées et ils imaginent difficilement qu'il soit possible d'intégrer pleinement l'évaluation aux situations d'apprentissage.

KEY WORDS: formative assessment, assessment for certification, normative assessment, conceptions of future teachers, experience as a learner

Based on a questionnaire submitted at the very beginning of initial training, the study aims to identify how future secondary school teachers perceive the links between assessment and learning. The study also looks at different functions of assessment, at different kinds of formative assessment that future teachers encounter during their schooling and at the way they can image different evaluation procedures. When the results coming from Likert-type questionnaires show that participants sustain an enlarged vision of formative assessment, open questions show indeed a narrower view of assessment in which regulation is poorly

represented and where students are few involved. Their experiences as students seem limited and they encounter difficulties to imagine how to fully integrate assessment in learning situations.

PALAVRAS-CHAVE: avaliação formativa, avaliação certificativa, avaliação normativa, concepções dos futuros professores, experiência como aprendiz

Com base num questionário submetido à entrada da formação inicial, o estudo pretende identificar o modo como os futuros professores do secundário concebem as ligações entre a avaliação e a aprendizagem. O estudo aborda as várias funções da avaliação, as diferentes formas de avaliação formativa com as quais os futuros professores se confrontam nos seus percursos escolares e o modo como podem imaginar modalidades de avaliação variadas. Embora os resultados dos questionários fechados mostrem que os participantes sustentam uma visão abrangente da avaliação formativa, as perguntas abertas refletem, em contrapartida, uma visão estreita na qual o lado de regulamentação não está suficientemente presente e onde os alunos parecem pouco envolvidos. As experiências vividas pelos futuros professores durante a sua escolaridade parecem limitadas, encontrando dificuldades para imaginar a possibilidade de integrar plenamente a avaliação nas situações de aprendizagem.

Introduction

Les recherches qui se sont intéressées aux conceptions des enseignants ou futurs enseignants en matière d'évaluation montrent qu'ils reconnaissent volontiers l'intérêt de l'évaluation formative pour soutenir l'apprentissage des élèves, tout en indiquant que les retombées dans les classes sont pourtant limitées (Issaieva, Yerly, Petkova, Marbaise & Crahay, 2015; Luisoni & Monnard, 2015; Soussi, Ducrey, Ferrez, Nidegger & Viry, 2006).

Les modifications de pratiques seraient en effet tributaires de résistances diverses, liées notamment à une vision trop étroite de l'évaluation formative (Goffin, Renson & Fagnant, 2014) conduisant les futurs enseignants à envisager une instrumentation lourde en vue de réaliser un test «à blanc», sans imaginer d'autres formes plus interactives et informelles s'intégrant davantage dans les activités d'enseignement/apprentissage elles-mêmes (Allal & Mottier Lopez, 2005; Morissette, 2010). Les résistances pourraient aussi s'expliquer par une non-prise en compte des conceptions à l'entrée en formation et par la difficulté de réellement les infléchir (Boraita & Crahay, 2013; Crahay, Wanlin, Issaieva & Laduron, 2010). Elles pourraient aussi être tributaires des tensions entre les rôles historiques de l'évaluation, axés sur le classement et la sélection, alors que les tendances pédagogiques actuelles invitent à concevoir l'évaluation comme une aide à l'apprentissage (Issaieva et al., 2015; Luisoni & Monnard, 2015; Mottier Lopez, 2015).

L'écart entre la vision positive de l'évaluation formative telle qu'elle apparaît dans la plupart des enquêtes par questionnaire et les pratiques de classe qui témoignent d'une faible mise en œuvre de celle-ci (Soussi et al., 2006; Younès & Gaime, 2012) pourrait aussi s'expliquer par un phénomène de «désirabilité sociale» résultant «de la volonté du répondant de se montrer sous un jour favorable» et par un «artefact de questionnement» les conduisant à répondre d'une certaine façon s'ils pensent avoir discerné l'objectif du questionnaire (Butori & Parguel, 2010, p. 3). Ainsi, conscients que ce type d'évaluation est aujourd'hui prôné dans le système

scolaire, les enseignants marqueraient un accord de principe avec les propositions relatives à l'importance de pratiquer différentes modalités d'évaluation formative pour soutenir les apprentissages des élèves, et ce, quelles que soient en réalité leurs conceptions profondes sur ces questions.

S'appuyant sur un questionnaire soumis à l'entrée en formation initiale, la présente étude s'intéresse aux conceptions des futurs enseignants du secondaire relativement au rapport que l'évaluation entretient avec l'apprentissage. Partant du principe selon lequel ces conceptions sont notamment influencées par leur vécu en tant qu'élève (ce que Vause [2010] appelle la socialisation secondaire), l'étude s'intéresse également aux différentes formes d'évaluation formative que les futurs enseignants ont rencontrées dans leur parcours scolaire et à la façon dont ils peuvent imaginer des modalités d'évaluation variées. Dépassant les accords de principe fréquemment observés dans les recherches s'appuyant sur des questionnaires fermés, l'analyse, réalisée au départ de questions ouvertes, cherche à approcher différemment les conceptions partagées par les enseignants en début de formation.

Cadre théorique

De nombreuses recherches se sont intéressées aux conceptions des enseignants ou futurs enseignants en matière d'évaluation scolaire (Boraita & Issaieva, 2013; Issaieva & Crahay, 2010; Issaieva, Pini & Crahay, 2011; Issaieva et al., 2015; Goffin et al., 2014; Luisoni & Monnard, 2015; Soussi et al., 2006). La plupart de ces recherches s'appuient sur des questionnaires fermés constitués d'une série d'items face auxquels les répondants doivent se positionner. Elles montrent des structurations différentes de ce champ de conceptions selon les contextes sociétaux et les publics interrogés (enseignants en fonction ou étudiants en formation, notamment). Elles permettent également de dégager les éléments jugés les plus importants par les personnes interrogées.

Issaieva et Crahay (2010) ont construit leurs premières échelles en s'appuyant sur la théorie des buts d'accomplissement, qui distingue deux types de buts principaux poursuivis par les élèves par rapport aux tâches scolaires: des buts de maîtrise ou d'apprentissage plaçant le développement de compétences au cœur des préoccupations, d'une part, et des buts de performance centrés sur la volonté de démontrer sa compétence ou d'évi-

ter de montrer son incompetence, d'autre part (Sarrazin, Tessier & Trouiloud, 2006). Issaieva et Crahay précisent que les travaux menés par Ames au début des années 1990 dans ce domaine ont permis de distinguer des pratiques pédagogiques et évaluatives distinctes, susceptibles de renforcer le développement de l'un ou l'autre type de buts. Ainsi, certaines pratiques évaluatives sont qualifiées de «formatives» dans la mesure où elles visent la régulation des apprentissages en misant notamment sur «l'encouragement des élèves en cas de difficulté, et l'ajustement de l'enseignement en vue de la maîtrise des connaissances» (p. 36). D'autres sont qualifiées de «normatives» dans la mesure où elles visent «la hiérarchisation des performances et le classement des élèves» (p. 36) en pénalisant l'erreur, en favorisant la compétition et en évaluant les capacités des élèves, plutôt que l'état de leurs connaissances.

L'étude réalisée par Issaieva et Crahay (2010) auprès d'une cinquantaine d'enseignants du primaire en France permet de dégager, à l'aide d'une analyse factorielle, une structuration des conceptions qui rend compte de la dimension formative (en distinguant un versant diagnostic et un versant régulation) et de la dimension normative (en distinguant trois sous-dimensions : repérage des capacités, bilan des acquis et sanction des élèves qui n'ont pas travaillé). S'appuyant sur ces mêmes échelles, la recherche d'Issaieva et ses collaborateurs (2011) révèle que plus de 80% des enseignants ont une attitude extrêmement positive envers l'évaluation formative, alors que plus de la moitié d'entre eux rejettent nettement sa fonction normative.

Les choix posés par les auteurs d'orienter leurs échelles au départ des travaux menés sur les buts d'accomplissement les ont conduits à proposer une opposition peu habituelle entre «formatif» et «normatif», alors qu'il est plus courant d'opposer les fonctions «formative» et «certificative» de l'évaluation¹. En s'appuyant sur cette opposition, ils faisaient aussi l'hypothèse selon laquelle la plupart des enseignants assimileraient la fonction certificative de l'évaluation «à un traitement normatif des résultats de cette évaluation» (Issaieva & Crahay, 2010, p. 37), ce qui les avait conduits à négliger le développement d'une échelle visant à cerner une fonction certificative de l'évaluation, appuyée cette fois sur une référence critériée.

Dans l'étude publiée en 2015, Issaieva et ses collaborateurs proposent dès lors un questionnaire comprenant trois dimensions : la première dimension «envisage l'évaluation comme un processus formatif visant la régulation

des apprentissages et la remédiation des difficultés scolaires» (p. 81); la deuxième se situe dans une optique certificative à référence normative visant le classement des élèves et la sanction; la troisième, située dans une approche critériée de l'évaluation certificative, conçoit l'évaluation comme «un processus visant à appréhender si les élèves ont atteint les objectifs poursuivis et ceci indépendamment de tout souci de classement des élèves» (p. 81)².

