

Pour citer cet article : Goffin, C., Baron, J., Ochelen, J.-P. & Fagnant, A. (2022). Freins et leviers pour œuvrer à la construction et à la communication d'une note à haute valeur informative : une étude exploratoire dans l'enseignement secondaire en Belgique francophone. *La Revue L'EeE*, 6. <https://doi.org/10.48325/rleee.006.04>

*FREINS ET LEVIERS POUR ŒUVRER À LA
CONSTRUCTION ET À LA COMMUNICATION
D'UNE NOTE À HAUTE VALEUR INFORMATIVE
Une étude exploratoire dans l'enseignement secondaire en
Belgique francophone*

Christelle GOFFIN, Jean BARON, Jean-Pascal OCHELEN, Annick FAGNANT

Version de la publication : juillet 2022

Évaluation ouverte et collaborative

Rétroactrices : Mylène DUCREY-MONNIER, Dominique LAFONTAINE

Résumé

Malgré les critiques couramment formulées concernant les notes scolaires (Butera et al., 2011 ; Merle, 2014), celles-ci continuent d'exister. Il ne semble d'ailleurs pas facile de s'en passer et il peut paraître raisonnable d'accepter qu'elles représentent une information incontournable que les élèves obtiennent sur le bilan de leurs apprentissages (Brookhart et al., 2016). Partant de ce constat, l'objectif principal de cette étude exploratoire est d'identifier les conditions qui, selon le prisme des enseignant-es interrogé-es, permettraient de pointer des freins et leviers à prendre en compte pour que la note soit à haute valeur informative (autrement dit, qu'elle constitue une source d'apprentissage pour les élèves). De futur-es enseignant-es ont invité des enseignant-es du secondaire supérieur (grades 10-12) à compléter un questionnaire et, sur cette base, à échanger sur les pratiques d'évaluation. Il ressort des 148 questionnaires analysés que la moitié des enseignant-es considèrent que la note finale ne reflète pas toujours le niveau de maîtrise des élèves, ce qui constitue un frein important à son caractère informatif. Certain-es élèves peuvent effectivement obtenir une note de réussite et pourtant passer à côté de l'objectif essentiel, ce qui se trouve renforcé lorsqu'elle est construite en additionnant les points obtenus pour chacune des questions et/ou des critères, comme le font la plupart des enseignant-es interrogé-es. L'analyse des questionnaires a aussi permis de pointer une série de leviers, parmi lesquels on peut mentionner différents éléments de communication en amont et en aval de l'évaluation.

Mots-clés : évaluation sommative, valeur de la note, soutien à l'apprentissage, enseignement secondaire supérieur, pratiques déclarées

Abstract

Despite the critics of school grades (Butera et al., 2011; Merle, 2014), they continue to exist. Moreover, it does not seem easy to do without them and it may appear reasonable to accept that they represent an inescapable piece of information that students obtain on their learning outcomes (Brookhart et al., 2016). On the basis of these findings, this exploratory study pursues the main objective of identifying the conditions that, according to the teachers interviewed, would allow to point out the threats and opportunities to take into account for the grade to be of high informative value (in other words, for it to become a source of learning for students). Future teachers invited secondary school teachers (grades 10-12) to answer a questionnaire and, on this basis, to discuss assessment practices. From the 148 questionnaires analyzed, it appears that half of the teachers considers that the grade students receive does not always reflect their real educational attainment, which is an important obstacle to its informative value. Indeed, some students may obtain a passing grade and yet miss the essential objective, which is reinforced when the grade is constructed by adding the points obtained for each of the questions and/or criteria, as most of the teachers interviewed do. The analysis of the questionnaires also made it possible to identify some opportunities, among which it is possible to show different elements of communication before and after the evaluation process.

Keywords: summative evaluation, value of the grade, assessment for learning, secondary school, related practices

1. Introduction

Véritable serpent de mer depuis les balbutiements des sciences docimologiques jusqu'à nos jours, la notation des élèves à l'école continue de susciter bien des débats, entre ses défenseurs et ses détracteurs. Malgré les critiques couramment formulées quant aux notes elles-mêmes, que ce soit d'un point de vue docimologique - erreurs de mesure par exemple - ou d'un point de vue psychosocial - impact négatif sur la motivation - (Butera et al., 2011 ; Merle, 2014), elles continuent d'exister en Fédération Wallonie-Bruxelles (FW-B)¹ et il semble très difficile de s'en passer. En effet, il est raisonnable d'accepter que la note représente une information incontournable que les élèves obtiennent sur le bilan de leurs apprentissages (Brookhart et al., 2016).

Avec ou sans note, toutes les parties reconnaissent que le fait d'évaluer les acquis des élèves est une étape importante dans le processus d'apprentissage. On s'accordera dès lors avec Bélair et Coen (2015) pour reconnaître que la formation (initiale et en cours de carrière) des enseignant·es à l'évaluation devrait constituer un enjeu majeur de leur formation professionnelle. D'ailleurs, de nombreuses recherches ont mis en évidence les effets positifs des compétences évaluatives des enseignant·es (*Assessment Literacy*) sur les apprentissages des élèves (Kruse et al., 2020). DeLuca et al. (2019) évoquent par exemple une augmentation de la motivation des élèves, un engagement plus important dans la tâche ou encore l'atteinte de plus hauts objectifs cognitifs. Ces auteur·rices soulignent également que les jeunes enseignant·es se sentent particulièrement mal préparé·es à l'évaluation et, même s'il semble que l'expérience du métier fasse en partie la différence par la suite (Mertler, 2004), la plupart des travaux continue de pointer des lacunes, même lorsqu'elles-ils sont plus expérimenté·es (DeLuca et al., 2016). D'autres auteur·rices (Tardif, 2012 ; Monchatre, 2008, cités par Tessaro, 2015) les rejoignent en soulignant les manquements quant à la codification et à l'opérationnalisation de l'évaluation chez les

¹ En Belgique, l'enseignement est communautarisé en fonction des trois communautés linguistiques (flamande, francophone et germanophone). La FW-B est le nom officiel de la partie francophone.



enseignant·es du secondaire (au niveau de la formulation des critères, de la pondération, du seuil de réussite, etc.). Dès lors, si ces dernier·ères peinent à mettre en œuvre leurs évaluations sommatives internes, il est légitime de s'intéresser aux conséquences qu'elles engendrent et notamment à la note que l'élève reçoit. Plus précisément, ces constats nous conduisent à interroger les conditions nécessaires pour que la note soit en mesure de fournir des informations pertinentes aux enseignant·es d'une part et aux élèves d'autre part à propos des apprentissages réalisés, afin de favoriser un processus (auto)régulateur.

C'est donc aux pratiques évaluatives notées, encore trop rarement prises en compte dans les recherches (Hadji, 2017 ; Mottier Lopez & Pasquini, 2017), que cette étude va s'intéresser en tentant de documenter à quelles conditions les notes peuvent être réellement informatives (Pasquini, 2019, 2021). En investiguant les modalités de conception et de communication des épreuves évaluatives des enseignant·es, un des enjeux majeurs de cette recherche est de voir dans quelle mesure les enseignant·es mettent en œuvre les éléments propices au développement d'une note « à haute valeur informative ». Pour ce faire, ces épreuves se doivent d'articuler les fonctions formative et sommative pour présenter l'acquisition d'apprentissages, via la note (qui peut se présenter sous forme d'un résultat chiffré ou d'une lettre), tout en fournissant des informations pour des régulations possibles (Harlen, 2005 ; 2012). Cette conceptualisation de la note – qui s'inscrit dans le courant de l'*Assessment for Learning* (AfL), l'une des finalités les plus importantes de l'évaluation scolaire aujourd'hui (Allal & Laveault, 2009 ; Assessment Reform Group, 2002 ; Laveault & Allal, 2016) – est à notre connaissance novatrice en francophonie (Pasquini, 2019).

Conçue dans une visée exploratoire, la présente étude tente ainsi de répondre à la question suivante : « quels sont les freins et leviers qui pourraient œuvrer à la construction et à la communication d'une note à haute valeur informative ? ». Pour ce faire, des futur·es enseignant·es ont invité des enseignant·es de l'enseignement secondaire supérieur de la FW-B (s'adressant à des élèves de 15 à 18 ans) à compléter un questionnaire et, sur cette base, à échanger sur les pratiques d'évaluation². L'alignement curriculaire élargi (Pasquini, 2019) est le modèle théorique principal qui a servi de fil conducteur à la conception du questionnaire. Ainsi, selon ce modèle, pour investiguer la valeur potentiellement informative de la note, il convient non seulement de s'intéresser à la note elle-même (à la façon dont elle est construite et à ce qu'elle représente), mais aussi de resituer l'épreuve d'évaluation dans un cadre plus large en interrogeant sa cohérence avec les objectifs poursuivis et les méthodes déployées. Concrètement, on s'intéressera alors à ce que l'enseignant·e met en œuvre en amont (pour construire son épreuve et y préparer les élèves), mais aussi en aval de l'évaluation (pour leur communiquer leurs résultats notamment). En vue de répondre à notre question de recherche, l'analyse du questionnaire cherchera donc à mettre en lumière ce qui, parmi les conceptions et les pratiques déclarées des enseignant·es interrogé·es, semble pouvoir jouer le rôle de frein ou de levier au caractère potentiellement informatif de la note.

² L'implication des étudiant·es dans le processus de recherche, les échanges qu'elles-ils ont eus avec les enseignant·es autour de leurs pratiques d'évaluation et l'exploitation, lors d'une séance de cours, des informations récoltées visaient des objectifs formatifs. Bien qu'il puisse être intéressant d'analyser cette démarche, ce n'est pas l'objet de cet article qui s'intéresse aux résultats issus des questionnaires.



