

Pour citer cet article : Pasquini, R. (2022). Dans le champ de l'évaluation aussi, la nouveauté naît probablement dans le retour aux sources : introduction du numéro 6 de La Revue LEeE. *La Revue LEeE*, 6. <https://revue.leee.online/index.php/info/article/view/152>

DANS LE CHAMP DE L'ÉVALUATION AUSSI, LA NOUVEAUTÉ NAÎT PROBABLEMENT DANS LE RETOUR AUX SOURCES¹

Introduction du numéro 6 de La Revue LEeE

Raphaël PASQUINI

Version de la publication : novembre 2022

Bien que certaines études soulignent la difficulté du corps enseignant à mettre en œuvre des démarches d'évaluation formative (Hayward & Spencer, 2010 ; Kingston & Nash, 2011 ; Klenowski, 2009 ; Moss, 2013 ; Webb & Jones, 2009), la recherche scientifique reconnaît leurs effets positifs sur les apprentissages des élèves (Black & Wiliam, 1998 ; Lui & Andrade, 2022 ; Mottier Lopez, 2015 ; Wiliam, 2010). Dans cet élan initié au cœur des années 1970 et visant à mettre l'évaluation au service des apprentissages, une quinzaine de chercheur-ses se réunissent au Royaume-Uni à la fin des années 1980 : l'*Assessment Reform Group* se constitue avec pour objectif de garantir que la politique et la pratique de l'évaluation tiennent compte de résultats de recherche. L'ouvrage « *Beyond testing* » est publié par l'une de ses membres, Caroline Gipps, en 1994. Il souhaite marquer un changement radical de paradigme entre la vision psychométrique et celle de l'évaluation pédagogique (*educational assesement*). À notre connaissance, c'est la première fois que le terme *Assessment for Learning (AfL)*², spécifiquement associé aux pratiques d'évaluation formatives jusque-là, inclut une réflexion sur les conditions nécessaires pour que les pratiques d'évaluation sommative et les pratiques de notation soient également au service de la progression de tous les élèves. Par-là, ce groupe confirme que l'évaluation réalisée en classe devrait toujours constituer un des outils les plus puissants pour soutenir des apprentissages significatifs et durables. Mais il souhaite également que l'AfL porte une ambition relativement novatrice : mieux comprendre dans quelle mesure la certification des apprentissages permette également au plus grand nombre d'élèves de progresser en favorisant notamment les processus de régulation (Laveault & Allal, 2016). L'acceptation de cette forme d'évaluation est présente depuis les années 2000 dans les prescriptions de la majorité des systèmes éducatifs occidentaux, qui travaillent à sa promotion (OCDE, 2013).

¹ Nous empruntons cette expression à Edgar Morin qui, dans son ouvrage « *Amour, poésie, sagesse* » de 1997, propose de revenir aux questions fondamentales de l'existence pour mieux en comprendre la complexité, arguant que c'est ainsi que la pensée progresse.

² Ce terme est traduit en français par évaluation-soutien d'Apprentissage (é-sA) par Allal et Laveault (2009), mais d'autres chercheur-ses, comme De Ketele ou Mottier Lopez parlent d'évaluation pour apprendre.

La réalité des pratiques semble toutefois résister à cette orientation. Une des raisons tiendrait au fait que l'évaluation sommative notée - dont les enjeux pour les élèves sont élevés puisque c'est souvent sur son unique base que leur parcours scolaire est déterminé - est souvent considérée négativement dans la sphère éducative, car elle est assimilée à une sanction au service de la sélection des élèves (Guskey, 2022 ; Taras, 2005). Plusieurs travaux menés dans des contextes divers ont effectivement souligné que les pratiques évaluatives restent la plupart du temps traditionnelles (Braxmeyer et al., 2004 ; Brookhart et al., 2016 ; Mottier Lopez, 2013 ; Ohlsen, 2007), c'est-à-dire centrées sur l'établissement d'une note et le calcul de moyennes, sans réels retours d'informations constructifs pour les acteur·rices concerné·es. Par ailleurs, nombre de ces pratiques produisent un effet que la recherche a appelé « teaching to the test » (Popham, 2001 ; Stobart, 2008), qui traduit un surentraînement des élèves à l'épreuve sommative réalisée en amont de l'enseignement, réduisant ainsi les opportunités d'apprentissage. Du point de vue des pratiques de notation qui y sont fortement articulées (McMillan & Nash, 2000), le constat est encore plus édifiant : un nombre conséquent d'études documente leur faible valeur informative pour les élèves et les enseignant·es ainsi que leur rapport régulièrement tenu aux apprentissages certifiés (Brookhart et al., 2016 ; Cross & Frary, 1999 ; Pasquini, 2019, 2021), soulignant ici aussi que la note est le plus souvent au service d'une logique sélective et de gestion des flux d'élèves (Butera et al., 2012 ; Crahay et al., 2019 ; Guskey, 2013).