À partir d'un échantillon d'enseignants du primaire en Suisse, en Belgique francophone, en Bulgarie et en Turquie, les résultats révèlent une structuration des conceptions dépendante du contexte socioculturel dans lequel les enseignants exercent. Dans tous les cas, à l'encontre du questionnaire construit pour mettre en évidence trois dimensions, ce sont des structures composées de deux facteurs qui ont été mises à jour. L'analyse des informations recueillies en Suisse fait émerger un premier facteur, qui regroupe les versants «formatif» et «certificatif critérié», et un deuxième facteur, qui reprend le versant «normatif». Les deux facteurs corrèlent positivement, ce qui, selon les auteurs, «confirme l'idée que les enseignants opèrent certaines distinctions conceptuelles, tout en souscrivant aux deux conceptions évaluatives, théoriquement opposées» (p. 84). Concernant les données belges, le premier facteur correspond à l'évaluation formative, tandis que le second renvoie à une perception de l'évaluation à visée normative. Les items relatifs au versant critérié de l'évaluation certificative n'entrent pas dans la structure factorielle. Dans le cas présent, les deux facteurs corrèlent négativement, témoignant de l'idée selon laquelle les tenants de l'évaluation formative rejettent la visée normative de l'évaluation.

Complémentairement, une étude menée en Suisse par Boraita et Issaieva (2013) compare les conceptions des futurs enseignants à celles d'enseignants en exercice dans l'enseignement primaire. Alors que les fonctions certificative critériée et formative s'entremêlent auprès des enseignants en fonction (comme dans l'étude d'Issaieva et ses collaborateurs, 2015), la structuration des conceptions des futurs enseignants met en exergue les trois fonctions initialement prévues dans le questionnaire. Les auteurs émettent l'hypothèse selon laquelle les futurs enseignants construiront leurs représentations à partir «de connaissances théoriques» et «d'idéaux» les conduisant à identifier clairement les différentes fonctions de l'évaluation. Toutefois, ces idéaux sont souvent bousculés lorsque les enseignants rencontrent les réalités de terrain.

En effet, cette question a été abordée par Luisoni et Monnard (2015), qui se sont intéressés aux conceptions des futurs enseignants du primaire et du début du secondaire lors de leur entrée en formation, en cours de formation et lors de leur entrée dans le métier. Ces auteurs, s'appuyant également sur des questionnaires fermés, structurent les conceptions en matière d'évaluation selon une dimension «critériée» et une dimension «normative». D'abord, la dimension critériée est elle-même constituée d'une sous-dimension formative (avec une visée diagnostique et de régulation) et de deux sous-dimensions certificatives à référence critériée («évaluation des objectifs» et «réussite des objectifs»). La dimension normative, quant à elle, recouvre notamment les notions de classement des élèves (idée selon laquelle la position dans le classement serait un indicateur du niveau de l'élève) et de mesure de capacités ou d'intelligence (idée selon laquelle les difficultés seraient liées au potentiel de l'élève). Globalement, les résultats montrent une tendance positive marquée, dès l'entrée en formation, pour «une évaluation centrée sur une définition claire et la réussite d'objectifs pédagogiques ainsi que sur une évaluation au service de l'apprentissage de l'élève» (p. 190-191). Les conceptions relatives à cette dimension évoluent peu pendant la formation puisqu'elles sont déjà très positives à l'entrée. Toutefois, les résultats montrent une baisse prononcée entre la troisième et la cinquième année d'expérience professionnelle. Si le versant «normatif» est déjà peu plébiscité dès l'entrée en formation, c'est aussi pour ce versant que l'effet de la formation est le plus important, conduisant à un rejet de plus en plus prononcé tout au long du parcours des études. Par contre, dès l'entrée dans le métier, l'accord avec cette conception de l'évaluation ne cesse d'augmenter, réduisant drastiquement l'effet de la formation sur cette vision de l'évaluation.

Dans toutes ces études, les résultats montrent une prégnance de la fonction formative et un rejet assez net de la fonction normative. Rappelons toutefois que les questionnaires fermés sont aisément soumis à un phénomène de désirabilité sociale et à un «artefact de questionnement» (Butori & Parguel, 2010) conduisant les enseignants à marquer un accord de principe avec les propositions soutenant l'importance d'une évaluation au service de l'apprentissage, et ce, même si la pratique de ce type d'évaluation reste malgré tout marginale, en particulier au secondaire (Younès & Gaime, 2012). Notons aussi que les différents questionnaires susmentionnés appréhendent une vision assez classique de l'évaluation formative et qu'ils présentent en ce sens une limite importante. En effet,

l'évaluation formative est essentiellement envisagée dans la perspective de Bloom, avec un versant « diagnostic » permettant de donner des repères sur les acquis et sur les lacunes des élèves, et un versant « remédiation » ou « régulation » de l'enseignement dans le but d'aider les élèves (Issaieva & Crahay, 2010). Cette vision « classique » se retrouve également dans l'enquête de Soussi et ses collaborateurs (2006), où les résultats des entretiens rapportent que la moitié des enseignants interrogés (surtout du secondaire) ont tendance à réduire la fonction formative à l'importance de donner des commentaires formatifs à la suite d'évaluations à visée certificative.

En fin de compte, si l'on s'accorde à reconnaître que les questionnaires envisagent une vision relativement classique de l'évaluation formative *au service de* l'apprentissage, qu'en est-il d'une perspective « élargie » conduisant à dépasser la vision d'une évaluation classique, se présentant sous la forme d'un test papier-crayon ayant lieu en fin de séquence, pour envisager l'intégration de l'évaluation au cœur même des activités d'enseignement/apprentissage (Allal & Mottier Lopez, 2005)? Qu'en est-il d'une perception de l'évaluation *en tant qu'*apprentissage visant à « étayer l'apprentissage » et vue comme un « moyen efficace de gérer l'utilisation et le développement de stratégies nécessaires à l'atteinte d'objectifs » (Rey & Feyfant, 2014, p. 22)? Qu'en est-il de l'implication de l'élève dans le processus d'évaluation, par l'intermédiaire des formes d'évaluation mutuelle, de coévaluation ou d'autoévaluation (Allal, 2007; Campanale, 2007; Mottier Lopez & Laveault, 2008; Mottier Lopez, 2015)?

Dispositif et questions de recherche

Public concerné

Le public concerné est constitué des futurs enseignants inscrits à l'agrégation de l'enseignement secondaire supérieur ou en master à finalité didactique dans une université belge francophone. Ils se destinent à enseigner à des élèves de 15 à 18 ans. Un questionnaire a été soumis à 220 étudiants à l'entrée de leur formation pédagogique dans le cadre d'un cours qui rassemble des étudiants issus de 18 disciplines différentes (mathématiques, histoire, langues modernes, éducation physique, etc.). La grande majorité des étudiants ne dispose d'aucune expérience d'enseignement et seuls ceux-ci (N = 187) sont retenus pour les analyses qui suivent³.

Questionnaire utilisé pour cerner les conceptions des futurs enseignants

Un questionnaire comprenant cinq volets a été proposé. Le premier volet vise à approcher les conceptions les plus spontanées en demandant aux futurs enseignants d'explicitier les différents liens qu'ils perçoivent entre «évaluation» et «apprentissage». La question posée était la suivante: *Dans l'enseignement, «évaluation» et «apprentissage» sont liés de différentes façons. Au moyen de trois phrases courtes, exprimez les différents liens que vous percevez entre ces deux termes, ceux-ci devant apparaître dans chaque phrase.*

Le deuxième volet propose une étude de cas invitant les futurs enseignants à se projeter dans une situation de stage durant laquelle ils ne sont pas autorisés à réaliser d'évaluation traditionnelle (une interrogation écrite, p. ex.). On leur demande alors d'imaginer ce qu'ils pourraient mettre en place pour évaluer néanmoins les apprentissages de leurs élèves. On vise ici à percevoir leur capacité à envisager des formes variées d'évaluation pouvant s'intégrer dans une situation d'enseignement/apprentissage.

Le troisième volet est constitué d'un questionnaire fermé comprenant 14 items face auxquels les futurs enseignants doivent se positionner sur une échelle de type Likert à 6 degrés allant de «pas du tout d'accord» à «tout à fait d'accord». Il reprend les items du questionnaire validé par Issaieva et Crahay (2010) pour les fonctions formative (5 items) et normative (5 items) de l'évaluation, auxquels ont été ajoutés les items de Boraita et Issaieva (2013) permettant d'appréhender une perspective certificative à référence critériée (4 items). Ainsi conçu, le questionnaire correspond à la même logique de construction que celui utilisé par Issaieva et ses collaborateurs (2015) dans leur étude comparant plusieurs contextes européens.

Le quatrième volet propose un questionnaire conçu par nos soins et inspiré des trois fonctions de l'évaluation telles qu'elles sont définies par le gouvernement canadien (voir travaux du Protocole de l'Ouest et du Nord canadiens de collaboration concernant l'éducation, 2006). Ce deuxième questionnaire fermé comprend 18 items visant à identifier les dimensions «évaluation de l'apprentissage» (fonction certificative: 7 items), «évaluation au service de l'apprentissage» (fonction formative classique, avec un versant diagnostique et un versant régulation: 5 items) et «évaluation en tant qu'apprentissage» (fonction formative élargie, invitant l'apprenant

à s'autoévaluer et à analyser les stratégies d'apprentissage qu'il met en œuvre : 6 items). Il s'agit cette fois de se positionner sur une échelle de type Likert à 6 degrés allant de « pas du tout important » à « extrêmement important ».