2. Cadre théorique

Ce cadrage théorique se centre sur la dimension sommative de l'évaluation et plus particulièrement sur la note qu'elle engendre. Après avoir rappelé l'omniprésence de la note et pointé quelques questionnements quant à cette dernière, le caractère potentiellement formatif de l'évaluation sommative sera évoqué. Pour que la note scolaire - qui vise à résumer les apprentissages des élèves grâce à un chiffre (i.e., 12/20) ou une lettre (i.e., S pour satisfaisant) - puisse avoir un réel pouvoir informatif, plusieurs éléments doivent être pris en compte, en amont et en aval de l'évaluation. Le modèle de l'alignement curriculaire élargi (Pasquini, 2019) apportera en ce sens un éclairage pertinent. Enfin, centré sur la communication de la note et les commentaires éventuels qui l'accompagnent, le dernier point de ce cadrage théorique correspond à l'étape où s'arrêtent les éléments investigués dans la présente étude. En effet, les régulations qui devraient pourtant constituer les étapes ultérieures d'une évaluation ne font pas l'objet de cette recherche.

2.1 L'omniprésence de la note

Les études internationales telles que PISA et TALIS révèlent la place prépondérante des évaluations sommatives internes, élaborées par l'enseignant-e, passées en classe ou à domicile, sous une forme traditionnelle « papier-crayon » (OECD, 2012 ; Quittre et al., 2019). En FW-B, l'enquête TALIS 2018 indique ainsi que 90% des enseignant-es déclarent évaluer leurs élèves en administrant un test élaboré par leurs soins (Quittre et al., 2019). Malgré un discours positif sur l'évaluation formative, c'est l'évaluation sommative qui reste la plus largement pratiquée dans de nombreux systèmes scolaires (Issaieva et al., 2015).

Parallèlement, on note une progression du cadrage légal entourant l'évaluation sommative dans nombre de systèmes scolaires. Citons par exemple la volonté d'instaurer des critères plus précis, la mise en œuvre de conditions adéquates pour construire ces évaluations, mais aussi la manière d'élaborer la note (CNESCO, 2014 ; Ministre de l'Enseignement obligatoire et de Promotion sociale, 2014). À ce sujet, en FW-B, l'Arrêté Ministériel du 01/07/2014 précise que :

Tout ce qui fait l'objet d'une évaluation sommative doit correspondre strictement aux objectifs annoncés ... L'essentiel dans la note n'est pas qu'elle soit exprimée à l'aide de chiffres ou de lettres mais qu'elle rende compte avec clarté de la performance accomplie par rapport à celle qui était attendue (p. 5).

Les notes sont donc utilisées et soumises à des directives dans les établissements scolaires de beaucoup de pays, même si elles prennent souvent des formes diverses. Outre ces considérations au niveau de l'encadrement de l'évaluation sommative, soulignons que les notes peuvent également être accompagnées de commentaires. D'ailleurs, l'enquête TALIS 2018 révèle qu'en FW-B, 74% des enseignant-es déclarent ajouter des commentaires écrits à la note des élèves (Quittre et al., 2019). Reconnaissons toutefois que, majoritairement, ceux-ci n'ont pas de valeur rétroactive et se limitent souvent à enjoindre aux élèves de travailler davantage (De Ketele & Frères, 2009). Et, même quand les commentaires ont une visée formative (Calone & Lafontaine, 2018), ils risquent d'être peu pris en compte par les élèves qui préfèrent comparer leurs notes à celles de leur voisin-e dans une visée normative (Merle, 2014 ; Younès & Gaime, 2012).



La note est donc omniprésente, mais ses modes de structuration questionnent, notamment parce qu'ils sont insuffisamment référés aux apprentissages (Pasquini, 2019). Pourtant, il s'agit là d'un élément essentiel pour qu'une évaluation sommative puisse se montrer au service des apprentissages (Allal, 2010 ; Broadbent et al., 2018 ; Harlen, 2005, 2012). Les possibles complémentarités entre les fonctions sommatives et formatives de l'évaluation sont esquissées dans le point suivant, avant de recentrer plus spécifiquement le propos sur la note et son processus de construction.

2.2 La combinaison des fonctions sommative et formative

Suite aux travaux de Bloom au milieu du siècle dernier, la distinction entre évaluation sommative et formative s'est largement répandue auprès des enseignant·es, mais ces dernier·ères les distinguent souvent de façon un peu caricaturale. À titre d'exemple, et alors qu'elle possède d'autres vertus, l'évaluation sommative est presque uniquement associée à la note (Butera et al., 2011 ; De Vecchi, 2014 ; Lipnevich & Smith, 2009, cités par Calone & Lafontaine, 2018) tandis que la version formative est trop souvent caractérisée par le fait qu'elle ne « compte pas pour des points » (Fagnant & Goffin, 2017).

Pour prendre le contrepied de ces croyances, Broadbent et ses collaborateurs (2018) se sont intéressés à la manière de les combiner afin de tirer les bénéfices et avantages de chacune d'elles ; l'évaluation sommative s'enrichirait alors des apports de la formative. Harlen (2005, 2012) et Allal (2010) abondent dans ce sens lorsque toutes deux affirment que l'harmonisation des volets sommatif et formatif servirait l'élève dans la compréhension de ses résultats. De plus, elle permettrait aussi de dégager des pistes de régulation pour l'élève.

2.3 La construction de la note

En s'appuyant sur le modèle classique de l'alignement curriculaire (Anderson, 2002)³, Pasquini (2019) réalise un saut conceptuel en proposant un modèle « élargi » qui semble particulièrement porteur lorsque l'on s'intéresse aux pratiques d'évaluation sommative. En effet, tout en s'appuyant sur les trois piliers fondamentaux sur lesquels tout dispositif de formation doit reposer (objectifs/O ; méthodes/M ; évaluation/É), l'auteur déploie quatre composantes permettant d'appréhender finement la problématique de la notation. Ainsi, s'intéresser à la note scolaire nécessite *de facto* d'appréhender plus en profondeur ces différentes facettes, illustrées dans la figure 1 (cadres rouges) :

- la *référenciation restreinte* (Vial, 2012, cité par Pasquini, 2021) renvoie aux objectifs d'apprentissage à évaluer que les enseignant·es considèrent comme pertinents ;
- le *design* (Wiggins & McTighe, 2005, cités par Pasquini, 2021) est la manière dont les enseignant·es construisent leurs questions, leurs activités, leurs tâches évaluatives concernant les contenus et les compétences visées ;
- la *pondération* représente le poids attribué aux apprentissages. Des outils fréquemment utilisés sont les critères, les points, etc. ;

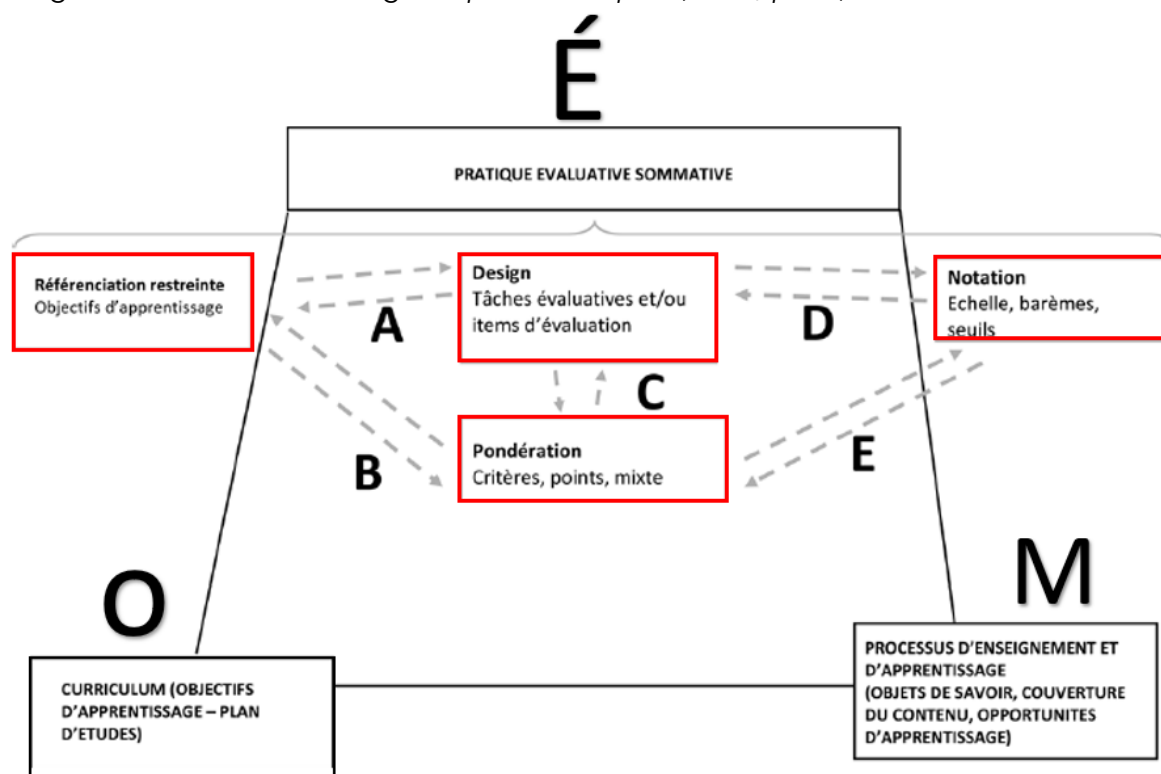
³ Ce modèle d'alignement curriculaire est à rapprocher du modèle de « triple concordance » (objectifs/méthode/évaluation) introduit par Tyler à la fin des années 1940 (Leclercq, 2008). Ces sont les termes habituels de ces trois piliers que nous retiendrons ici.



