



Mesure et évaluation

au-delà du testing



Résumé

- La mesure et l'évaluation sont des outils essentiels dans le sens qu'ils permettent au personnel scolaire d'adapter ses pratiques pédagogiques, et aux élèves de réfléchir, et de se questionner sur leur propre apprentissage, aux parents la capacité d'accompagner leur enfant dans leur réussite scolaire. Également, elles aident les preneurs de décision dans la mise en œuvre des politiques en matière d'éducation.
- Les compétences en créativité, en apprentissage socio-affectif, en citoyenneté et en santé devraient être évaluées pour les mêmes raisons que l'on évalue la littératie ou la numératie, dans le but de fournir des données pertinentes et spécifiques au sujet de l'apprentissage des élèves dans ces domaines essentiels.
- L'évaluation de ces compétences est complexe, faisant en sorte que l'on ne puisse pas avoir recours aux mêmes outils et stratégies utilisés régulièrement dans l'évaluation des compétences et des connaissances régulières.
- En classe, une série de stratégies peuvent servir à évaluer le cheminement scolaire de l'élève.
- Il est possible d'évaluer ces compétences dans différentes juridictions (locale, provinciale, nationale ou internationale); cependant, l'évaluation standardisée ou les enquêtes fournissent des données limitées compte tenu de la complexité des compétences.

é•va•lu•a•tion (n)

L'évaluation est un processus par lequel sont colligées des données sur les élèves, la salle de classe, l'école ou le système scolaire même par rapport à un ensemble de critères d'apprentissage ou de programmes d'étude (Ontario, 2010).

me•su•re (n)

La mesure est un processus par lequel une expérience d'apprentissage, un phénomène ou un contexte se traduit en variables numériques.

La mesure et l'évaluation sont des concepts inhérents à l'éducation.

On peut s'en servir à des fins diverses, notamment ;

- rendre des comptes aux élèves, aux parents et au public en général ;
- aider les élèves à comprendre les objectifs en matière d'éducation et leur propre cheminement scolaire ;
- aider le personnel scolaire à comprendre efficacement leurs propres pratiques pédagogiques ;
- améliorer la collaboration à l'échelle de l'école ;
- fournir des données qui permettent de constater si le système fonctionne et s'il est équitable pour chacun des élèves.

Complexité de l'évaluation

Partout dans le monde, le personnel scolaire, les preneurs de décision et les experts dans le domaine de l'éducation s'entendent tous pour dire que la réussite, tant à l'école comme dans la vie, implique plus que les habiletés et les compétences en numératie, en littératie et celles définies par les programmes d'étude (p.ex. Christensen et Lane, 2016; OCDE, 2016). Plusieurs systèmes d'éducation tentent d'enclôser des compétences plus élargies, par exemple des compétences du 21^e siècle allant jusqu'aux compétences globales, dans les programmes d'étude, dans les résultats d'apprentissage et dans les stratégies d'évaluation.

Cependant, l'expansion des stratégies d'évaluation pour aller au-delà de la base en matière de compétences n'est pas chose facile. Il est difficile, même inapproprié selon certains, d'établir des repères en matière de compétences des élèves dans les domaines tels que l'apprentissage socio-affectif, la créativité, la santé et la citoyenneté. Certaines compétences sont liées à un « processus » plutôt qu'à un « produit »; de plus, certaines de ces compétences sont observables uniquement dans le cadre des interactions sociales.

L'évaluation commence en classe

Depuis 2015, le personnel scolaire à travers la province de l'Ontario met à l'essai une série de compétences concrètes et observables dans les domaines de la santé, de la créativité, de l'apprentissage socio-affectif, et de la citoyenneté. On cherche à trouver des stratégies pour enseigner ces compétences en salle de classe, à suivre le progrès des élèves et leur offrir de la rétroaction bénéfique tout au long de leur cheminement scolaire.

Ces méthodes d'évaluation sont composées de listes de vérifications, d'observations, d'un journal de bord tenu par les élèves et de tâches collaboratives pour établir des barèmes à savoir à quelle fréquence les élèves font appel à une compétence particulière.