Enfin, dans le cinquième et dernier volet du questionnaire, différentes formes d'évaluation sont proposées : évaluation formative, évaluation mutuelle, coévaluation et autoévaluation. On demande aux futurs enseignants de chercher des souvenirs d'expérience de ce type dans leur vécu en tant qu'élève et de les décrire.

Questions de recherche

L'étude cherche à répondre à trois questions principales visant à appréhender les conceptions, le vécu et les projections de ces futurs enseignants.

- 1) Comment les futurs enseignants du secondaire perçoivent-ils les liens entre évaluation et apprentissage ?
 - a. *Sur la base du premier questionnaire fermé*, quelle est la structuration des conceptions qu'on retrouve avec un public différent sous deux aspects : futurs enseignants du secondaire (alors que les études appuyées sur ce questionnaire portent sur l'enseignement primaire) à l'entrée en formation (alors que les études portent sur des enseignants en fonction [Issaieva & Crahay, 2010 ; Issaieva et al., 2011 ; Issaieva et al., 2015] ou sur des étudiants en fin de formation [Boraita & Issaieva, 2013]) ?
 - b. *Sur la base du deuxième questionnaire fermé*, est-il possible de mettre à jour une structuration des conceptions qui distingue une évaluation de l'apprentissage, au service de l'apprentissage et en tant qu'apprentissage ?
 - c. *Sur la base de ces deux questionnaires*, retrouve-t-on une vision positive de l'évaluation formative et de l'évaluation certificative à référence critériée ainsi qu'un rejet de l'évaluation normative, comme dans les études susmentionnées ?
 - d. Lorsque les futurs enseignants évoquent spontanément des propositions visant à relier les concepts d'évaluation et d'apprentissage, à quelle(s) fonction(s) de l'évaluation font-ils référence ?
- 2) Leur vécu en tant qu'apprenant a-t-il permis aux futurs enseignants du secondaire de rencontrer différentes formes d'évaluation ?

- 3) Les futurs enseignants peuvent-ils se projeter en tant qu'enseignant devant évaluer ses élèves et parviennent-ils à intégrer cette évaluation au cœur d'une activité d'enseignement/apprentissage?

Types d'analyses réalisées

Les questions fermées ont été analysées à l'aide du logiciel Statistica (version 12). La consistance interne des dimensions présumées (alpha de Cronbach) a tout d'abord été vérifiée. Une analyse en composantes principales a ensuite été réalisée pour observer si une autre structuration des données apparaissait. Finalement, les dimensions qui ressortaient de cette analyse et qui maximisaient la consistance interne ont été retenues. Pour analyser les résultats des questions ouvertes, une grille d'analyse a été construite a priori et une analyse inductive (inspirée de l'approche proposée par Blais et Martineau, 2006) a été réalisée pour affiner la grille et pour dégager des sous-catégories significatives. Une cinquantaine de copies a fait l'objet d'un double codage. Les désaccords ont donné lieu à un débat qui a permis d'encore préciser la grille et les critères de classement. Pour les copies suivantes, quelques sondages ont permis de s'assurer d'un accord intercodeurs pour chacune des questions; les cas jugés problématiques (au sens de difficiles à coder selon les critères établis) ont été discutés.

Résultats

Conceptions relatives aux fonctions de l'évaluation telles qu'induites par les questionnaires fermés

La première analyse porte sur le questionnaire tiré des études d'Issaieva et ses collaborateurs (2010, 2013 et 2015). L'analyse en composantes principales fait ressortir une structure à deux facteurs dont la consistance interne s'est révélée satisfaisante et qui explique 37,7% de la variance totale (respectivement 20,7% et 17,0% pour chaque facteur). Le premier facteur combine les fonctions formative (*L'analyse des résultats des élèves aux évaluations permet d'ajuster l'enseignement en fonction de leurs faiblesses*) et certificative critériée (*Les évaluations permettent de voir si les objectifs du programme sont atteints*). Le second facteur traduit la dimension normative de l'évaluation (*Les évaluations permettent de repérer les bons et les mauvais élèves ou encore de sanctionner les élèves qui n'ont pas travaillé*).

La seconde analyse porte sur le questionnaire construit par nos soins. L'analyse en composantes principales fait également ressortir une structure à deux facteurs qui explique 40,5% de la variance totale (respectivement 23,1% et 17,4%). Le premier facteur combine une vision de «l'évaluation *au service de* l'apprentissage» (*Les évaluations visent à déterminer ce que l'enseignant peut faire pour aider les élèves à progresser*) et «*en tant qu'*apprentissage» (*Les évaluations visent à inciter les élèves à ajuster leurs stratégies d'apprentissage*). Le second facteur correspond à la dimension «évaluation *de* l'apprentissage» (*Les évaluations visent à établir un bilan de ce que les élèves ont appris*). Le premier facteur traduit une vision élargie de l'évaluation formative et le second une vision certificative, sans référence normative.

Le tableau 1 synthétise les dimensions retenues et présente le nombre d'items, les alphas de Cronbach et les scores moyens obtenus par les futurs enseignants.

Tableau 1
Scores moyens pour les différentes dimensions des deux échelles

Dimension	N ^{bre} d'items	Alpha de Cronbach	Score moyen (N = 187)
Évaluation à visée formative et certificative critériée	9	0,72	4,69
Évaluation à visée normative	5	0,66	2,71
Évaluation <i>au service de</i> l'apprentissage et <i>en tant qu'</i> apprentissage	10	0,81	4,51
Évaluation <i>de</i> l'apprentissage	7	0,77	4,11

Les résultats montrent clairement une adhésion aux fonctions formative et certificative de l'évaluation ainsi qu'un net rejet de la fonction normative.

Conceptions relatives aux fonctions de l'évaluation spontanément formulées par les futurs enseignants

En orientant les futurs enseignants sur les liens entre évaluation et apprentissage, nous avons comme intention de voir dans quelle mesure ils distinguaient les fonctions formative et certificative de l'évaluation. Les répondants devaient produire trois propositions. L'objectif était d'essayer de les pousser à distinguer certains éléments de l'évaluation formative (p. ex. «*au service de*» ou «*en tant qu'apprentissage*») ou à distinguer le versant «*diagnostic*» du versant «*régulation*», voire à scinder les régulations situées du côté de l'élève ou de l'enseignant (p. ex., adapter sa manière d'apprendre ou modifier ses pratiques pédagogiques). Le tableau 2 est organisé en suivant cette logique : il distingue l'évaluation *de* l'apprentissage, l'évaluation *au service de* l'apprentissage et l'évaluation *en tant*

Tableau 2
Répartition des propositions spontanément émises par les futurs enseignants qui s'inscrivent dans une logique d'évaluation de, au service de ou en tant qu'apprentissage

Évaluation de l'apprentissage	N^{bre} de propositions	
Pas de destinataire précisé	141	
L'élève est le destinataire	6	
L'enseignant est le destinataire	8	
L'élève et l'enseignant sont les destinataires	1	
Les parents sont les destinataires	1	
<i>Total</i>	<i>157 (28 %)</i>	
Évaluation au service de l'apprentissage	N^{bre} de propositions	
	Informier	Réguler
Pas de destinataire précisé	30	6
L'élève est le destinataire	18	3
L'enseignant est le destinataire	22	21
L'élève et l'enseignant sont les destinataires	1	–
<i>Sous-total</i>	<i>71 (13 %)</i>	<i>30 (5 %)</i>
<i>Total</i>	<i>101 (18 %)</i>	
Évaluation en tant qu'apprentissage	N^{bre} de propositions	
L'élève est de facto le destinataire	22	
<i>Total</i>	<i>22 (4 %)</i>	

qu'apprentissage. Les lignes précisent le destinataire (enseignant et/ou élève, voire parents) et le caractère informatif ou régulateur de l'évaluation *au service de* l'apprentissage. Seules 50% des propositions (280 sur 561 ; 3×187 participants) entrent dans ces catégories, les autres n'étant pas orientées sur ces fonctions.

La catégorie la plus représentée et recouvrant presque un tiers des propositions est celle qui traduit l'idée d'une évaluation *de* l'apprentissage. Cette catégorie se caractérise par une prégnance des verbes *vérifier* et/ou *contrôler*. Elle consiste majoritairement à faire référence à l'idée de « bilan des acquis », mais contient aussi quelques nuances comme l'idée de *s'assurer* du bon apprentissage, de *vérifier* qu'il a *porté ses fruits*, de *vérifier* le *degré* de maîtrise, voire la *compréhension*. Le destinataire n'est généralement pas précisé (*L'évaluation permet de tester l'apprentissage [...], de vérifier s'il a bien été assimilé*), même si l'élève est parfois mentionné dans la phrase (*L'évaluation permet de vérifier les apprentissages que l'élève a faits pendant x heures de cours*). Dans 14 propositions, l'élève ou l'enseignant est clairement identifié comme le destinataire de ce bilan (*L'évaluation permet à l'élève de mesurer ce qu'il a appris; L'évaluation permet au professeur de voir que l'apprentissage d'une matière est effectué*). Une seule proposition précise l'enseignant et l'élève comme destinataires (*Une évaluation est utile aussi bien pour l'élève que pour le professeur*) et une fait référence aux parents (*Les évaluations permettent aux parents d'avoir une idée de l'apprentissage fait par leurs enfants*).