- la *notation* concerne la manière dont les enseignant·es construisent la note. Les outils sont les échelles, les barèmes (Brookhart, 2017 ; Marzano & Heflebower, 2011, cités par Pasquini, 2021).

Figure 1

L'alignement curriculaire élargi adapté de Pasquini (2019, p. 63)



À la manière du modèle initial présentant l'alignement curriculaire (O-M-É), les facettes propres aux pratiques évaluatives sont en cohérence entre elles. Dans la figure 1, les doubles flèches en pointillés modélisent d'ailleurs les liens réciproques les unissant (relations A-B-C-D et E). Le degré d'information que la note véhicule découle de la cohérence (ou de l'incohérence) de ces relations, mais aussi de la façon dont l'épreuve s'insère dans le modèle plus large, prenant notamment en compte les opportunités d'apprentissages offertes aux élèves avant l'évaluation (cf. le M de « Méthode »). On peut également considérer que « communiquer clairement les objectifs d'apprentissage et d'évaluation aux élèves » fait partie de la « méthode », tout comme la façon de « communiquer les résultats de l'évaluation ». En proposant cette lecture, on considèrera donc que le modèle de l'alignement curriculaire élargi permet d'interroger la valeur informative de la note en s'intéressant à la façon dont cette dernière est construite, mais aussi à ce qui se passe en amont et en aval de l'évaluation.

Revenons à présent sur la note elle-même. Idéalement, les tâches proposées aux apprenant·es doivent effectivement évaluer les objectifs énoncés (A), ceux-ci se retrouvant eux-mêmes dans les critères choisis (B). De même, le poids alloué aux différentes activités d'évaluation doit être en accord avec le niveau de complexité des contenus évalués et des niveaux taxonomiques engagés (C). Ces mêmes activités étant nécessairement présentes dans les grilles de correction ; ceci pour permettre d'explicitier clairement ce qui se cache derrière ces tâches (D). Enfin, pour construire la note, il faut veiller à ce que les critères soient conçus de manière à pouvoir établir des repères significatifs tel le seuil de réussite



(E) (Pasquini, 2021). Dans la construction de la note, la pondération est une étape cruciale. En effet, elle définit le poids/l'importance que l'enseignant·e accorde à une tâche (et donc à un apprentissage) puis lui sert de cadre de référence pour attribuer les points.

À ce sujet, nous relevons particulièrement l'idée de Feldman (2018, cité par Pasquini, 2021) qui définit les points comme « la devise des salles de classe pour évaluer les apprentissages des élèves..., ce qui en fait l'outil de pondération le plus courant » (p. 75). Si l'on s'accorde généralement à reconnaître la difficulté de la notation lorsqu'il s'agit d'évaluer des tâches complexes, le travail de l'enseignant·e peut sembler simple, en apparence, lorsqu'il s'agit de corriger une série d'exercices. On peut en effet penser qu'il suffit de se servir des différents résultats (et donc des points) pour calculer des pourcentages ou des sommes afin de faire le bilan des apprentissages réalisés. Toutefois, une analyse plus fine de la situation met en lumière que ce processus peut se révéler délicat. Nous faisons allusion ici au « phénomène de compensation » (Pasquini, 2021, p. 78). Celui-ci, intimement lié à la pondération, consiste à « rattraper » de potentielles lacunes grâce à l'attribution de points issus de tâches jugées plus abordables par l'enseignant·e. De cette manière, pour des niveaux de difficultés différents (de contenus ou d'habiletés cognitives), il est possible que l'on attribue un poids équivalent (ou peu cohérent par rapport aux objectifs visés) afin de contrebalancer les activités plus complexes.

Ainsi, la construction d'une note scolaire n'est pas une sinécure et dépasse largement la simple action d'additionner des scores intermédiaires ci et là pour aboutir à une note finale représentative des véritables capacités de l'élève. Cette construction est notamment fonction de la (difficile) cohérence que l'utilisateur applique à ses composantes. Mais comme nous l'avons pointé ci-dessus en rappelant l'importance de réinsérer l'évaluation au sein du modèle classique de l'alignement curriculaire (Anderson, 2002), le sens que la note pourra avoir dépendra aussi de ce qui est fait en amont et en aval de l'évaluation.

2.4 La communication de la note

Enseigner, c'est avant tout communiquer avec les autres. Tout comme l'enseignement dispensé en classe, l'épreuve sommative que l'enseignant·e propose à ses élèves constitue un message à destination de celles et ceux-ci (et de leurs parents) et exige donc un feedback afin d'assurer la nécessaire boucle de rétroaction présente dans tout modèle de communication. Ici, l'intention doit impérativement dépasser le simple fait de fournir une information destinée à classer les élèves selon leurs performances scolaires. Crahay et al. (2019) rappellent à ce sujet que l'évaluation normative constitue un terreau particulièrement fertile pour la mise en place du climat de compétition fréquemment rencontré au sein des classes. *A contrario*, lorsqu'on communique la note, l'élève devrait en tirer les informations nécessaires pour, d'une part, se situer par rapport à ses apprentissages et, d'autre part, envisager la régulation de ceux-ci.

Comme développé dans la section précédente, derrière la note scolaire se cachent de multiples facettes que même les professionnel·les de l'éducation peinent à maîtriser. Dès lors, il n'est pas étonnant que les élèves éprouvent beaucoup de difficultés à la comprendre quand elle leur est fournie telle la suite « logique » de l'évaluation effectuée. Pourtant, les enseignant·es ont tendance à réduire l'évaluation sommative à la note (Calone & Lafontaine, 2018), l'ajout de commentaires semblant concerner davantage les épreuves à visée plus formative.



Les commentaires liés à la note pourraient toutefois enrichir le seul feedback délivré par la note elle-même. En ce sens, les études mettant en avant les effets des feedbacks indiquent très souvent leurs bienfaits... sous certaines conditions (Crahay, 2007 ; Georges & Pansu, 2011 ; Hattie, 2009, 2013). En effet, à l'instar de plusieurs auteur·rices (Allal, 2007, 2011 ; Mottier Lopez, 2012 ; Perrenoud, 1993), soulignons que, si c'est bien l'enseignant·e qui endosse un rôle d'étayage en fournissant des feedbacks, c'est l'élève qui se régule au final. Pour y parvenir, l'élève doit non seulement comprendre ces feedbacks, mais aussi se les approprier pour pouvoir progresser dans ses apprentissages (Calone & Lafontaine, 2018).

Finalement, si l'ajout de ce type de commentaires accompagnant la note semble potentiellement porteur, encore faut-il que l'élève prenne la peine de les lire, de chercher à les comprendre et que l'enseignant·e lui ait appris à s'en servir. En effet, pour de nombreux élèves, l'obtention de la note et la comparaison de celle-ci avec celle de ses camarades suffit. Etant donné le caractère potentiellement chronophage d'une telle pratique, on peut aussi comprendre que nombre d'enseignant·es finissent par y renoncer. Les grilles critériées pourraient alors offrir une autre option pour communiquer aux élèves les éléments constitutifs de la note reçue.

À ce sujet, Balan et Jönsson (2018, cités par Pasquini, 2021) précisent d'abord que l'évaluation critériée est centrée sur les apprentissages et permet *de facto* d'identifier plus facilement d'éventuelles faiblesses. En effet, l'élève peut, grâce aux indicateurs, constater des signes observables pour percevoir la qualité de sa production. Ensuite, les critères sont étroitement liés aux objectifs d'apprentissage, ce qui favorise la prise de conscience de l'élève à propos de son niveau, indispensable à la transformation de sa pensée. Enfin, les critères sont communiqués de sorte à pouvoir être inscrits dans une réelle démarche (auto)évaluative. En ce sens, l'enseignant·e est amené·e à utiliser ce type de grilles à différents moments de façon à prévoir un temps d'apprentissage à son usage.

Finalement, on pourrait considérer que la note, les commentaires ou la grille critériée l'accompagnant constituent un ensemble de feedbacks visant à informer l'élève de ses acquis d'apprentissage (fonction sommative). Plus ces éléments sont en cohérence avec les principes de l'alignement curriculaire élargi, plus ils ont de chances d'être réellement informatifs aux yeux des enseignant·es et des élèves et de pouvoir ainsi insuffler différentes formes de régulation (fonction formative).

3. Méthodologie

3.1 La question de recherche et les sous-questions qui guident l'analyse

Pour rappel, cette étude est conçue dans une visée exploratoire à partir de pratiques déclarées d'enseignant·es. Elle cherche à répondre à la question suivante : « quels sont les freins et leviers qui pourraient œuvrer à la construction et à la communication d'une note à haute valeur informative ? ».

Comme signalé par Pasquini dans le modèle de l'alignement curriculaire élargi (2019), les pratiques évaluatives sont composées de différentes facettes. Celles-ci sont à étudier



individuellement et en interaction. Dans cette recherche, nous les prenons en considération en explorant plus spécifiquement les sous-questions suivantes qui nous permettront de situer les points essentiels à considérer avant et après le moment de passation des épreuves, et qui guideront nos analyses.

1. Comment les enseignant·es conçoivent-elles·ils leurs évaluations et leurs barèmes de notation ?
2. Comment perçoivent-elles·ils la note elle-même : considèrent-elles·ils que la note reflète le niveau de maîtrise des élèves et que pensent-elles·ils des seuils de réussite/échec ?
3. Quelles informations communiquent-elles·ils aux élèves avant l'évaluation et comment les y préparent-elles·ils ?
4. Quelles informations communiquent-elles·ils aux élèves après l'évaluation ?