L'accent sur la croissance :

En utilisant un processus d'évaluation basé sur la rétroaction régulière – de l'enseignant à l'élève, de l'élève à l'enseignant et d'un élève à un autre – le personnel scolaire est en moyen de privilégier une orientation basée sur la croissance au lieu de miser uniquement sur les niveaux de compétences. On crée ainsi des milieux d'apprentissage où les élèves ont l'occasion de mettre en pratique les compétences, tout en tenant compte de leur propre cheminement scolaire.

Ce genre d'évaluation, misant toujours sur la rétroaction au lieu des résultats scolaires (les notes), crée chez les élèves une motivation intrinsèque (Deci, Koestner et Ryan 1999).

Dans une école en particulier, le personnel enseignant utilise un barème à échelle variable, où les élèves s'auto-évaluent, les enseignantes ou les enseignants évaluent les élèves à leur tour, tout cela de façon indépendante pour après comparer les résultats. On a découvert que l'emploi de cette échelle variable mène tout naturellement à une discussion portant sur l'amélioration des élèves.

L'évaluation collaborative à l'appui de l'apprentissage: Plusieurs études font état de l'importance de la collaboration entre les membres du personnel scolaire, contribuant à l'amélioration dans les écoles, ainsi que sa valeur en matière d'évaluation (p. ex. Cochran-Smith et Lytle, 2009; Hargreaves, Earl, Moore et Manning, 2002).

Grâce à cette collaboration, les membres du personnel scolaire sont en mesure d'explorer et d'évaluer leurs propres pratiques pédagogiques, créer des outils d'évaluation et suivre sur une base régulière l'impact qu'ont les conditions d'apprentissage en classe sur le développement des compétences chez les élèves.

L'évaluation comme moyen d'adapter les conditions d'apprentissage: Les données recueillies par le personnel scolaire servent « sur le champ » à ajuster les occasions d'apprentissage (l'évaluation au service l'apprentissage), pour que les activités en salle de classe répondent plus précisément au niveau du rendement des élèves. Ce type d'évaluation fait en sorte que le niveau général du rendement des élèves s'améliore, même chez les élèves avec un rendement inférieur, réduisant ainsi l'écart entre les différents groupes (Black et Wiliam, 1998, p. 141).

Un langage ou lexique spécifique pour une évaluation efficace: Étant donné que les compétences sont définies en termes spécifiques et observables, elles peuvent être utiles dans la formulation d'objectifs d'apprentissage. Ceci permet au personnel scolaire d'utiliser les données recueillies de diverses manières pour évaluer le progrès des élèves dans l'atteinte de ces objectifs.

La spécificité des compétences signifie également que les enseignantes et les enseignants sont en mesure de fournir à leurs élèves de la rétroaction qui fait état de la mise en pratique des compétences dans le cadre d'une tâche, d'une expérience ou d'un processus. La rétroaction portant sur la tâche, au lieu de la personne, est importante dans l'amélioration du rendement et les interventions par rapport à l'apprentissage socio-affectif des élèves.

Avantages et risques de l'évaluation par juridiction (locale, provinciale, nationale ou internationale)

Le but de l'évaluation à l'échelle de la salle de classe diffère des évaluations aux niveaux local, provincial, national et international. Dans ce dernier contexte, la mesure et l'évaluation procurent plus souvent des données sur le fonctionnement du système d'éducation, destinées primordialement aux preneurs de décisions. Les évaluations sur une grande échelle peuvent inclure la performance dans le cadre des tâches, y compris des tests et des essais, un questionnaire destiné aux tiers partis, (par exemple, les enseignantes ou les enseignants ainsi que les parents) ou des enquêtes menées auprès des élèves (autogestion et autorégulation).

Ces enquêtes peuvent prendre la forme d'un recensement (où l'ensemble de la population est appelé à répondre au questionnaire) ou d'un échantillonnage (le questionnaire est destiné uniquement à une portion de la population).

Bien que plusieurs systèmes scolaires explorent les moyens pour évaluer les compétences dans les domaines tels que la santé, l'apprentissage socio-affectif, et la citoyenneté, toujours dans le but de recueillir des données sur les milieux d'apprentissage à travers le système, les deux moyens sont assortis d'avantages et de risques.