Un peu moins de 20% des propositions ont été comptabilisées dans la catégorie « Évaluation *au service de* l'apprentissage ». Relevant majoritairement du versant « informatif » (71 sur 101), elles se distinguent de la catégorie précédente par la présence des verbes *repérer*, *révéler*, *se situer dans son apprentissage* et par des qualificatifs ayant trait à des intentions diagnostiques (*repérer* les lacunes, faiblesses ou difficultés, *cibler* les points délicats) ou faisant référence à une notion de qualité de l'apprentissage réalisé (l'évaluation est un outil *révélateur* de la *qualité*/de *l'efficacité* de l'apprentissage). Pour le versant « informatif », les propositions sont assez proches de celles de la catégorie précédente (évaluation *de* l'apprentissage), mais on perçoit ici une volonté de dépasser le simple bilan pour informer de façon plus précise un destinataire plus souvent identifié (*L'évaluation permet aux élèves de faire le point de manière personnelle sur leurs acquis et faiblesses; L'évaluation permet au professeur de vérifier que sa méthode*

d' [enseignement/]apprentissage est efficace ; À l'aide d'une évaluation, on peut vérifier si l'élève a une bonne méthode d'apprentissage). La régulation se situe majoritairement du côté des pratiques pédagogiques (*Réajuster sa façon d'enseigner après l'évaluation*). Seules trois propositions impliquent explicitement l'élève (*Les évaluations permettent aux élèves de voir où ils en sont dans leurs apprentissages et d'améliorer ce qui doit l'être*).

Enfin, 4% des propositions ont pu être situées dans la catégorie « Évaluation en tant qu'apprentissage ». La plupart des propositions mettent l'accent sur les techniques ou stratégies d'apprentissage (*L'évaluation permet à l'élève d'acquérir des stratégies d'apprentissage de la matière ; Les démarches à mettre en œuvre pour acquérir l'apprentissage peuvent être suggérées par le biais de l'évaluation*). Certaines propositions invitent à un recul métacognitif (*L'évaluation permet à l'élève de savoir si ses méthodes d'apprentissage personnelles sont valables, efficaces*), voire à une régulation (*L'évaluation permet de revoir les méthodes d'apprentissage*). D'autres propositions mentionnent l'idée selon laquelle l'évaluation constitue une méthode pour apprendre ou pour évoluer (*L'évaluation est elle-même une forme d'apprentissage*) ou fait partie intégrante de l'apprentissage (*L'évaluation fait partie intégrante de l'apprentissage et doit être un élément contribuant à la formation de l'élève*). Quelques propositions mentionnent explicitement l'autoévaluation, parfois de façon très générale (*L'évaluation fait partie intégrante de l'apprentissage [autoévaluation, métacognition]*), et quelques autres évoquent sa fonction de régulation (*Durant l'apprentissage, tout étudiant a la possibilité de s'autoévaluer et cibler son étude sur les difficultés rencontrées*).

Parmi les nombreuses propositions non reprises dans le tableau de synthèse, seules deux propositions font référence à une évaluation normative consistant à classer les élèves ou à les comparer (*L'évaluation permet de comparer le degré d'apprentissage des élèves d'une même classe*). Par contre, 41 propositions (soit 7%) interrogent le rôle de l'évaluation comme moteur à l'apprentissage (autre élément caractérisant un climat évaluatif normatif favorisant les buts de performance). Si 29 propositions soutiennent que « *L'évaluation est nécessaire pour encourager le processus d'apprendre* », 12 propositions sous-tendent au contraire qu'il peut s'agir d'un frein (*L'évaluation peut freiner l'apprentissage en démotivant l'élève*). Une autre série de propositions (environ 5%) se regroupent autour d'une idée de « temporalité ». Généralement, l'idée est que l'évaluation doit se faire

« après » l'apprentissage, ce qui peut faire penser à une évaluation certificative plutôt que formative, selon la distinction héritée des travaux de Bloom en pédagogie de la maîtrise (Allal & Mottier Lopez, 2005).

Enfin, l'analyse inductive a permis de dégager d'autres catégories, dont les plus significatives expriment l'idée selon laquelle l'évaluation ne reflète pas toujours l'apprentissage (voire que c'est impossible de réellement évaluer les apprentissages) ou selon laquelle d'autres facteurs interviennent, notamment la qualité de l'enseignement. D'autres propositions s'articulent autour de l'idée qu'il y a plusieurs façons d'évaluer et plusieurs objets possibles d'évaluation et/ou précisent que l'évaluation nécessite une couverture de la matière et de coller aux objectifs. Plusieurs propositions se réfèrent au fait que l'évaluation permet de fixer l'apprentissage ou qu'un bon apprentissage doit permettre de réussir l'évaluation. Une série de propositions relativement isolées expriment des idées variées, par exemple le fait que l'évaluation ne concerne que quelques points de passage alors qu'on apprend toute sa vie ; qu'il est complexe pour l'enseignant de jouer à la fois le rôle d'aide à l'apprentissage et d'évaluateur ; qu'évaluation et apprentissage préparent au monde du travail ; ou encore que l'apprentissage peut être différent, mais que l'évaluation doit être la même pour tous. Enfin, notons qu'environ 7% des réponses correspondent à des omissions, une trentaine d'étudiants n'ayant réussi à trouver qu'une ou deux propositions reliant les termes évaluation et apprentissage.

Expériences d'évaluation formative vécues en tant qu'élève

Les expériences d'évaluation formative vécues en tant qu'apprenant à l'école secondaire ont été classées selon deux dimensions : la première fait référence au type de situation rencontrée et est liée aux propositions du questionnaire⁴, tandis que la seconde, issue de l'analyse inductive, reprend les caractéristiques principales de ces situations (voir Tableau 3). Les étudiants témoignent d'expériences limitées en la matière : seules 104 propositions sur 187 (56%) ont ainsi pu être catégorisées, les autres n'étant pas suffisamment précises ou témoignant d'une confusion avec d'autres types d'évaluations. Notons aussi qu'une proportion non négligeable d'étudiants indiquent (8%) n'avoir jamais vécu ce type d'expérience au secondaire et qu'une série d'étudiants (6%) n'écrivent rien, sans en préciser la raison.

Tableau 3
*Répartition des expériences d'évaluation formative vécues
 en tant qu'apprenant*

	Évaluation formative	Autoévaluation ou coévaluation	Évaluation mutuelle
Préparation à l'évaluation certificative	7	6	–
Expériences de non-notation/correction	4	6	32
Estimation/commentaires formatifs	–	25	16
Accent sur régulation et/ou métacognition	5	3	–

Préparation à l'évaluation certificative

En ce qui concerne les expériences de « préparation à l'évaluation certificative », parmi les exemples d'évaluation formative, on trouve l'idée que c'est une évaluation qui se déroule « avant » l'évaluation certificative et qui vise à s'y préparer (*Dans mon cours de français, on nous donnait l'occasion d'être évalué de manière formative régulièrement avant l'évaluation certificative*), à voir où on en est (*Dans mon école secondaire, il arrivait quelques fois que les professeurs fassent des évaluations formatives pour, d'après eux, voir où nous en étions et qui ne compteront pas dans le bulletin*) dans des conditions très proches de celles qui seront à l'œuvre lors de l'examen (*On a dû se préparer pour le cours de maths comme si on passait le véritable examen de juin*). Elles permettent de cerner les attentes de l'enseignant, notamment en ce qui concerne la précision des réponses attendues (*L'évaluation formative [...] m'a toujours permis de comprendre ce que le professeur voulait qu'on apprenne et ce qui était considéré comme important*). En matière d'autoévaluation, on trouve des expériences où l'enseignant fournit des exemples de questions que les élèves doivent réaliser seuls (en classe ou à domicile). L'enseignant procure ensuite le corrigé ou réalise la correction au tableau (*Les enseignants donnent aux élèves une préparation et ceux-ci la corrigent en classe*). On suppose que l'autoévaluation se traduit ici par l'idée que l'élève se trouve seul face à la tâche à réaliser de façon à voir où il en est.