L'analyse de ces questions nous permettra de mettre en lumière, de façon indirecte, ce qui, parmi les conceptions et les pratiques déclarées des enseignant·es concernant la conception de l'épreuve et de la note elle-même (Q1 et Q2) mais aussi concernant ce qu'elles et ils mettent en place en amont (Q3) et en aval (Q4) de l'évaluation, semble pouvoir jouer un rôle de frein ou de levier au caractère potentiellement informatif de la note.

3.2 Le contexte et le public concerné

Avant de décrire l'échantillon sur lequel se base cette recherche, il convient de resituer le contexte dans lequel celle-ci s'est déroulée. Au début de l'année académique 2020-2021, des futur·es enseignant·es⁴ ont été amené·es à interroger des enseignant·es sur les pratiques d'évaluation mises en place dans leurs classes du secondaire supérieur. Concrètement, elles et ils leur ont soumis un questionnaire structuré et ont échangé oralement à son propos. Précisons que les données ainsi recueillies ont été anonymisées et partagées lors d'une séance en petits groupes (composés de 25 à 30 étudiant·es) dans le cadre d'un cours de pédagogie. En début de cursus de formation, ces étudiant·es sont alors pour la plupart confronté·es pour la première fois en tant qu' « observateur·rices externes » à un écho des pratiques évaluatives déclarées qu'elles et ils avaient pourtant vécues précédemment en tant qu'élèves. L'objectif poursuivi était que les étudiant·es puissent se faire une première idée de différentes façons de concevoir l'évaluation et d'en dégager des pistes pour leur pratique future. Autrement dit, le but était de leur permettre de construire les prémises de leur identité professionnelle relative à cette facette spécifique du métier⁵.

Dans le cadre de cette étude, le public est composé de 148 enseignant·es ayant donné leur accord pour que les résultats soient également utilisés à des fins de recherche⁶. Les données sont exploitées afin d'investiguer les freins et leviers au pouvoir potentiellement

⁴ Ces étudiant·es, inscrit·es à l'agrégation de l'enseignement secondaire supérieur ou en master à finalité didactique dans une université belge francophone, se destinent à enseigner à des élèves de 15 à 18 ans.

⁵ Comme précisé en introduction, ce n'est toutefois pas sur l'analyse de ce dispositif de formation que porte le présent article, mais sur les résultats des questionnaires ainsi récoltés.

⁶ Le projet avait en effet été soumis au Comité d'éthique de la Faculté et les participant·es (étudiant·es et enseignant·es) ont donné un accord écrit. Les étudiant·es devaient réaliser le travail dans le cadre du cours, mais elles et ils pouvaient préciser ne pas souhaiter que les données récoltées soient utilisées à des fins de recherche. Les enseignant·es qui acceptaient que les informations les concernant soient utilisées à des fins de recherche ont signé un formulaire de consentement éclairé.



informatif que pourrait avoir la note aussi bien pour aider l'enseignant·e à réguler ses pratiques que pour aider l'élève à réguler ses apprentissages. Notons que, si les répondant·es enseignent tou·tes dans le secondaire supérieur, leur profil est hétérogène, notamment au niveau de leur ancienneté de service (25% ont moins de cinq ans d'ancienneté, 22% ont entre 5 et 10 ans d'ancienneté, 53% ont plus de 10 ans d'ancienneté) et des disciplines enseignées (biologie, chimie, éducation à la philosophie et à la citoyenneté, éducation physique, français, géographie, histoire, latin, langues modernes, mathématiques, physique, religion, sciences économiques, sciences sociales...).

3.3 Le questionnaire relatif aux pratiques d'évaluation

Le questionnaire a été conçu par l'équipe de recherche de façon à apporter des éléments de réponse aux quatre questions-guides. Il est structuré en cohérence avec les différentes étapes d'une démarche d'évaluation sommative. En effet, pour déterminer la qualité informative d'une note, nous nous intéressons à ce qui se passe en amont de l'épreuve pour analyser la façon dont l'épreuve et la note sont construites (Q1 et Q2) et pour appréhender certains éléments relatifs à la manière dont les élèves y sont préparés (Q3). Nous observons aussi les modalités avec lesquelles les enseignant·es communiquent les résultats (Q4).

Plus concrètement, la première partie vise à dresser un état des lieux sur les modalités concrètes de conception des épreuves d'évaluation. Elle est constituée de questions ouvertes et de questions fermées ; ces dernières étant composées d'items face auxquels les enseignant·es ont dû se positionner sur une échelle de type Likert à quatre degrés allant de « jamais » à « toujours ». Nous envisageons la façon dont les épreuves sont construites et interrogeons la concordance des facettes propres aux pratiques évaluatives des enseignant·es telles que la référencement restreinte, la pondération et la notation. En référence aux éléments précisés par Pasquini (2019) dans son modèle d'alignement curriculaire élargi, cette concordance fait partie des conditions essentielles pour que la note puisse être informative. Pour appréhender cela, les questions portent ici sur différents aspects : l'utilisation des critères, le poids alloué aux questions et la construction de la note finale. Et puisqu'un des objectifs visés par la note est qu'elle reflète le niveau de maîtrise, la perception des enseignant·es à cet égard est également questionnée, de même que leur perception quant au choix du seuil de maîtrise.

La deuxième partie a pour but d'identifier les informations communiquées aux élèves *avant* l'évaluation. En référence au modèle classique de l'alignement curriculaire (Anderson, 2002), l'hypothèse sous-jacente est que le potentiel informatif de la note dépend de la façon dont les élèves sont préparés à l'évaluation et clairement informés de celle-ci. Cette partie du questionnaire commence par une question ouverte qui concerne les raisons de prévenir (ou pas) les élèves avant une évaluation et, le cas échéant, les informations que les enseignant·es déclarent spontanément communiquer avant l'épreuve. S'ensuivent des questions plus orientées sur les raisons de communiquer (ou non) les objectifs des leçons, les objectifs des évaluations, les critères de correction et leur pondération. La dernière question de ce volet demande aux enseignant·es de décrire brièvement la façon dont les élèves sont préparé·es aux épreuves auxquelles elles et ils seront soumi·ses : types d'exercices, conseils, devoirs...

La troisième partie appréhende les informations communiquées aux élèves *après* l'épreuve. Comme explicité plus haut, l'hypothèse sous-jacente est ici que le pouvoir informatif de la note seule est limité, et qu'elle gagnerait à être accompagnée de



commentaires (Calone & Lafontaine, 2018). Cette partie comprend des questions fermées du même type que celles du premier volet. Les enseignant·es doivent se situer face aux types de feedbacks transmis aux élèves et expliciter ce qu'elles·ils mettent en œuvre après les évaluations.

3.4 Les types d'analyses réalisées

L'analyse des questions ouvertes est inspirée de l'approche inductive proposée par Blais et Martineau (2006). Pour chaque question ouverte, les 148 réponses récoltées ont été analysées afin de dégager des catégories d'éléments mentionnés par les enseignant·es. Chaque réponse a été lue une première fois pour avoir une vue d'ensemble des éléments évoqués. Une fois ce travail effectué, les différents éléments ont été regroupés en catégories significatives. Chaque réponse a ensuite été relue et intégrée dans l'une de celles-ci ou a conduit à en créer une nouvelle. Notons que certaines réponses entrent dans une et une seule catégorie tandis que d'autres ont trait à deux (parfois même trois) catégories. Quant aux questions fermées, elles ont simplement fait l'objet de calculs de pourcentages.

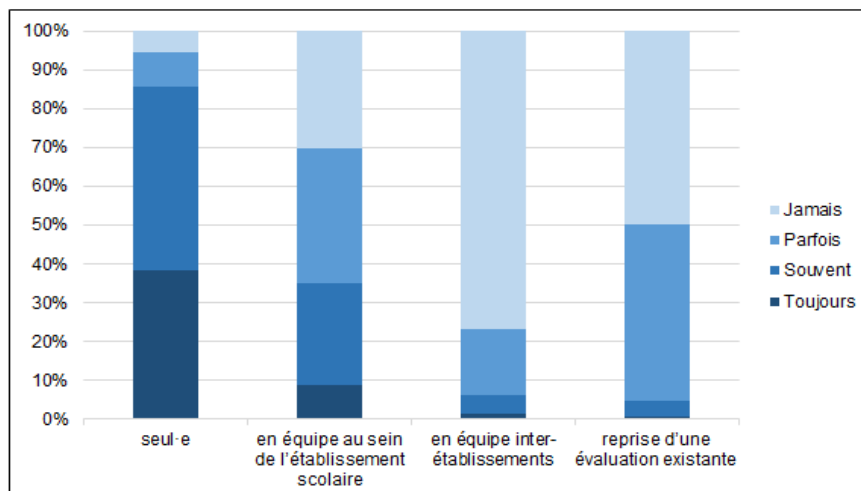
4. Résultats

4.1 La construction de l'évaluation et de la note

Attardons-nous sur la façon dont les enseignant·es construisent les évaluations et plus particulièrement sur l'attention accordée à la pondération. La figure 2 présente le pourcentage d'enseignant·es qui disent à quelle fréquence (« jamais », « parfois », « souvent » ou « toujours ») elles·ils conçoivent leurs évaluations seul·es, en équipe (intra- ou inter-établissements) ou reprennent des évaluations existantes telles que des évaluations externes (certificatives ou non) – autrement dit, elles·ils ne les créent pas.