Ces constats partagés dans divers contextes mettent en lumière la nécessité d'approfondir la problématique de l'évaluation-soutien d'apprentissage lorsque cette dernière a pour but premier de certifier. Tout en reconnaissant que les pratiques d'évaluation sommative et de notation sont d'une grande complexité (DeLuca et al., 2019 ; McMillan & Nash, 2000 ; Mottier Lopez, 2015 ; Mottier Lopez & Pasquini, 2017 ; Pasquini, 2019 ; Pasquini & DeLuca, 2021), nous manquons encore de recherches empiriques permettant de mieux comprendre les conditions nécessaires pour que ces pratiques soutiennent les apprentissages des élèves. Corollairement, nous avons aussi besoin de nouvelles conceptualisations des phénomènes évaluatifs spécifiques à ces pratiques, en prenant en compte les apports de la recherche ainsi que les spécificités des contextes et des disciplines. Il s'agit ici de réaffirmer que toute démarche d'évaluation sommative notée doit prioritairement poursuivre trois objectifs : certifier les apprentissages considérés comme importants dans les curricula et les établissements, mais aussi fournir des rétroactions formatives aux élèves et donner des informations aux enseignant·es pour leur permettre de planifier les apprentissages à poursuivre (Carey & Carifio, 2012). Pour appréhender cette acception du concept d'évaluation-soutien d'apprentissage, un retour aux sources de ce dernier est nécessaire : le fait d'affronter la complexité inhérente à l'articulation des fonctions de certification et de régulation est un défi, qui oblige notamment à dépasser le débat opposant les pro et les anti (notes, examens, tests, certificats) qui divise fréquemment le monde de l'éducation depuis plus d'un siècle, et qui n'a notamment pas permis à ce jour aux pratiques de notation de progresser significativement à large échelle (Brookhart et al., 2016).

C'est l'ambition de ce numéro, qui présente les travaux et réflexions de chercheur·ses et d'enseignant·es américain·es, néo-zélandais·es, suédois·es, belges, suisses romand·es et tessinois·es. Leurs textes proposent, dans divers contextes et avec des questions ciblées, d'interroger rigoureusement les habitudes, voire les habitus, souvent conditionnés par des contextes eux-mêmes saisis dans des modes opératoires a priori inflexibles.

Ainsi, huit articles constituent ce numéro. Tout d'abord, dans un article théorique recensant de nombreuses recherches réalisées en contexte anglophone, Susan Brookhart met en évidence le rôle clé que jouent les critères dans la perspective d'articuler évaluation formative et notation, de permettre aux élèves d'auto-réguler leurs apprentissages, d'outiller les différentes formes de rétroactions, d'améliorer la dimension informative de la note ou encore de permettre au corps enseignant de prendre des décisions relativement aux apprentissages des élèves. Puis, dans une recherche empirique menée très récemment en Suède, Anders Jönsson investigate comment des élèves de fin du primaire pourraient utiliser des critères de notation explicites comme moyen



d'autorégulation de leur apprentissage et d'amélioration leurs performances scolaires. Suit un article de Frances Edwards qui montre, dans une recherche empirique menée en Nouvelle-Zélande, comment les croyances d'enseignant-es novices ont évolué suite à leur formation pour mettre l'évaluation sommative au service des apprentissages. Christelle Goffin, Jean Baron, Jean-Pascal Ochelen et Annick Fagnant présentent ensuite les résultats d'une recherche empirique menée en Fédération Wallonie-Bruxelles auprès d'enseignant-es du secondaire où ils-elles tentent d'identifier les freins et leviers qui pourraient œuvrer à la construction et à la communication d'une note à haute valeur informative. Réalisée à Fribourg, la recherche empirique de Gonzague Yerly et Jean-Louis Berger vise quant à elle à discuter des conditions qui permettent à des examens universitaires de soutenir les processus d'apprentissage des étudiant-es. De leur côté, Miriam Salvisberg et Lucie Mottier Lopez interrogent la pluralité des buts qui organisent et orientent l'activité évaluative certificative d'enseignants et d'enseignantes du primaire et du secondaire dans le canton du Tessin. Enfin, deux récits de pratique rendent compte de réflexions également enrichissantes : Yoann Buyck, Bernard Poussin, Adriàn Cordoba, Bénédicte Desbiolles, Philippe Delmege, Yann Fischer, Esteban Iglesias et Benoît Lenzen présentent le processus de construction d'un nouveau canevas pour l'évaluation certificative en éducation physique à l'école primaire de Genève en s'appuyant sur la notion de notation constructive. Quant à Jan Mrazek et Fernando Morales, ils donnent à voir les réflexions d'un enseignant vaudois de mathématiques du secondaire qui tente de rendre ses démarches d'évaluation sommatives constructives en traitant de la résolution de problèmes arithmétiques verbaux avec ses élèves.