Expériences de non-notation ou de correction

La catégorie « expériences de non-notation/correction » regroupe les descriptions qui mettent explicitement l'accent sur ces éléments. En matière d'évaluation formative, il s'agit alors de non-notation (*Au cours suivant, les élèves faisaient une interro « surprise » mais formative. Elle ne comptait pas dans la moyenne de l'année*). En ce qui concerne l'autoévaluation et la coévaluation, l'accent est mis sur l'autocorrection (*L'enseignant écrit les réponses au tableau et les élèves corrigent chacun leur propre feuille*), avec une éventuelle confrontation avec la note de l'enseignant (*Après une évaluation de version latine, notre prof [...] nous a expliqué son système de notation et de correction des fautes. Nous nous sommes corrigés nous-mêmes. Ensuite, notre prof a repris les feuilles afin de confronter nos deux évaluations*). En matière d'évaluation mutuelle, il s'agit essentiellement de corriger la feuille d'un pair ; l'accent est mis sur la correction (juste ou faux), et non sur l'éventuelle rétroaction qui pourrait être donnée pour pointer les forces et faiblesses de la copie. Dans la majorité des cas, il est assez clair qu'un corrigé est donné (*La correction se faisait en classe directement après l'évaluation et chacun corrigeait la feuille de son voisin selon les bonnes réponses fournies par le professeur*) ou que l'enseignant intervient pour orienter la correction (*On ramassait les feuilles et on les redistribuait au hasard ; on corrigeait tous ensemble en classe, un élève corrigeant la feuille d'un autre, tout en corrigeant oralement avec le prof*).

Estimation/commentaires formatifs

Certaines expériences dépassent cette logique de « simple correction » et sont reprises dans une catégorie que nous avons choisi d'intituler « estimation/commentaires formatifs » pour la distinguer de la précédente. Les exemples d'autoévaluation et de coévaluation se limitent généralement à estimer sa note (*En 5^e secondaire, à la fin d'un examen oral de néerlandais, le prof nous demandait si on avait le sentiment d'avoir répondu correctement et si nous pensions savoir nous exprimer correctement. Nous devions nous coter nous-mêmes sur une échelle de 1 à 10*) et parfois à assortir ses réponses de degrés de certitude (*Au cours de sciences, nous devons évaluer nos réponses par des coefficients de certitude*). Ce qui distingue les exemples repris ici de ceux de la catégorie précédente, c'est que la correction semble dépasser le juste/faux et que les critères de correction ne semblent pas fournis d'emblée. C'est alors aux élèves de rechercher les forces et faiblesses de leur production (*Autoévaluation d'une interro sous forme d'un QCM sur*

une matière. Il s'agissait d'estimer son score total et donc où se trouvent les bonnes réponses certaines et les réponses possiblement fausses). Dans certains cas, la pertinence de l'autoévaluation est valorisée et entre en compte dans la note finale (*Lors des évaluations dans le cadre des cours de maths en 5^e secondaire, pour chaque question, nous devons nous autoévaluer. Ensuite, nous devons nous mettre une cote finale. Si celle-ci était identique à notre note finale, accordée par le prof, nous avons un point bonus*). La plupart des exemples de coévaluation se distinguent peu de ceux d'autoévaluation (la confrontation avec la note de l'enseignant pouvant être mentionnée dans les deux cas). Les discussions entre l'élève et l'enseignant pourraient être riches et permettre d'impulser une certaine réflexion métacognitive, mais ce type d'argument n'est généralement pas mentionné. Les exemples d'expériences d'évaluation mutuelle nécessitent généralement d'explicitier les éléments sur lesquels se fonde l'évaluation. Par exemple, lorsqu'il s'agit de commenter la production d'autrui, il est important de préciser les forces et les faiblesses de celle-ci (*Dissertations orales et exposés en français: [...] les élèves étaient impliqués dans une discussion soulignant les forces et les faiblesses de la production*). Lorsqu'il s'agit d'attribuer une note à autrui, il s'agit de la justifier (*Un camarade et moi-même devons faire un dialogue en anglais devant la classe. À la fin, le prof nous a demandé d'évaluer notre camarade et de justifier notre cotation*) et/ou d'expliquer les erreurs (*Durant un cours d'anglais, le prof a donné un test de grammaire et vocabulaire. Au bout de 30 minutes, on a échangé notre copie avec notre voisin et on l'a corrigée, puis on devait expliquer les erreurs faites*). Même si ce n'est pas mentionné explicitement, l'objectif est sans doute de retravailler ensuite sa propre production sur cette base et donc de procéder à une certaine régulation.

Accent sur régulation et/ou métacognition

Quelques exemples mettent l'accent sur les buts poursuivis en matière de régulation et/ou sur le développement de réflexions métacognitives. En matière d'expériences d'évaluation formative, la régulation se situe du côté de l'enseignant (*L'enseignant demande aux élèves de revoir la matière pour le cours suivant durant lequel ils auront un test sur la matière vue; l'enseignant corrige le test et propose ensuite des exercices sur les points de matière qui étaient le moins bien compris par l'ensemble de la classe*) ou du côté de l'élève (*En 1^{re} secondaire, un de mes profs nous permettait, après avoir corrigé la première interro formative, de pouvoir repasser autant de*

fois qu'on le devait les interros pour les réussir). En ce qui concerne l'autoévaluation, deux propositions évoquent un recul métacognitif, où l'étudiant est invité à s'interroger sur ses méthodes d'apprentissage (*À la fin de notre journal de classe, nous avons des fiches comportant différentes questions sur notre méthode de travail à domicile, le travail réalisé en classe, les résultats aux tests. Ces fiches permettaient d'évaluer nous-mêmes la façon dont nous percevions notre travail*) ou sur ses lacunes et sur la façon de s'améliorer (*L'autoévaluation m'a permis de mieux comprendre vers où j'arrivais. Au fur et à mesure que j'analysais les réponses, je comprenais mieux mes lacunes et où je devais m'améliorer*). En matière de coévaluation, le seul exemple entrant dans cette catégorie évoque une expérience de mathématiques durant laquelle le professeur réalise une première correction et fournit des explications complémentaires en cas d'erreur ; l'élève peut alors reprendre sa copie (réguler), mais il perd un point.

Les autres expériences décrites n'ont pas pu être classées dans le tableau : certaines semblent témoigner d'une confusion entre évaluation formative et évaluation certificative classique (14%) ou d'autres types de confusion (7% ; l'évaluation mutuelle est notamment confondue avec des travaux de groupe). D'autres sont peu compréhensibles ou trop peu informatives pour en saisir l'enjeu (7%). Quelques exemples présentent des expériences visant à se préparer à une évaluation certificative menées à l'initiative de l'étudiant (se poser des questions ou revoir la matière ensemble comme exemples d'évaluation mutuelle) et ne sont pas repris dans le tableau (2%).

Projections en tant que futur enseignant

En invitant les futurs enseignants à imaginer ce qu'ils pourraient mettre en place pour évaluer les apprentissages de leurs élèves sans recourir à une évaluation traditionnelle, nous visons ici plus particulièrement à voir s'ils envisagent d'intégrer l'évaluation dans la situation d'enseignement/apprentissage et d'impliquer chacun des élèves dans le processus d'évaluation. Le tableau 4 présente les différentes projections proposées par les futurs enseignants, en fonction des modalités organisationnelles du travail (collectif, en groupe ou individuel).

Tableau 4
Répartition des projections en tant que futur enseignant

	Modalité organisationnelle du travail		
	Collectif	En groupe	Individuel
Questions-réponses	100	–	–
Participation active	8	–	–
Présentation orale	–	17	11
Travail écrit	–	13	15
Partage des forces et faiblesses	–	5	–
Élaboration d'une synthèse	13	–	–
Exercices	30	–	13
Évaluation formative (dont autoévaluation et coévaluation)	–	–	8

Notons d'emblée que, si trois étudiants ne donnent pas de réponses (deux d'entre eux indiquent clairement qu'ils ne voient pas comment évaluer les apprentissages autrement que par une interrogation écrite ou orale traditionnelle), d'autres en donnent plus d'une, ce qui porte le nombre total de projections répertoriées dans le tableau 4 à 233.

Les futurs enseignants imaginent des situations d'évaluation assez classiques qui, dans la très grande majorité des cas (100 projections sur 233, soit 43%), s'en tiennent à une évaluation des connaissances déclaratives, proposée de façon collective (*Inclure quelques questions orales ponctuelles de certains points vus en classe; Je les interrogerai oralement à chaque début de cours pour savoir ce qu'ils ont retenu des cours précédents*). Parmi ces 100 exemples de questions-réponses collectives, 21 mentionnent l'idée de poser des questions d'évaluation à chacun des élèves, les autres se limitant à prendre en compte le collectif-classe. Soulignons également que 8 projections (soit 3%) font référence à l'idée selon laquelle attester de la participation active des élèves constituerait la meilleure façon d'évaluer leurs apprentissages.

Près d'un quart des dispositifs proposés (56 projections) font référence à des travaux écrits (*Un devoir fait à la maison peut permettre à l'enseignant d'évaluer les apprentissages des élèves*) ou à des présentations orales

(*Je donnerais quelques consignes aux élèves intégrant les points de grammaire et vocabulaire vus au cours et ceux-ci devraient réaliser une saynète ou un sketch en groupe en appliquant ces consignes*), à préparer en classe ou à domicile et à réaliser en groupe ou individuellement. Dans la plupart des exemples, les tâches semblent être proposées en plus ou en marge des autres activités d'enseignement/apprentissage, comme s'il était nécessaire de prévoir une tâche différente pour la dédier spécifiquement à l'évaluation. Dans la colonne se référant aux évaluations réalisées en groupe, on trouve aussi quelques rares exemples (5 projections, soit 2%) dans lesquels les futurs enseignants expliquent qu'un travail de groupe est mis en place pour inviter les élèves à partager leurs points forts et/ou leurs difficultés (*Je mettrai des moyens d'évaluation plus interactifs faisant intervenir plusieurs groupes d'élèves afin qu'ils se présentent leurs difficultés et diverses compréhensions de la matière*). Là encore, les tâches imaginées sont spécifiquement dédiées à l'évaluation.