Figure 2

Fréquence à laquelle les enseignant·es construisent les évaluations, selon la modalité de rédaction



En général, les enseignant·es (86%) mentionnent construire souvent, voire toujours, seul·es leurs évaluations. Les raisons évoquées relèvent tantôt d'un aspect pédagogique tantôt d'un aspect organisationnel. Le premier peut s'illustrer par la volonté d'adapter des évaluations déjà utilisées dans les classes précédentes ou encore de correspondre à ce qui a été mis en place en classe. Le second peut s'expliquer par le fait que l'enseignant·e est le ou la seul·e à donner ce cours au sein de l'établissement scolaire ou tout simplement par facilité d'organisation (notamment horaire). Il est toutefois intéressant de souligner qu'une part non négligeable d'enseignant·es (35%) déclarent créer souvent (ou toujours) leurs évaluations en équipe intra-établissement dans un souci de cohérence ou dans un but constructif (ouverture à d'autres idées), voire dans une optique de contact.

Lorsque les enseignant·es conçoivent leurs évaluations, la majorité d'entre elles et eux (85%) signalent accorder une attention particulière au poids de chaque question par rapport au total. Elles et ils attribuent un nombre de points en fonction de son importance, au vu du programme, des compétences et/ou des points de matière travaillés en classe et considérés comme essentiels, voire du type de question. En effet, certain·es mentionnent spontanément accorder plus de poids quand il y a un développement plus élaboré. Autrement dit, plus il y a d'éléments attendus dans la réponse, plus la question vaut de points (en comptant généralement un point par élément correct). De la même façon, la note finale se construit alors en additionnant les points de chaque question ou critère (lorsque les enseignant·es indiquent utiliser une grille d'évaluation critériée). Outre le fait de procéder à une simple addition, un·e enseignant·e sur dix signale qu'un certain poids peut aussi être attribué aux différents types de tâches (par exemple, 25% pour la théorie, 50% pour les exercices, 25% pour les tâches complexes). Le total obtenu aux questions correspondant à chaque type de tâche est ainsi ramené au pourcentage établi. Et le tout est à nouveau additionné pour aboutir à la note finale. Enfin, même si ils/elles sont peu nombreux·ses à le pointer, il est intéressant de souligner que cinq enseignant·es évoquent la notion de seuil en expliquant que certains critères doivent impérativement être atteints pour obtenir une note de réussite à l'évaluation.

4.2 Le sens de la note

4.2.1 La note, reflet du niveau de maîtrise des élèves

À la question de savoir si les enseignant·es considèrent que la note finale (telle qu'elles ou ils l'élaborent) reflète le niveau de maîtrise des élèves, 36% répondent de manière affirmative, 53% abordent la question de façon nuancée et 9% sont clairement réservés. La moitié des optimistes affirment que, si la grille d'évaluation critériée est construite correctement, la note reflète logiquement les acquis des élèves. Un autre quart explique qu'il s'agit de l'objectif poursuivi par la note ! Quant au dernier quart, il précise une condition : celle que l'épreuve porte sur l'évaluation de compétences. Pour les plus nuancés·es, la note obtenue ne rend pas toujours compte de la réelle maîtrise des apprentissages qu'elle est censée refléter et ce, pour trois raisons principales. Premièrement, si quasiment un cinquième de ces enseignant·es évoque également les grilles d'évaluation critériée, celles-ci ou ceux-ci estiment qu'il est toutefois possible de gagner facilement des points grâce à certaines questions qualifiées de simples ou à des critères liés davantage au respect des consignes qu'au fond voire même à la forme (ce qui revient à questionner à nouveau l'élaboration adéquate de la grille). Par conséquent, certain·es élèves peuvent obtenir une note de réussite tout en passant à côté de l'objectif essentiel et/ou à côté d'un certain type de tâches. Si l'on se réfère à l'exemple



susmentionné, il suffit de répondre correctement aux questions de théorie et aux exercices simples pour réussir avec une moyenne de 50%. Ceci n'est pas sans rappeler un effet de compensation, des lacunes dans certaines questions ou certains critères pouvant effectivement être comblées par les acquis dans les autres questions ou critères. Et cet effet se produit sans doute plus couramment lorsque la note finale résulte, comme c'est généralement le cas ici, d'une addition de points. Dans le même ordre d'idées, quatre enseignant·es mettent en évidence qu'une même note ne reflète pas forcément les mêmes difficultés. L'aspect informatif de la note se perd : il s'avère difficile d'arriver à distinguer par la note uniquement ce qui est acquis et ce qui ne l'est pas encore. Deuxièmement, un tiers des enseignant·es nuancé·es rejoignent les optimistes en stipulant aussi que la note peut refléter le niveau de maîtrise à condition que la matière évaluée soit complexe. Elles et ils clarifient ensuite le fait qu'il ne s'agit d'évaluer qu'une partie de la matière, à un moment donné et vont jusqu'à interroger le degré de rétention à long terme. Troisièmement, près de la moitié de ces enseignant·es énoncent de nombreux facteurs, extérieurs à l'évaluation proprement dite, pouvant influencer sur la note (sur ou sous-évaluation) : la triche, la chance... ou au contraire la distraction, le stress, le manque de motivation... peuvent donc invalider la note. Enfin, il est à noter que les raisons avancées par les pessimistes sont identiques à celles décrites ci-avant, leur avis est seulement plus tranché.

4.2.2 Le seuil de maîtrise

Les enseignant·es ont également été interrogé·es à propos du seuil de réussite qu'elles-ils fixeraient si le choix leur était donné. En effet, en FW-B, lorsque les écoles ont recours aux pourcentages, le seuil de réussite est fixé à 50%. Les enseignant·es qui maintiendraient ce seuil sont 32% : elles et ils le trouvent clair et juste, expriment une habitude qui convient tant aux élèves qu'à elles-eux-mêmes et ne voient pas d'intérêt à changer ce seuil. À ce sujet, elles et ils arguent que, étant donné que ce dernier est arbitraire, il y a toujours moyen d'aider les élèves à l'atteindre en modifiant la pondération des questions par exemple. Les justifications des enseignant·es qui augmenteraient le seuil (39%) sont de deux ordres. Le premier a trait aux apprentissages : augmenter le seuil de réussite permettrait de mieux asseoir les prérequis afin d'éviter que les lacunes s'accumulent et d'atteindre un meilleur niveau de maîtrise des apprentissages. Le second fait appel au facteur motivationnel : augmenter le seuil de réussite offrirait l'opportunité d'augmenter la motivation des élèves, de ne pas succomber à la tentation de se limiter à faire le minimum (donc 50%). Les 29% restants ne diminueraient pas le seuil de réussite mais, en quelque sorte, l'abandonneraient. Si certain·es se contentent de dénoncer l'inefficacité des points, d'autres émettent des alternatives à la note : elles-ils évalueraient soit en pointant les apprentissages acquis, en cours d'acquisition et non acquis, soit en prenant en compte les progrès, soit en privilégiant les auto-évaluations... D'autres encore adopteraient un seuil variable en fonction de la matière (seuil élevé s'il s'agit d'un nouvel apprentissage nécessaire aux apprentissages ultérieurs, par exemple).

4.3 La communication avant l'évaluation

Dans la plupart des cas, les enseignant·es (91%) signalent prévenir les élèves de l'approche d'une évaluation, principalement via le journal de classe ou oralement. Quand les élèves ne sont pas prévenu·es systématiquement, les enseignant·es précisent que ce n'est pas utile.



Plus de 80% des enseignant·es interrogé·es signalent toujours communiquer les objectifs d'apprentissage en début de séquence. *A contrario*, elles-ils sont 14% à ne pas les communiquer systématiquement et précisent dans ce cas que cela dépend du type de séquence. Lorsqu'elles-ils transmettent et explicitent les objectifs, près de la moitié des enseignant·es précisent le faire à des fins de clarification et de meilleure compréhension des objectifs finaux. Certain·es précisent que ce faisant, les élèves peuvent être mieux préparé·es ; cela les aide à donner du sens à la séquence et à avoir une vue d'ensemble de celle-ci. Il est d'ailleurs parfois précisé que le fait de communiquer les objectifs permet de montrer l'importance des apprentissages effectués ou encore de préparer les apprentissages suivants. Quelques enseignant·es évoquent l'importance de communiquer les objectifs de la séquence pour favoriser le travail des élèves en évitant le décrochage, permettant de s'améliorer ou encore préparant l'évaluation en fonction de ce qui est attendu ou encore pour motiver les élèves. Enfin, toujours parmi celles et ceux qui disent communiquer les objectifs, 13% déclarent ne pas être convaincu·es de l'utilité de cette pratique pour les élèves.

On retrouve des proportions similaires quand il s'agit de donner aux élèves les objectifs spécifiques des évaluations : 74% indiquent toujours le faire et 13% disent les communiquer à certaines conditions. Les principales raisons évoquées sont le fait de favoriser la préparation au travail demandé ou la motivation des élèves, l'autonomie, l'intérêt porté à l'étude ou encore le fait de les rassurer une fois les objectifs connus. Notons toutefois qu'un peu plus de 10% des enseignant·es consulté·es déclarent ne pas rappeler les objectifs visés lors des évaluations. Elles-ils sont plusieurs à évoquer que ce sont les mêmes objectifs que ceux donnés en début de séquence.

Dans la même logique, 85% des répondant·es signalent communiquer les critères de correction ; 60% signalent toujours le faire et 29% les communiquer « sous conditions ». Parmi les enseignant·es qui communiquent les critères, près d'un tiers font le lien avec la finalité du travail ; tantôt en définissant les attendus et fixant l'attention sur des points spécifiques, tantôt en précisant le but de ce qui est appris. L'idée de favoriser l'autoévaluation ou l'amélioration du travail, que ce soit pendant les apprentissages et/ou le contrôle, est précisée par 15% des enseignant·es. Au travers des différentes raisons évoquées, on peut constater une volonté de guidage des élèves. Seul·es trois enseignant·es justifient le fait de ne pas donner les critères de correction parce que d'après eux/elles, cela peut influencer négativement les résultats. Quatre autres disent ne pas trouver cela utile. Enfin, certain·es des enseignant·es ne communiquant pas systématiquement les critères indiquent que cela dépend du type de tâche. D'autres indiquent que cela serait trop long, trop détaillé, ou encore que seuls les critères « de base » sont partagés et pas ceux plus spécifiques.