Les avantages

Les avantages de l'évaluation à grande échelle peuvent :

- identifier les domaines où sont nécessaires un appui supplémentaire et la reddition de comptes portant sur la performance du système.
- fournir de l'information sur l'impact des politiques et des programmes en relation à la créativité, à la citoyenneté, à l'apprentissage socio-affectif et à la santé.
- suivre de près les progrès du système en matière de conditions d'apprentissage qui soutiennent le concept de résultats équitables dans ces domaines essentiels du cheminement scolaire.
- promouvoir la visibilité et l'importance de ces domaines dans un contexte de redevabilité en éducation qui se fie actuellement sur des indicateurs de rendement plus pointus.

Les défis

L'évaluation à grande échelle peut :

- fournir des données erronées pouvant être mal interprétées. Par exemple, il est possible d'évaluer une ou deux compétences dans le domaine de la créativité à un niveau ou une juridiction, sans pour autant conclure que ces données représentent le niveau de performance systémique dans ce domaine.
- créer la supposition, parfois fausse, que la réussite à un test fournit un bilan complet des habiletés des élèves dans la compétence mesurée (p. ex. Johansson, 2016; Sellar, Thompson et Turkowski, 2017).
- influencer les politiques, le financement et la reddition de comptes d'un système ayant comme résultats la restriction dans les pratiques pédagogiques et l'expérimentation ainsi que dans la prise de risques.
- augmenter les ressources et le temps de préparation de l'évaluation et de sa prestation en classe en tenant compte des sommes dépensées dans le cadre de l'administration de l'évaluation.

Réduire les risques

Étant donné l'importance de recueillir des données portant sur ces domaines d'apprentissage, il serait utile de trouver des moyens pour limiter les risques de l'évaluation à l'échelle locale, provinciale, nationale ou internationale.

Par exemple, l'évaluation standardisée administrée à un échantillon de la population d'élèves, portant sur une sélection de compétences, pourrait procurer des données au sujet de la performance du système et ce, selon le niveau ou l'année scolaire, tout en évitant les conséquences négatives entraînées par une évaluation à grande échelle (p. ex. FINEEC, 2016). Ou bien, des équipes au niveau local pourraient élaborer des normes et mener à bien des évaluations collaboratives basées sur le niveau ou l'année scolaire. Tout cela fournirait des informations ou des données au niveau de la juridiction (locale, provinciale, nationale ou internationale) sur la performance des élèves, ce qui exigerait bien sûr, des ressources, des infrastructures et des compétences additionnelles (p. ex. Mooney et Mausbach, 2008).



L'évaluation constitue une composante centrale du système d'éducation. Cela dit, avant de prendre la piste des évaluations à l'échelle des juridictions (locale, provinciale, nationale et internationale) portant sur la performance du système dans le développement des compétences en créativité, en apprentissage socio-affectif, en santé et en citoyenneté, il est important de comprendre davantage comment appuyer le personnel scolaire dans l'évaluation de ces domaines clés, et comment empêcher les évaluations de juridiction, c'est-à-dire à l'échelle locale, provinciale, nationale ou internationale, de dicter les pratiques pédagogiques au niveau de la salle de classe.

Évaluation en classe : au service de l'apprentissage, en tant qu'apprentissage et de l'apprentissage

Les spécialistes considèrent généralement qu'il existe trois types d'évaluation en classe, en fonction de l'intention, c'est-à-dire, dans une large mesure, de la façon dont l'information est utilisée (Black et Wiliam, 2007; Ontario, 2010).

- *Évaluation au service de l'apprentissage* – l'évaluation vise à déterminer ce que les élèves savent déjà et ce qu'ils doivent apprendre. Il peut s'agir d'une « évaluation diagnostique », utilisée avant l'enseignement d'un nouveau concept pour déterminer dans quelle mesure l'élève est prêt à apprendre, ou d'une évaluation continue des progrès de l'élève pendant le processus d'apprentissage. La rétroaction est un élément clé de l'évaluation au service de l'apprentissage.
- *Évaluation en tant qu'apprentissage* – les élèves s'évaluent eux-mêmes ou évaluent leurs camarades pendant une leçon.

L'évaluation en tant qu'apprentissage et l'évaluation au service de l'apprentissage sont formatives. La boucle de rétroaction continue créée pendant le processus est considérée autant comme un processus d'apprentissage que comme un processus d'évaluation (Ramaprasad, 1983; Ontario, 2010).