À n'en pas douter, ces textes offrent un regard rigoureux et pluriel sur une thématique encore peu explorée dans la communauté scientifique et professionnelle francophone. La diversité des genres de textes met en lumière les enjeux de recherche et de pratique importants pour de futurs travaux. Bonne lecture !

Références

- Black, P. & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education, Principles, Policies & Practice*, 5(1), 7-74. <http://dx.doi.org/10.1080/0969595980050102>
- Braxmeyer, N., Guillaume, J.-C., & Levy, J.-F. (2004). Les pratiques d'évaluation des enseignants au collège. *Les dossiers du Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Direction de l'évaluation et de la prospective*, 160, 5-117. <https://archives-statistiques-depp.education.gouv.fr/Default/doc/SYRACUSE/10844/les-pratiques-d-evaluation-des-enseignants-au-college-nicole-braxmeyer-jean-claude-guillaume-et-jean?lg=fr-FR>
- Brookhart, S., 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). *L'évaluation, une menace ?* PUF.
- Carey, T., & Carifio, J. (2012). The minimum grading controversy: Results of a quantitative study of seven years of grading from an urban high school. *Educational Researcher*, 41(6), 201-208. <https://doi.org/10.3102/0013189X12453309>
- 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. 358-425). De Boeck.
- Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, 12(1), 53-72. https://doi.org/10.1207/s15324818ame1201_4
- 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. *Frontiers in Education*, 4(94). <https://doi.org/10.3389/educ.2019.00094>
- Guskey, T. R. (2013). The case against percentage grades. *Educational, School, and Counseling Psychology Faculty Publications*. 22. https://uknowledge.uky.edu/edp_facpub/22



- Guskey, T. R. (2022). Can grades be an effective form of feedback? *Kappan*, 104(3), 36-41. <https://kappanonline.org/grades-feedback-guskey/>
- Hayward, L., & Spencer, E. (2010). The complexities of change: Formative assessment in Scotland. *Curriculum Journal*, 21(2), 161-177. <https://doi.org/10.1080/09585176.2010.480827>
- Kingston, N., & Nash, B. (2011). Formative assessment: A meta-analysis and a call for research. *Educational Measurement, Issues and Practice*, 30(4), 28-37. <https://doi.org/10.1111/j.1745-3992.2011.00220.x>
- Laveault, D., & Allal, L. (Eds.) (2016). *Assessment for learning: Meeting the challenge of implementation*. Springer International.
- Lui, A. M., & Andrade, H. L. (2022). The next black box of formative assessment: A model of the internal mechanisms of feedback processing. *Frontiers in Education*, 7. <https://doi.org/10.3389/educ.2022.751548>
- McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making*. <https://files.eric.ed.gov/fulltext/ED447195.pdf>
- Moss, C. (2013). Research on classroom summative assessment. In J. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 235-256). SAGE Publications.
- Mottier Lopez, L. (2013). Pour une culture partagée de l'évaluation : du retour à des pratiques d'évaluation traditionnelles vers la construction d'un nouveau rapport entre la recherche, le terrain et le politique. *Revue Education Canada*. <http://www.ceaace.ca/educationcanada/article/pour-une-culturepartagée-de-l'évaluation>
- Mottier Lopez, L. (2015). *Évaluations formative et certificative des apprentissages : enjeux pour l'enseignement*. De Boeck.
- 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>
- Ohlsen, M. T. (2007). Classroom assessment practices of secondary school members of NCTM. *American Secondary Education*, 36(1), 4-14.
- Organisation de Coopération et de Développement Economiques (OCDE) (2013). *Les grandes espérances : comment les notes et les politiques éducatives façonnent-elles les aspirations des élèves ?* Éditions OCDE. [https://www.oecd.org/pisa/pisaproducts/pisainfocus/pisa%20in%20focus%20n°26%20\(fra\)--Final.pdf](https://www.oecd.org/pisa/pisaproducts/pisainfocus/pisa%20in%20focus%20n°26%20(fra)--Final.pdf)
- 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 : évaluer pour certifier et soutenir les apprentissages*. Presses de l'Université Laval.
- Pasquini, R., & Deluca, C. (2021). Grading in a dilemmatic space: An exploratory cross-cultural analysis of Mathematics and Language Secondary teachers. *Comparative and International Education*, 49(2), 51-70. <https://doi.org/10.5206/cieeci.v50i1.14133>
- Popham, W. J. (2001). Teaching to the Test? *Educational Leadership*, 58(6), 16-20.
- Stobart, G. (2008). *Testing times: The uses and abuses of assessment*. Routledge.
- Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466-478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>
- Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 16(2), 165-184. <https://doi.org/10.1080/09695940903075925>
- William, D. (2010). An integrative summary of the research literature and implications for a new theory of formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). Routledge.