Comme signalé précédemment, la pondération est essentielle à prendre en compte pour permettre une lecture éclairée de la note. La plupart des répondant·es signalent d'ailleurs la communiquer à leurs élèves : 77% signalent toujours le faire et 8% selon les besoins. Ils/elles sont 15% à dire ne pas la partager. Parmi ceux/celles qui communiquent la pondération, ils/elles sont 37% à évoquer un besoin de clarté et/ou de transparence obtenu par ce biais. Comme indiqué précédemment, on retrouve cette idée de clarté et la possibilité donnée aux élèves de poser des choix parmi les items, en prenant en compte ce qui est important pour l'enseignant·e par exemple. Une bonne partie d'entre eux/elles ajoutent que cela permet à l'élève de mesurer l'importance de chaque item. Quatre autres répondant·es signalent cependant que derrière cela, il y a encore une connotation négative car il y a un risque de bâcler certaines questions en faisant des choix. 12% des répondant·es



signalent que la pondération permet une certaine régulation dans le travail. Elles-ils évoquent l'auto-évaluation, l'indication sur l'attendu ou encore la quantité de « points » à noter. Elles-ils sont 13% à signaler que la pondération peut également avoir un effet sur la gestion du temps de travail ou l'organisation face aux questions à traiter. 5% des enseignant-es évoquent à nouveau la vue d'ensemble indiquée par la pondération.

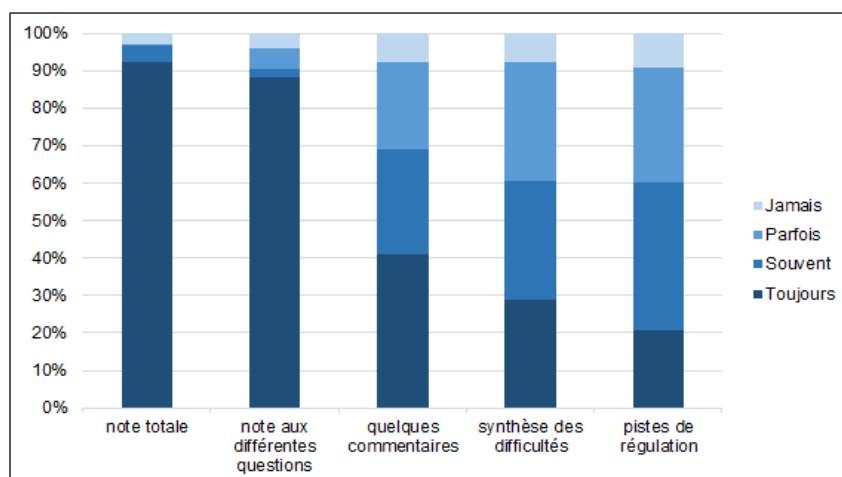
La préparation aux évaluations semble aussi trouver sa place dans les pratiques de classe. Dans le cadre de cette investigation, une question ouverte invite les enseignant-es à décrire brièvement la façon dont les élèves sont préparés aux évaluations. Pour deux tiers des répondant-es, le fait de proposer des exercices d'entraînement (formatifs) en classe permet aux élèves de s'y préparer. Il est souvent fait mention des interactions avec le groupe-classe. On cite, par ordre d'importance la correction de travaux (avec l'enseignant-e mais pouvant également être faite entre pairs), les feedbacks et les moments d'échanges entre élèves et enseignant-es pour des phases de questions-réponses. Si la plupart des feedbacks semblent porter sur les tâches elles-mêmes (ou ne sont pas précisés par les répondant-es), des feedbacks portant spécifiquement sur les méthodes de travail à mettre en œuvre ont été cités par huit enseignant-es différent-es : faire inventer des questions qui pourraient être posées, étudier la théorie, mettre le focus sur les méthodes de travail, favoriser l'auto-évaluation et relire ses notes après le cours. Notons aussi que 32% des enseignant-es parlent de besoin de clarification pour favoriser une bonne préparation : 22% signalent préciser les points importants, plus essentiels ; certains évoquent même les « tuyaux ». De plus, 10% rappellent ici le partage des critères d'évaluation et 5% précisent donner des objectifs très clairs évitant toute interprétation. Elles-ils sont 7% à évoquer la remédiation ou l'individualisation. Rappelons que les questions portant sur la préparation à l'évaluation étaient ouvertes et non orientées, on peut bien entendu supposer qu'ils/elles sont probablement plus nombreux-ses à mettre en œuvre certains éléments au sein de leurs pratiques de classe.

4.4 La communication après l'évaluation

La figure 3 présente le pourcentage d'enseignant-es déclarant à quelle fréquence (« jamais », « parfois », « souvent » ou « toujours ») ils/elles transmettent différents types de feedbacks à tous les élèves suite aux évaluations.

Figure 3

Fréquence à laquelle les enseignant-es transmettent des feedbacks, selon leur type



Sans grande surprise, la quasi-totalité des enseignant-es communique directement la note finale aux élèves (97%) et leur montre le détail des différentes notes aux diverses questions lors de la remise des copies (90%). Au niveau des commentaires, 69% affirment en noter régulièrement quelques-uns, ce chiffre descendant à 61% lorsqu'il s'agit de cibler davantage les difficultés. Ils/elles sont également majoritaires (60%) à prétendre ajouter régulièrement (toujours ou souvent) quelques conseils pour aider l'élève à remédier à ses difficultés.

Lorsque nous les interrogeons pour savoir si les informations qu'elles-ils communiquent aux élèves après l'évaluation permettent à celles et ceux-ci de bien se situer par rapport à leurs apprentissages, une courte majorité répond de manière affirmative. Elles-ils évoquent principalement que ce sont les feedbacks qui jouent un rôle important (27%), que l'étape de correction collective est nécessaire (14%) ou encore soulignent l'importance de la remédiation (7%). Les enseignant-es qui émettent plus de réserves sont plus tranché-es sur les raisons évoquées : 33% disent que de toute façon, l'élève ne s'intéresse qu'à la note, 30% pensent que les élèves ne comprennent pas les informations plus qualitatives communiquées et 21% arguent que le climat de compétition lié à la note qu'elles-ils reçoivent est trop néfaste pour permettre aux élèves de prendre le recul nécessaire pour bien se situer par rapport à leurs propres apprentissages.

5. Conclusion et discussion

Dans le domaine de l'éducation, l'évaluation ne peut être basée sur des jugements de valeur (Nadeau, 1978) mais doit au contraire s'appuyer sur des normes et données objectives. Si la vision actuelle, faisant notamment appel à la notion de « jugement évaluatif » (Mottier Lopez & Allal, 2008), est plus ouverte et accepte une part de subjectivité voire « d'arrangements évaluatifs » (Mottier Lopez et al., 2012), elle demande tout autant voire davantage de compétences professionnelles pour aboutir à une notation cohérente. Et ceci est important puisque les notes, certificats, diplômes, etc. que nous obtenons au terme d'une formation peuvent conduire à des conséquences non négligeables sur nos orientations futures.

Effectivement, le parcours scolaire d'un-e élève est fonction des performances obtenues à ses évaluations et trop souvent, celles-ci ne servent pas suffisamment ses apprentissages. Pour remédier à ce problème, nous rejoignons le point de vue des nombreux-ses auteur-es qui proposent d'associer certaines vertus formatives au fort potentiel de l'évaluation sommative (Allal, 2010 ; Broadbent et al., 2018 ; Harlen, 2005 ; Mottier Lopez, 2015 ; Mottier Lopez & Pasquini, 2017 ; Pasquini, 2019, 2021). Rendre la note davantage informative apparaît dès lors comme une piste d'action réaliste dans ce processus complexe.

Partant des pratiques déclarées des enseignant-es, notre recherche souhaitait apporter un éclairage à cette problématique, en resituant la note au sein du continuum allant de l'amont à l'aval de l'évaluation. L'analyse, guidée par les quatre questions-guides, a ainsi permis de mettre en évidence un certain nombre d'éléments que l'on peut considérer comme jouant potentiellement le rôle de frein ou de levier pour conférer à la note une potentielle valeur informative.



En ce qui concerne la façon dont les enseignant·es conçoivent leurs évaluations et leurs barèmes de notation (Q1), on peut relever un premier élément qui pourrait jouer à la fois un rôle de levier ou, *a contrario*, de frein. Ainsi, lorsque les enseignant·es déclarent construire leurs évaluations en correspondance à ce qui a été mis en place, on peut y voir une volonté d'assurer un alignement curriculaire (Anderson, 2002). La note pourrait alors être informative au sens où elle reflèterait les performances des élèves par rapport à ce qu'ils/elles ont réellement appris. Cependant, si l'enseignant·e crée seul·e son épreuve, en restant trop centré·e sur le contexte de sa classe, elle·il risque d'être soumis·e à l'effet Posthumus et de produire une épreuve discriminante (Crahay et al., 2019). La note obtenue par l'élève reflèterait alors son niveau par rapport à celui des autres élèves de la classe plutôt que sa réelle maîtrise des attendus des référentiels. En ce sens, on peut considérer qu'elle serait peu informative pour lui permettre de se réguler. Si l'on s'intéresse plus spécifiquement à la construction de la note elle-même, en référence à l'élargissement du concept d'alignement curriculaire proposé par Pasquini (2009), on peut pointer deux leviers potentiels : d'une part le souhait de prendre en compte le poids des différentes parties de l'évaluation dans le résultat général de l'épreuve pour qu'il y ait un « équilibre » entre les aspects travaillés ; d'autre part, l'idée de proposer des « critères incontournables » qui doivent être remplis pour pouvoir réussir l'épreuve. En effet, en activant ces leviers, la note devrait davantage refléter le degré de maîtrise des élèves.