- *Évaluation de l'apprentissage* – dans ce cas, l'évaluation vise à déterminer la somme des connaissances et des habiletés que les élèves ont tirées de leurs apprentissages. C'est le type d'évaluation le plus souvent associé aux notes attribuées pour la réalisation d'un projet ou pour un test, ainsi qu'aux bulletins. Les résultats d'un test standardisé sont une autre forme d'évaluation sommative. L'évaluation de l'apprentissage est sommative plutôt que formative.

La distinction entre les trois types d'évaluation décrits ci-dessus tient à qui utilise l'information et à la manière dont il l'utilise. Un même instrument – un test écrit, par exemple – peut servir aux trois types. En effet, le test donne aux élèves la possibilité de comparer leur compréhension à une norme (évaluation en tant qu'apprentissage). C'est aussi un moyen que prend l'enseignant pour déterminer où en sont ses élèves dans l'apprentissage d'un concept (évaluation au service de l'apprentissage). Enfin, c'est aussi l'occasion de réfléchir à ce qui a été appris et de noter le rendement de l'élève (évaluation de l'apprentissage).

Commentaire de quelques enseignants sur les essais...

« On voit que peu importe où elle se déroule, la conversation à partir de l'échelle variable incite l'élève à annoncer lui-même ses prochaines étapes. Notre plus grande découverte! En inscrivant son point sur l'échelle, l'élève dit : "Eh bien, si j'avais fait ça, je serais ici [à un point supérieur] sur l'échelle". »

« Si je ne sais pas ce que je cherche, je ne peux pas le voir et je ne peux pas le nommer. Je ne peux pas y réagir. Je ne peux pas planifier la suite. [Les compétences visées par le programme Mesurer ce qui compte] nous aident à découvrir à quoi ressemble un apprentissage significatif, ce qui nous donne l'autonomie et la créativité nécessaires pour agir. C'est tout l'art et la science de la pédagogie. »

« Je trouve que les domaines et les compétences nous permettent de parler d'apprentissage et pas seulement de rendement. En nous attardant seulement aux attentes du curriculum, nous avons tendance à ne parler que de rendement. Soit l'élève peut faire telle chose soit il ne peut pas. On ne pense pas à la progression. On oublie les conditions d'apprentissage que nous sommes pourtant chargés de créer, que ce soit en tant qu'enseignant ou administrateur dans un bâtiment. »



References

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.

Black, P. & Wiliam, D. (2007). Large-scale assessment systems design principles drawn from international comparison. *Measurement*, 5(1), 1-53.

Cochran-Smith, M., & Lytle, S. (2009). *Inquiry as Stance: Practitioner Research in the Next Generation*. New York, NY: Teachers College Press.

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *The Psychological Bulletin*, 125(6), 627-668.

Finnish Education Evaluation Centre (FINEEC). (2016). *National Plan for Education Evaluation 2016-2019*. Helsinki, Finland: FINEEC.

Hargreaves, A., Earl, L., Moore, S. & Manning, S. (2002) *Learning to Change: Teaching Beyond Subjects and Standards*. San Francisco, CA: Jossey-Bass Education, Wiley.

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.

Johansson, S. (2016). International Large-Scale Assessments: What Uses, What Consequences? *Educational Research* 58(2), 139-48. doi:10.1080/00131881.2016.1165559.

Kempf, A. (2014). "Revising the pedagogic form: Test-oriented teaching and learning" In *The Pedagogy of Standardized Testing*. New York, NY: Palgrave MacMillan, 71-97.

Kyllonen, P. C. (2013). *Measurement of 21st Century Skills Within the Common Core State Standards*. US: K-12 Center, Education Testing Service. Retrieved from https://cerpp.usc.edu/files/2013/11/Kyllonen_21st_Cent_Skills_and_CCSS.pdf

Mooney, N. & Mausbach, A. (2008). *Align the Design: A Blueprint for School Improvement*. Alexandria, USA: ASCD.

Ontario Ministry of Education. (2010) *Growing Success: Assessment, Evaluation, and Reporting in Ontario Schools*. Queen's Printer for Ontario.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13.

Sellar, S., Thompson, G., & Rutkowski, D. (2017). *The Global Education Race: Taking the Measure of PISA and International Testing*. Canada: Brush Education.

Ce projet reçoit le soutien
des organismes suivants :



R. HOWARD WEBSTER
FOUNDATION

McConnell

The Murray R. O'Neil
Charitable Foundation



The Counselling
Foundation of Canada