A contrario, dans environ un quart des cas (56 projections), les futurs enseignants développent des stratégies visant à intégrer l'évaluation au cœur de la leçon, dans un processus collectif (43 projections) ou individuel (13 projections). Dans les activités collectives, on peut distinguer les tâches consistant à élaborer une synthèse en commun (*Réalisation d'une synthèse en commun. L'objectif est de les faire s'exprimer sur la matière vue en mobilisant certaines compétences, sans qu'une cotation soit en jeu*) de celles consistant à réaliser des exercices tous ensemble, généralement avec l'appui du tableau (*En distribuant des exercices que certains élèves viendront résoudre au tableau, on peut notamment évaluer le degré de compréhension de ceux au tableau. Si, en plus de cela, lorsqu'un exercice est incorrect, on demande si un autre élève peut venir corriger, on peut éventuellement repérer une lacune dans l'apprentissage de la classe*). Si ces projections visent à permettre à l'enseignant de repérer les difficultés des élèves, voire d'opérer différents gestes de régulation, elles ne permettent pas une évaluation des apprentissages de *chacun* des élèves. Les 13 projections classées dans la colonne retraçant des activités individuelles se donnent davantage de chances de répondre à ce critère. Le recours à des tâches à réaliser seul permet à chaque élève de voir où il en est (*Je mets en place un exercice permettant à l'élève de voir si oui ou non il maîtrise les différentes compétences demandées*). Cette modalité de travail permet aussi d'informer l'enseignant (dans la mesure où il circule dans la classe et observe les élèves au travail),

voire lui donne l'occasion de procéder à des régulations (*Réaliser des exercices individuellement, donner les corrections, les élèves ayant des erreurs doivent se manifester et je réexplique ce qui n'a pas été compris*).

Enfin, 8 projections (soit 3%) font explicitement référence à des tâches d'évaluation à caractère formatif, intégrées ou non au cœur de la leçon et tentant d'impliquer chaque élève dans le processus d'évaluation, certes à des degrés divers. Lorsque l'évaluation formative est citée, c'est le caractère informatif qui est généralement prégnant (*Je mettrai en place une évaluation formative afin que les élèves et moi-même sachions où ils en sont dans leur apprentissage. Cela permettrait de voir ce qui est acquis et ce qui doit encore être travaillé*). La coévaluation est quant à elle envisagée en tant que modalité permettant des discussions, voire des conseils (*Je mettrai en place un système de coévaluation, une discussion ouverte qui permettrait aux élèves de se sentir moins jugés ou humiliés par une mauvaise note et qui proposerait plutôt des remarques/critiques/conseils utiles et pertinents*). Enfin, l'occasion de pratiquer une autoévaluation est évoquée quatre fois (*Je mettrais en œuvre des autoévaluations d'apprentissage où les élèves devraient noter sur une échelle ordinale leur sentiment sur des apprentissages acquis ou non ; On peut développer une autoévaluation avec un espace où les élèves pourraient exprimer les points non compris*).

Conclusion et discussion

La première question de recherche visait à éclairer les liens entre évaluation et apprentissage au moyen de trois « instruments » complémentaires.

Sur la base du questionnaire tiré des études d'Issaieva et de ses collaborateurs (2010, 2013 et 2015), les résultats révèlent une structuration des conceptions comprenant deux dimensions : une dimension « formative et certificative », d'une part, et une dimension « normative », d'autre part. Cette structuration s'accorde à celle observée chez les enseignants suisses plutôt qu'à celle correspondant aux enseignants belges ou encore aux étudiants suisses. Pour rappel, la structuration des conceptions observée auprès des futurs enseignants suisses distinguait les trois fonctions de l'évaluation ; l'hypothèse formulée par les auteurs étant qu'ils s'appuyaient sans doute sur les connaissances théoriques et sur les idéaux développés au cours de leur formation (Boraita & Issaieva, 2013). Notre étude se situant

à l'entrée en formation, nous pouvons penser que les futurs enseignants s'appuient plutôt sur le système scolaire qu'ils ont connu en tant qu'élève et sur la façon dont ils se projettent en tant qu'enseignant. Pourquoi se distinguent-ils alors des enseignants belges ? Pour rappel, ceux-ci opposaient les fonctions « formative » et « normative », les items relevant du versant critérié de l'évaluation certificative n'entrant pas dans la structure factorielle. Pour expliquer ces résultats, Issaieva et ses collaborateurs (2015) font référence aux tensions éprouvées par les enseignants entre, d'une part, les prescriptions pédagogiques (soutenues par les responsables politiques) valorisant l'évaluation formative pour lutter contre l'échec scolaire et, d'autre part, les contraintes d'un système scolaire historiquement à dominante sélective, comme en témoigne notamment la pratique courante du redoublement (Marcoux & Crahay, 2008 ; Draelants, 2006). L'omniprésence des notes dans le système scolaire et la tendance à les interpréter de manière normative (Merle, 2014) pourraient aussi expliquer ces résultats. Dans notre étude, les futurs enseignants n'étant pas encore aux prises avec la gestion de ces tensions, nous pouvons estimer que la structuration de leurs conceptions oppose les fonctions pédagogiquement attendues de l'évaluation (fonctions formative et certificative) à celles qui pourraient être considérées comme moins nobles, comme le classement des élèves et la sanction (fonction normative). Les valeurs moyennes des échelles, témoignant d'un accord net envers les fonctions formative/certificative (score moyen de 4,69) et d'un rejet tout aussi net de la fonction normative (score moyen de 2,71), corroborent cette hypothèse interprétative (voir Tableau 1).

Le deuxième questionnaire fermé a lui aussi mis en évidence une structuration des conceptions en deux facteurs, opposant cette fois l'évaluation à visée formative (évaluation *au service de* l'apprentissage et *en tant qu'*apprentissage) et l'évaluation à visée certificative (évaluation *de* l'apprentissage). Alors que, dans la première échelle, l'évaluation normative semble jouer le rôle de « repoussoir » contre l'ensemble des autres items, un questionnaire centré sur les fonctions pédagogiquement attendues de l'évaluation semble permettre aux futurs enseignants de distinguer clairement les versants formatif et certificatif. Les valeurs moyennes assez élevées (4,51 et 4,11 ; voir Tableau 1) soulignent la reconnaissance de l'importance de ces deux fonctions. En ce sens, les résultats s'accordent avec les propos de Mottier Lopez (2015), qui invite à combiner harmonieusement ces deux fonctions en considérant que « l'évaluation est un moyen d'ensei-

nement, parmi d'autres, dans les mains du maître» et que c'est aussi «un moyen pour apprendre, parmi d'autres, dans les mains de l'élève» (p. 30).

Rappelons toutefois que ces résultats doivent être relativisés, dans la mesure où les questionnaires fermés sont aisément soumis à un phénomène de désirabilité sociale et à un biais de questionnement (Butori & Parguel, 2010) pouvant conduire les répondants à marquer un accord de principe avec les propositions pédagogiquement attendues, quelles que soient en réalité leurs conceptions profondes sur ces questions. D'ailleurs, la fonction formative «classique» (évaluation *au service de* l'apprentissage), plébiscitée dans les différentes études appuyées sur des questionnaires fermés (Issaieva et al., 2011 ; Luisoni et Monnard, 2015), est moins mentionnée spontanément par les futurs enseignants, lorsqu'on leur demande de relier les concepts d'évaluation et d'apprentissage, que la fonction certificative. L'évaluation formative élargie à l'idée d'évaluation *en tant qu'*apprentissage est, quant à elle, très peu mentionnée, alors que les étudiants marquent un accord prononcé dans le questionnaire fermé avec les items pointant l'importance de cette fonction. Par ailleurs, alors qu'Issaieva et Crahay (2010) notaient que les enseignants distinguaient généralement deux dimensions à la fonction formative («diagnostiquer les difficultés d'apprentissage des élèves et viser à installer un processus régulateur de l'enseignement dans le but d'aider les élèves», p. 51), le versant «régulation» est relativement peu mentionné spontanément par les futurs enseignants puisqu'il ne concerne que 5% des propositions⁵.