Pourtant, la question portant sur la façon dont les enseignant·es perçoivent la note (Q2) a montré qu'une proportion non négligeable (plus de 60% si on additionne les très réservé·es - 9% et les nuancé·es - 53%) considèrerait qu'elle ne reflétait pas nécessairement le niveau de maîtrise des élèves, ce qui constitue en soi un frein important à sa potentielle valeur informative. De nombreux éléments sont mentionnés pour exprimer ce doute, certains relevant de l'élève (chance, triche, stress...) et d'autres pouvant être imputés à la façon dont l'enseignant·e construit lui-même sa note. Plus précisément, il semblerait que ce qui freine le caractère informatif de la note provienne essentiellement de l'effet de compensation, mécanisme grâce auquel l'épreuve peut être réussie alors que l'élève obtient une série de points acquis à certaines questions mais ne témoignant pourtant pas de la maîtrise de la compétence. Pour dépasser cette difficulté, et comme mentionné précédemment, il pourrait être intéressant de creuser l'idée des critères « incontournables » qui, s'ils ne sont pas réussis, ne permettraient pas à l'élève de dépasser le seuil de réussite. À ce sujet, les critiques exprimées quant à ce seuil attestent également un frein important, au sens où la note perd de son caractère informatif quand elle atteste une réussite alors que l'enseignant·e considère pourtant que l'élève ne maîtrise pas la matière. À l'inverse, la proposition consistant à augmenter ce seuil, pour qu'il témoigne davantage d'une maîtrise que d'une réussite « sur le fil », peut être considérée comme un levier pouvant tendre à conférer à la note un pouvoir informatif plus important.

En référence au modèle classique de l'alignement curriculaire (Anderson, 2002), nous avons également appuyé nos travaux sur l'hypothèse selon laquelle le potentiel informatif de la note dépendait de la façon dont les élèves étaient préparés à l'évaluation et clairement informés de celle-ci. Dès lors, les informations communiquées aux élèves avant l'évaluation et la façon de les y préparer (Q3) constituent à nos yeux autant de leviers potentiels. Ainsi, près de 80% des enseignant·es déclarent communiquer les objectifs, les critères et même la pondération, ce qui est essentiel pour aider à clarifier les attentes tant au niveau de l'apprentissage que de l'évaluation qui suivra. Dans le même ordre d'idées, les nombreux éléments que les enseignant·es mettent en place pour préparer les élèves aux évaluations sont essentiels pour aligner les méthodes d'enseignement/apprentissage aux objectifs visés ainsi qu'à l'évaluation qui doit permettre d'en attester la maîtrise. Un frein



évident à la valeur informative de la note serait que celle-ci soit le résultat d'une évaluation à laquelle les élèves n'ont pas été préparés. Toutefois, peu d'éléments en ce sens transparaissent des propos des enseignant·es. Par ailleurs, un frein potentiel à la communication des objectifs, des critères et de la pondération doit toutefois être mentionné : le caractère chronophage de cette activité, pointé par une partie non négligeable d'enseignant·es.

Comme plusieurs écrits permettent d'interroger le caractère informatif de la note elle-même (e.g., Butera et al., 2011 ; Merle, 2014 ; Younès & Gaime, 2012), nous avons aussi investigué les informations communiquées aux élèves après l'évaluation. En termes de leviers, on peut noter la communication du détail de la note (puisque ceci peut aider les élèves à cerner les questions ou les critères maîtrisés ou non) et les commentaires l'accompagnant, surtout lorsqu'ils pointent les difficultés ou qu'ils offrent des pistes de remédiation. Près de 70% des enseignant·es déclarent accompagner fréquemment les notes de commentaires, ce qui est proche de la proportion relevée dans l'enquête TALIS qui portait sur le début du secondaire (élèves de 12-14 ans). Par ailleurs, le fait que plus de 60% des enseignant·es déclarent proposer des commentaires constructifs (ciblés sur les difficultés et/ou donnant des conseils pour aider l'élève à remédier à ses difficultés) est un constat bien plus positif que celui pointé par De Ketele et Frères (2009) qui considéraient que ces commentaires se limitaient souvent à enjoindre aux élèves de travailler davantage. Par ailleurs, les raisons évoquées par les enseignant·es plus réfractaires aux commentaires témoignent de freins importants à la communication d'une note à haute valeur informative. Pointons ici le caractère néfaste d'un climat compétitif, qui semblerait détourner les élèves d'une analyse de leurs difficultés, d'une part, et la difficulté que sembleraient rencontrer certain·es d'entre eux/elles à comprendre les commentaires qualitatifs associés aux notes d'autre part. Concernant le premier frein, on pourrait imaginer que la mise en œuvre d'une pratique d'évaluation soutien d'apprentissage (*Assessment for learning*) au travers d'évaluations formatives, mais aussi en œuvrant à donner un caractère informatif aux évaluations sommatives (Allal & Laveault, 2009 ; Laveault & Allal, 2016) pourrait contribuer à améliorer le climat de classe (Sarrazin et al., 2006). Concernant le second frein, on peut penser qu'il serait important que l'enseignant·e apprenne aux élèves à décoder ce type de commentaires et qu'il/elle les aide à s'en saisir pour réguler leurs apprentissages. En effet, si dispenser des feedbacks précis et qualitatifs semble être une condition favorable pour rendre la note plus constructive (Broadbent et al., 2018), s'assurer qu'ils soient compris et utilisés semble tout aussi essentiel pour qu'ils puissent réellement remplir leur rôle informatif (William, 2011).

Finalement, bien que cette étude exploratoire comporte certaines limites ou biais, tels que la non-représentativité de l'échantillon et le risque de désirabilité sociale, l'enquête réalisée a permis de mettre en lumière un certain nombre de leviers et de freins, tant au niveau des conceptions que des pratiques déclarées des enseignant·es, pour donner à la note et à sa communication un caractère potentiellement informatif. Evidemment, des études complémentaires seraient nécessaires afin de mieux comprendre comment agir sur ces éléments pour améliorer les compétences évaluatives des enseignant·es, en vue d'agir positivement sur les apprentissages des élèves (DeLuca et al. 2019 ; Kruse et al., 2020).

Références

Allal, L. (2007). Régulations des apprentissages : orientations conceptuelles pour la recherche et la pratique en éducation. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en*



- situation scolaire et en formation* (pp. 7-23). De Boeck. <https://doi.org/10.3917/dbu.motti.2007.01.0007>
- Allal, L. (2010). Assessment and the Regulation of Learning. In: P. Peterson, E. Baker & B. McGaw (Eds.), *International Encyclopedia of Education. Volume 3* (pp. 348-352). Elsevier. <http://unige.ch/fapse/people/allal/doc/Allal2010.pdf>
- Allal, L. (2011). Pedagogy, didactics and the co-regulation of learning: a perspective from the French-language world of educational research. *Research Papers in Education*, 26(3), 329-336. <http://dx.doi.org/10.1080/02671522.2011.595542>
- Allal, L., & Laveault, D. (2009). Assessment for Learning: Evaluation-soutien d'Apprentissage. *Mesure et évaluation en éducation*, 32(2), 99-106. <https://doi.org/10.7202/1024956ar>
- Anderson, L. W. (2002). Curricular alignment: A re-examination. *Theory Into Practice*, 41, 255-260. https://doi.org/10.1207/s15430421tip4104_9
- Assessment Reform Group (2002). *Assessment for learning: 10 principles. Research-based principles to guide classroom practice.* https://www.researchgate.net/publication/271849158_Assessment_for_Learning_10_Principles_Research-based_principles_to_guide_classroom_practice_Assessment_for_Learning
- Bélair, L., & Coen, P. (2015). Introduction. In P. Coen & L. Bélair (Eds.), *Évaluation et autoévaluation: Quels espaces de formation?* (pp. 9-12). De Boeck Supérieur. <https://doi.org/10.3917/dbu.ceon.2015.02>
- Blais, M., & 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. <https://www.researchgate.net/publication/242085520>
- Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavour: A case study in a large class. *Assessment & Evaluation In Higher Education*, 43(2), 307-322. <https://doi.org/10.1080/02602938.2017.1343455>
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803-848. <https://doi.org/10.3102/0034654316672069>
- Butera, F., Buchs, C., & Darnon, C. (2011). Introduction. In F. Butera, C. Buchs & C. Darnon (Eds.), *L'évaluation, une menace?* (pp. 9-12). Presses universitaires de France. <https://doi.org/10.3917/puf.darno.2011.01.0009>
- Calone, A., & Lafontaine, D. (2018). Feedback normatif vs feedback élaboré : quel impact sur la performance et le sentiment de contrôlabilité des élèves. *Évaluer. Journal international de recherche en éducation et formation*, 4(2), 47-76. <https://journal.admee.org/index.php/ejiref/article/view/156/86>
- Conseil national d'évaluation du système scolaire (CNESCO). (2014). *L'évaluation des élèves par les enseignants dans la classe et les établissements : réglementation et pratiques. Une comparaison internationale dans les pays de l'OCDE.* CNESCO. <https://hal.archives-ouvertes.fr/hal-01112954>
- Crahay, M. (2007). Feedback de l'enseignant et apprentissage des élèves : revue critique de la littérature de recherche. In L. Allal & L. Mottier Lopez (Eds.), *Régulation des apprentissages en situation scolaire et en formation* (pp. 47-70). De Boeck. <https://doi.org/10.3917/dbu.motti.2007.01.0247>
- Crahay, M., Mottier Lopez, L., & Marcoux, G. (2019). L'évaluation des élèves : Docteur Jekyll and Mister Hyde de l'enseignement. In M. Crahay (Ed.), *Peut-on lutter contre l'échec scolaire?* (pp. 357-424). De Boeck. <https://doi.org/10.3917/dbu.craha.2019.01.0357>