C'est au vécu en tant qu'élève au secondaire que s'intéressait notre deuxième question de recherche. La question ouverte les invitait à faire référence aux expériences d'évaluation formative élargie aux expériences d'autoévaluation, de coévaluation ou d'évaluation mutuelle vécues en tant qu'apprenant. Les résultats témoignent d'expériences assez limitées en la matière. Parmi les expériences relatées, nombreuses sont celles qui font référence à des pratiques d'autonotation ou de conotation (Rey & Feyfant, 2014). Lorsque l'accent est mis sur la correction, les élèves sont centrés sur la confrontation de leurs réponses avec un référentiel ou un corrigé donné par l'enseignant. Ce type d'autoévaluation n'appelle pas un réel recul métacognitif et conduit généralement à des corrections de surface, plutôt qu'à une réelle régulation des processus cognitifs impliqués (Campanale, 2007). L'évaluation mutuelle devrait s'avérer plus porteuse,

dans la mesure où elle offre l'occasion d'amener les élèves à expliciter et à confronter leurs démarches. Elle devrait dès lors susciter davantage d'interactions verbales favorisant les régulations interactives entre pairs (Mottier Lopez, 2015). Toutefois, lorsqu'elle est exclusivement orientée sur la « correction » de la copie d'autrui au départ d'une grille ou d'un corrigé donné par l'enseignant (comme cela semble être le cas dans la majorité des expériences décrites), l'évaluation mutuelle rencontre les mêmes travers que l'autocorrection et engage peu au recul métacognitif et à la régulation. En ce sens, des « questions ouvertes relatives aux critères de réalisation essentiels, suivies d'une rubrique "propositions d'amélioration", sollicitent davantage les interactions verbales » (Campanale, 2007, p. 200) et devraient s'avérer plus porteuses. On peut penser que les quelques expériences reprises dans la catégorie « commentaires formatifs » liée à l'évaluation mutuelle s'approchent de cette proposition, sans être aussi précises. Enfin, parmi les expériences décrites, un recul métacognitif et la régulation ne sont que très rarement évoqués, et on voit mal comment l'autonotation ou la conotation permettraient d'atteindre des objectifs aussi ambitieux.

La troisième question de recherche a mis en évidence que, lorsqu'on empêche les futurs enseignants d'imaginer une évaluation classique, très peu projettent une vision « élargie » d'une évaluation à visée formative, qui pourrait s'intégrer au cœur même des activités d'enseignement/apprentissage et qui viserait à impliquer chacun des élèves dans le processus d'évaluation, notamment par de l'autoévaluation ou de l'évaluation mutuelle (Allal, 2007 ; Allal & Mottier Lopez, 2005 ; Mottier Lopez & Laveault, 2008).

Quelques croisements (non présentés ici) entre les trois volets du questionnaire ouvert révèlent qu'il ne paraît ni nécessaire ni suffisant d'avoir vécu des expériences d'évaluation formative *au service de* l'apprentissage ou *en tant qu'*apprentissage pour répondre en ce sens aux autres parties du questionnaire. Nous pouvons toutefois penser que le fait de vivre peu d'expériences d'évaluation formative stimulantes au secondaire n'aide pas les futurs enseignants à imaginer la multiplicité des rôles que peut jouer l'évaluation. S'il est important que la formation initiale prenne en compte les conceptions des futurs enseignants lorsque divers contenus relatifs à l'évaluation sont travaillés, il semble tout aussi essentiel de faire vivre aux futurs enseignants des expériences d'évaluation formative riches et variées,

puis de les amener à réfléchir sur les éléments essentiels à prendre en compte pour transposer ce vécu d'étudiant universitaire vers leurs pratiques d'enseignants (Campanale, 2007 ; Goffin & Fagnant, à paraître). Inviter les enseignants à réfléchir sur leurs pratiques, tout au long de leur carrière, constitue également un enjeu important (Mottier Lopez, 2016). Sans cela, on peut en effet craindre que la réalité de l'exercice du métier au quotidien conduise *in fine* les enseignants à se concentrer sur la seule fonction certificative qui, si elle est certes importante, ne devrait sans doute pas être la seule à laquelle sont couramment soumis les élèves (Soussi et al., 2006 ; Younès & Gaime, 2012).

En fin de compte, si les futurs enseignants s'accordent positivement avec les propositions relatives aux différents versants de l'évaluation formative dans les deux questionnaires fermés, les réponses aux questions ouvertes témoignent assez peu d'une perspective « élargie » de l'évaluation formative. Les futurs enseignants semblent peiner à dépasser la vision de l'évaluation (qu'elle soit formative ou certificative) comme relevant d'un événement spécifique se présentant sous la forme d'un test papier-crayon pour tendre vers une vision plus interactive ou informelle (Allal & Mottier Lopez, 2005 ; Morissette, 2010 ; Morissette & Compaoré, 2012) impliquant davantage l'élève dans le processus (Allal & Mottier Lopez, 2005 ; Campanale, 2007 ; Mottier Lopez, 2015). Dans une étude exploratoire précédente, nous avons montré qu'une conception « ouverte » de l'évaluation formative, imaginant d'intégrer l'évaluation au cœur des activités d'enseignement/apprentissage, n'était pas aisément perçue par les futurs enseignants du secondaire, même après un an de formation (Goffin et al., 2014). À nos yeux, il s'agit pourtant là d'un enjeu important si l'on souhaite lutter contre la tendance qui consiste à justifier la non-implantation de pratiques d'évaluation formative dans les classes par « la complexité et la lourdeur » qui caractérise sa mise en œuvre (Boraita & Issaieva, 2013) ou par la non-motivation des élèves à réaliser des tests qui ne comptent pas pour des points du bulletin.

Cette étude présente évidemment certaines limites, liées notamment aux choix posés dans les mises en situation proposées pour les questions ouvertes. En effet, celles-ci ne sont jamais neutres. Par exemple, le choix de demander de relier les concepts d'évaluation et d'apprentissage pourrait, d'une certaine façon, avoir orienté les réponses. Ce choix se justifiait par la volonté d'essayer de distinguer évaluation *de* l'apprentissage, *au service de*

l'apprentissage et *en tant qu'*apprentissage. C'est d'ailleurs dans l'optique de «forcer» les étudiants à distinguer différentes fonctions que nous avons demandé qu'ils formulent trois propositions. Dans cette logique, c'est essentiellement sous ces trois angles que nous avons analysé ces propositions. Même si l'analyse inductive a permis de mettre en avant d'autres catégories significatives, ces informations d'une grande richesse restent en partie sous-exploitées. Par ailleurs, même si les exemples relatés sont produits par des étudiants venant de disciplines diversifiées et même si les expériences relatées s'inscrivent dans des disciplines différentes, une limite de l'étude réside dans le fait d'avoir réalisé les analyses toutes disciplines confondues. Ce choix a été fait dans la mesure où cette situation traduit la réalité de certains cours de la formation où les étudiants sont ainsi «mêlés», notamment pour parler d'évaluation. Si ce type d'analyse non disciplinaire semblait dès lors défendable dans le cadre du présent article, une telle situation n'est néanmoins pas tenable lorsqu'il s'agit de s'intéresser plus finement aux qualités des évaluations elles-mêmes, notamment à leur potentiel diagnostique nécessitant de cerner les processus cognitifs mis en œuvre dans des tâches spécifiques. En effet, évaluer (que ce soit de façon certificative ou formative, formelle ou informelle) nécessite une réelle prise en compte des contenus disciplinaires impliqués ainsi que des connaissances didactiques propres à ces disciplines (Fagnant & Demonty, 2017; Vantourout & Goasdoué, 2014).

Tout en restant prudentes quant aux limites liées au caractère exploratoire de l'étude présentée dans cet article, nous pensons que la combinaison de questionnaires fermés et de questions ouvertes (s'intéressant notamment au vécu en tant qu'élève et invitant à se projeter dans un rôle d'enseignant) permet d'apporter un éclairage original à la problématique des conceptions relatives à l'évaluation. Cette approche permet non seulement de relativiser les accords de principe trop souvent rencontrés dans les questionnaires fermés par rapport aux fonctions pédagogiquement attendues de l'évaluation, mais aussi de mettre en question une vision «élargie» de l'évaluation formative, généralement peu présente dans ce type d'enquête.

Réception : 19 février 2016

Version finale : 2 décembre 2016

Acceptation : 7 décembre 2016

NOTES

1. Plusieurs auteurs établissent de nombreux liens entre évaluation sommative et certificative (Allal, 2007 ; Mottier Lopez & Laveault, 2008), les conduisant parfois, à la suite de Tourneur (1985, cité par Mottier Lopez & Laveault, 2008), à privilégier l'appellation d'évaluation «certificative» à celle d'évaluation «sommative». C'est également ce choix qui est posé par Mottier Lopez (2015) dans son ouvrage intitulé *Évaluations formative et certificative des apprentissages*. Dans la lignée de ces travaux, c'est le terme «certificatif» que nous utilisons dans le texte.
2. Même si l'on peut suivre la logique des auteurs qui les conduit à opposer ces deux types d'évaluation, on peut aussi penser que les résultats d'un bilan, souvent présentés sous la forme d'une note chiffrée, peuvent être interprétés de manière normative, et ce, même si cette note chiffrée est le fruit d'une analyse critériée. En effet, les notes donnent une indication (certes imparfaite) du niveau des élèves; elles leur permettent de savoir où ils se situent par rapport à la moyenne, mais aussi les uns par rapport aux autres (Merle, 2014 ; Younès & Gaime, 2012).
3. Les étudiants qui disposaient déjà d'une pratique professionnelle ont été écartés, dans la mesure où l'on peut penser que celle-ci a déjà sans doute influencé leurs conceptions en matière d'évaluation.
4. Les expériences d'autoévaluation et de coévaluation ont été regroupées dans le tableau et dans les analyses parce qu'elles sont peu différenciées dans les propos des étudiants.
5. L'étude de Monney et Fontaine (2016), appuyée elle aussi sur des questions ouvertes, donne quant à elle des résultats assez différents de ceux observés dans la présente étude. Les chercheuses ont invité les futurs enseignants à produire librement cinq mots ou expressions qui portent sur l'évaluation des apprentissages, puis à hiérarchiser ces propositions en fonction de l'importance qu'ils accordent à chacune d'elles. Leurs résultats montrent que la fonction de «régulation» fait partie du noyau central des conceptions des futurs enseignants, alors que la fonction «certificative», tout en étant bien présente, se situe plutôt dans le noyau périphérique. Précisons toutefois que l'étude de Monney et Fontaine porte sur des étudiants qui terminent leur formation à l'enseignement, ce qui peut expliquer les différences observées puisque leurs conceptions peuvent ici être structurées selon les attendus théoriques de la formation, comme dans l'étude de Boraita et Issaieva (2013).