- De Ketele, J.-M., & Frères G. (2009) Les commentaires des enseignants et des élèves : simples jugements ou processus évaluatifs ? *Les Cahiers de Recherche en Education et en Formation*, 74, 1-38. <https://halshs.archives-ouvertes.fr/halshs-00561526>
- DeLuca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251-272. <https://doi.org/10.1007/s11092-015-9233-6>
- DeLuca, C., Coombs, A., MacGregor, S., & Rasooli, A. (2019). Toward a differential and situated view of assessment literacy: Studying teachers' responses to classroom assessment scenarios. *Front. Educ*, 4(94). <https://doi.org/10.3389/feduc.2019.00094>
- Fagnant, A., & Goffin, C. (2017). Les conceptions des futurs enseignants du secondaire en matière d'évaluation : entre un accord de principe et une vision limitée de l'évaluation formative. *Mesure et Evaluation en Education*, 40(1), 1-32. <https://doi.org/10.7202/1041002ar>
- Georges F., & Pansu, P. (2011). Les feedbacks à l'école : un gage de régulation des comportements scolaires. *Revue Française de Pédagogie*, 176, 101-124. <https://doi.org/10.4000/rfp.3239>
- Hadji, C. (2017). Savoir mettre en œuvre une évaluation constructive. In A. Bentolila (Ed.). *L'essentiel de la pédagogie* (pp. 229-250). Nathan.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Hattie, J. (2013). *Visible learning for teachers: Maximizing impact on learning*. Routledge. <https://doi.org/10.4324/9780203181522>
- Harlen, W. (2005). Teacher's summative practices and assessment for learning - tensions and synergies. *The Curriculum Journal*, 16(2), 207-223. <https://doi.org/10.1080/09585170500136093>
- Harlen, W. (2012). On the relationship between assessment or formative and summative purposes. In J. Gardner (Ed.), *Assessment and learning* (pp. 87-102). Sage. <https://doi.org/10.4135/9781446250808.n6>
- Issaieva E., Yerly, G., Petkova, I., Marbaise, C., & 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 ? In P.-F. Coen & L. M. Bélair (Eds.), *Évaluation et autoévaluation. Quels espaces de formation ?* (pp. 73-98). De Boeck Supérieur. <https://doi.org/10.3917/dbu.ceon.2015.02.0073>
- Kruse, L., Impellizeri, W., Witherel, C. E., & Sondergeld, T. A. (2020). Evaluating the Impact of an Assessment Course on Preservice Teachers' Classroom Assessment Literacy and Self-Efficacy. *Mid-Western Educational Researcher*, 32(2). <https://www.mwera.org/MWER/volumes/v32/issue2/V32n2-Kruse-FEATURE-ARTICLE.pdf>
- Laveault, D., & Allal, L. (2016). Implementing Assessment for Learning: Theoretical and Practical Issues. In D. Laveault, & L. Allal (Eds.). *Assessment for Learning: Meeting the Challenge of Implementation* (pp. 1-18). Springer. <https://doi.org/10.1007/978-3-319-39211-0>
- Leclercq, D. (2008). A la recherche de la Triple concordance en Education. Illustration sur un cours de 1^o Bac universitaire en grand groupe. LabSET IFRES. Université de Liège. https://orbi.uliege.be/bitstream/2268/22286/1/A1240_Leclercq_A_la_recherche_de_la_T_C_Illustration_fevrier_2008_b.pdf
- Merle, P. (2014). Faut-il en finir avec les notes ? *La vie des idées*. <https://lavedesidees.fr/Faut-il-en-finir-avec-les-notes.html>
- Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American secondary education*, 49-64. <https://www.jstor.org/stable/41064623>
- Ministre de l'Enseignement obligatoire et de Promotion sociale. (2014, 1^{er} juillet). Arrêté ministériel portant approbation du règlement des études de l'Enseignement secondaire ordinaire



- organisé par la Communauté française. *Moniteur belge*.
https://www.galilex.cfwb.be/fr/leg_res_02.php?ncda=40656&referant=I01
- Mottier Lopez, L. (2012). *La régulation des apprentissages en classe*. De Boeck. <https://archive-ouverte.unige.ch/unige:24531>
- Mottier Lopez, L. (2015). *Evaluation formative et certificative des apprentissages : enjeux pour l'enseignement*. De Boeck. <https://archive-ouverte.unige.ch/unige:106923>
- Mottier Lopez, L., & Allal, L. (2008). Le jugement professionnel en évaluation: un acte cognitif et une pratique sociale située. *Revue Suisse des Sciences de l'Education*, 30(3), 465-482.
<https://doi.org/10.25656/01:4229>
- Mottier Lopez, L., Tessaro, W., Dechamboux, L., & Morales Villabona, F. (2012). La modération sociale : un dispositif soutenant l'émergence de savoirs négociés sur l'évaluation certificative des apprentissages des élèves. *Questions vives recherches en éducation*, 6, 159-175.
<https://doi.org/10.4000/questionsvives.1235>
- Mottier Lopez, L., & Pasquini, R. (2017). Professional controversies between teachers about their summative assessment practices: a tool for building assessment capacity. *Assessment in Education: Principles, Policy & Practice*, 24(2), 228-249.
<https://doi.org/10.1080/0969594X.2017.1293001>
- Nadeau, M.-A. (1978). L'évaluation de l'apprentissage en milieu scolaire : un modèle d'évaluation continue. *Revue des sciences de l'éducation*, 4(2), 205-221.
<https://doi.org/10.7202/9000075ar>
- OECD. (2012). *Grade Expectations: How marks and education policies shape student's ambitions*. PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264187528-en>
- Pasquini, R. (2019). Élargir conceptuellement le modèle de l'alignement curriculaire pour comprendre la cohérence des pratiques évaluatives sommatives notées des enseignants : enjeux et perspectives. *Mesure et évaluation en éducation*, 42(1), 63-92.
<https://doi.org/10.7202/1066598ar>
- Pasquini, R. (2021). *Quand la note devient constructive. Evaluer pour certifier et soutenir les apprentissages*. Presses de l'Université Laval. <http://hdl.handle.net/20.500.12162/4900>
- Perrenoud, P. (1993). Vers des démarches didactiques favorisant une régulation individualisée des apprentissages. In L. Alla, D. Bain, & Ph. Perrenoud (Eds.), *Evaluation formative et didactique du français* (pp. 31-50). Delachaux et Niestlé.
http://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_1993/1993_17.html
- Quittre, V., Dupont, V., & Lafontaine, D. (2019). *Enseigner au quotidien*. Talis 2018.
<https://events.uliege.be/talis-fwb/resultats-talis-2018-en-fwb/>
- Sarrazin, P., Tessier, D., & 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. <https://doi.org/10.4000/rfp.463>
- Tessaro, W. (2015). Formation permanente des enseignants à l'évaluation des apprentissages : une continuité à construire. In P.-F. Coen & L. Bélair (Eds.), *Évaluation et autoévaluation : Quels espaces de formation ?* (pp. 55-71). De Boeck Supérieur.
<https://doi.org/10.3917/dbu.ceon.2015.02.0055>
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, 37(1), 3-14.
<https://doi.org/10.1016/j.stueduc.2011.03.001>
- Younès, N., & Gaime, E. (2012). L'évaluation formative en contexte : points de vue d'enseignants, points de vue d'élèves. *Diversité, VEI*, 161-166. <https://hal.archives-ouvertes.fr/hal-01121431>



Remerciements

Les auteur·rices remercient les enseignant·es qui se sont confié·es sur leurs pratiques évaluatives et les futur·es enseignant·es qui ont accepté de partager les informations récoltées. Cette étude a pu être menée suite à l'impulsion d'un subside de recherche octroyé par la Faculté de Psychologie, Logopédie et Sciences de l'Éducation de l'Université de Liège.

Christelle Goffin

Assistante de formation au centre interfacultaire de formation des enseignant·es à l'Université de Liège, membre principale de l'Unité de recherche DIDACTIFEN et membre secondaire d'EQUALE. Mon domaine de recherche prioritaire touche aux conceptions et pratiques d'évaluation des (futur·es) enseignant·es du secondaire supérieur.

Jean Baron

Assistant de formation au centre interfacultaire de formation des enseignant·es à l'Université de Liège après avoir exercé comme instituteur primaire. Mon domaine de recherche principal concerne tant les connaissances et pratiques d'évaluation des enseignant·es que l'enseignement de celles-ci.

Jean-Pascal Ochelen

Assistant de formation au centre interfacultaire de formation des enseignant·es à l'Université de Liège, maître-assistant en pédagogie et méthodologie à la Haute École pédagogique de la ville de Liège et professeur de sciences dans le secondaire. Mes domaines de prédilection sont la formation des enseignant·es et l'évaluation des apprentissages scolaires.

Annick Fagnant

Professeure en sciences de l'Éducation à l'Université de Liège, Directrice de l'Unité de recherche EQUALE et membre secondaire du DIDACTIFEN. Mes domaines de recherche portent notamment sur les premiers apprentissages mathématiques et sur les pratiques d'évaluation des enseignant·es.