RÉFÉRENCES

- Allal, L. (2007). Évaluation : lien entre enseignement et apprentissage. Dans V. Dupriez et G. Chapelle (dir.), *Enseigner* (pp. 139-149). Paris : Presses universitaires de France.
- Allal, L. et Mottier Lopez, L. (2005). L'évaluation formative de l'apprentissage : revue de publications en langue française. Dans Centre pour la recherche et l'innovation dans l'enseignement (dir.), *L'évaluation formative pour un meilleur apprentissage dans les classes secondaires* (pp. 265-290). Paris : OCDE.
- Blais, M. et Martineau, S. (2006). L'analyse inductive générale : description d'une démarche visant à donner un sens à des données brutes. *Recherches qualitatives*, 26(2), 1-18. Repéré à [www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero26\(2\)/blais_et_martineau_final2.pdf](http://www.recherche-qualitative.qc.ca/documents/files/revue/edition_reguliere/numero26(2)/blais_et_martineau_final2.pdf)
- Boraita, F. et Crahay, M. (2013). Les croyances des futurs enseignants : est-il possible de les faire évoluer en cours de formation initiale et, si oui, comment ? *Revue française de pédagogie*, 183, 99-158. Repéré à <http://orbi.ulg.ac.be/handle/2268/184128>
- Boraita, F. et Issaieva, É. (2013, janvier). *Les conceptions de l'évaluation chez les futurs enseignants et enseignants en exercice : une étude dans le canton de Genève*. Communication présentée au 25^e Colloque international de l'ADMEE-Europe, Fribourg (Suisse).
- Butori, R. et Parguel, B. (2010). *Les biais de réponse : impact du mode de collecte des données et de l'attractivité de l'enquêteur*. France : AFM. Repéré à <https://halshs.archives-ouvertes.fr/halshs-00636228/document>
- Campanale, F. (2007). Évaluation réflexive en formation professionnelle et évaluation interactive dans les classes. Dans L. Allal et L. Mottier Lopez. (dir.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 191-206). Bruxelles : De Boeck.
- Crahay, M., Wanlin, P., Issaieva, É. et Laduron, I. (2010). Fonctions, structuration et évolution des croyances (et connaissances) des enseignants : une revue critique de la littérature anglo-saxonne. Note de synthèse. *Revue française de pédagogie*, 172, 85-129. doi: 10.4000/rfp.2296
- Draelants, H. (2006). Le redoublement est moins un problème qu'une solution : comprendre l'attachement social au redoublement en Belgique francophone. *Les Cahiers de recherche en éducation et formation*, 52, 1-25. Repéré à <https://halshs.archives-ouvertes.fr/halshs-00563874/document>
- Fagnant, A. et Demonty, I. (2017). Comment former à une vision élargie de l'évaluation formative en vue de réguler efficacement les apprentissages mathématiques des élèves ? Dans P. Detroz, M. Crahay et A. Fagnant (dir.), *L'évaluation à la lumière des contextes et des disciplines* (pp. 257-273). Bruxelles : De Boeck.
- Goffin, C. et Fagnant, A. (à paraître). Faire vivre une expérience d'évaluation mutuelle en formation initiale : quelle transposition vers les pratiques d'enseignement projetées ? *Contextes et Didactiques*.
- Goffin, C., Renson, J. et Fagnant, A. (2014, janvier). *Comment les futurs enseignants du secondaire conçoivent-ils l'évaluation formative et ses difficultés de mise en œuvre ?* Communication présentée au 26^e Colloque international de l'ADMEE-Europe, Marrakech (Maroc). Repéré à <http://hdl.handle.net/2268/164492>

- Issaieva, É. et Crahay, M. (2010). Conception de l'évaluation scolaire des élèves et des enseignants : validation d'échelles et études de leurs relations. *Mesure et évaluation en éducation*, 33(1), 31-62. doi: 10.7202/1024925ar
- Issaieva, É., Pini, G. et Crahay, M. (2011). Positionnements des enseignants et des élèves face à l'évaluation : une convergence existe-t-elle? *Revue française de pédagogie*, 176, 5-26. doi: 10.4000/rfp.3362
- Issaieva É., Yerly, G., Petkova, I., Marbaise, C. et Crahay, M. (2015). Conceptions et prises de positions des enseignants face à l'évaluation scolaire dans quatre systèmes éducatifs européens : quel est le reflet des cultures et politiques éducatives? Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation : quels espaces de formation?* (pp. 73-98). Louvain-la-Neuve, Belgique : De Boeck Supérieur.
- Luisoni, M. et Monnard, I. (2015). L'évolution des conceptions de l'évaluation scolaire chez les futurs enseignants en formation initiale. Dans P.-F. Coen et L. M. Bélair (dir.), *Évaluation et autoévaluation : quels espaces de formation?* (pp. 175-196). Louvain-la-Neuve, Belgique : De Boeck Supérieur.
- Marcoux, G. et Crahay, M. (2008). Mais pourquoi continuent-ils à faire redoubler? Essai de compréhension du jugement des enseignants concernant le redoublement. *Revue suisse des sciences de l'éducation*, 30(3), 501-518. Repéré à <https://core.ac.uk/download/pdf/33977257.pdf>
- Merle, P. (2014). Faut-il en finir avec les notes? *La vie des idées*, 2 décembre 2014. Repéré à www.laviedesidees.fr/Faut-il-en-finir-avec-les-notes.html
- Monney, N. et Fontaine, N. (2016). La représentation sociale de l'évaluation des apprentissages chez des finissants d'un programme universitaire en éducation préscolaire et en enseignement primaire. *Mesure et évaluation en éducation*, 39(2), 59-85. doi: 10.7202/1038242ar
- Morissette, J. (2010). Un panorama de la recherche sur l'évaluation formative des apprentissages. *Mesure et évaluation en éducation*, 33(3), 1-27. doi: 10.7202/1024889ar
- Morissette, J. et Compaoré, G. (2012). Le savoir-faire enseignant sur l'évaluation formative informelle. *Formation et profession*, 20(3), 26-35. doi: 10.18162/fp.2012.22
- Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages : enjeux pour l'enseignement*. Bruxelles : De Boeck.
- Mottier Lopez, L. (2016). L'évaluation formative des apprentissages des élèves : entre innovations, échecs et possibles renouveaux par des recherches participatives. *Questions vives*, 23, 1-17. doi: 10.4000/questionsvives.1692
- Mottier Lopez, L. et Laveault, D. (2008). L'évaluation des apprentissages en contexte scolaire : développements, enjeux et controverses. *Mesure et évaluation en éducation*, 31(3), 5-34. doi: 10.7202/1024962ar
- Protocole de l'Ouest et du Nord canadiens de collaboration concernant l'éducation. (2006). *Repenser l'évaluation en classe en fonction des buts visés*. Winnipeg: Éducation, Citoyenneté et Jeunesse Manitoba. Repéré à www.edu.gov.mb.ca/m12/frpub/me/docs/repenser_eval
- Rey, O. et Feyfant, A. (2014). Évaluer pour (mieux) faire apprendre. *Dossier de veille de l'Institut français de l'éducation*, 94, 1-44. Repéré à <http://ife.ens-lyon.fr/vst/DA-Veille/94-septembre-2014.pdf>

- Sarrazin, P., Tessier, D. et Trouilloud, D. (2006). Climat motivationnel instauré par l'enseignant et implication des élèves en classe : l'état des recherches. *Revue française de pédagogie*, 157, 147-177. doi: 10.4000/rfp.463
- 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)*. Rapport du Service de la recherche en éducation. Genève : Université de Genève. Repéré à <http://edudoc.ch/record/3603>
- Statistica (version 12) [Logiciel]. (2015). Paris : StatSoft
- Vantourout, M. et Goasdoué, R. (2014). Approches et validité psycho-didactiques des évaluations. *Éducation & Formation*, e-302, 139-156. Repéré à <https://halshs.archives-ouvertes.fr/halshs-01239551/document>
- Vause, A. (2010). Les croyances et connaissances des enseignants de l'école primaire à propos de l'acte d'enseigner. *Éducation & Formation*, 294, 14-19.
- Younès, N. et Gaime, E. (2012). L'évaluation formative en contexte : points de vue d'enseignants, points de vue d'élèves. *Diversité, VEI*, 161-166. Repéré à <https://hal.archives-ouvertes.fr/hal-01121431>