

Évaluations internationales des systèmes éducatifs et évaluations externes nationales...

Points de convergences et de tension

Les élèves de la Communauté française de Belgique, comme dans la plupart des systèmes éducatifs du monde économiquement avancés, sont régulièrement soumis à ce que l'on appelle des évaluations externes. Conçues par des organismes ou groupes externes aux établissements scolaires, ces évaluations sont administrées à de grands nombres d'élèves de différents niveaux d'étude ou âges (qui constituent un échantillon représentatif), voire à l'ensemble des élèves.

Nous allons nous intéresser ici aux deux principaux types d'évaluations externes en cours actuellement en Communauté française:

- les évaluations externes non certificatives organisées par le Service général du Pilotage du système éducatif;
- les évaluations internationales menées à l'initiative de l'OCDÉ (Organisation de coopération et de développement économiques) (PISA) ou de l'I.E.A. (Association internationale pour l'évaluation du rendement scolaire).

Pour des raisons de place et de manque de recul historique, nous n'envisagerons pas, dans le cadre de cet article, les évaluations externes certificatives conduisant

à l'octroi du CEB ou celles qui prendront place à l'avenir à l'issue du 1^{er} degré du secondaire ou en fin de secondaire (TESS).

Nous nous attacherons d'abord à présenter les deux types de dispositifs en fonction d'une série de paramètres caractéristiques et à identifier ainsi leurs points de convergence et de tension. Nous envisagerons ensuite dans quelle mesure ces deux types d'évaluation débouchent sur des constats convergents.

Les évaluations externes non certificatives

L'idée de ces évaluations externes, bien avant les décrets qui les ont formellement instituées (Décret du 2 juin 2006, modifié et complété par le Décret du

30 avril 2009), revient à Jean Magy, ancien secrétaire général de l'Administration générale de l'enseignement, qui a mis sur pied le premier dispositif d'évaluations externes non certificatives en 1994-95 (en 5^e année primaire, français et mathématiques). Le dispositif s'est ensuite progressivement renforcé au fil du temps, en s'étendant à davantage de disciplines et d'années d'études. Il concerne aujourd'hui l'ensemble des élèves de P2, P5, S2 et S4 ou S5, qui sont testés chaque année en novembre dans l'une des disciplines suivantes: lecture-production d'écrit, mathématiques, éveil historique-géographique et scientifique. Le calendrier, ainsi que l'ensemble des documents

relatifs aux évaluations externes non certificatives peuvent être consultés et téléchargés sur www.enseignement.be, au départ de la rubrique L'école de A à Z, « en savoir plus », Evaluations.

Pourquoi des évaluations externes non certificatives ?

L'objectif premier des évaluations externes non certificatives est l'amélioration de la qualité des apprentissages; le destinataire privilégié en est l'enseignant dans sa classe.

Comme leur nom l'indique, les évaluations externes « non certificatives », à la différence de l'épreuve qui permet l'octroi du CEB, n'ont pas d'incidence sur le parcours individuel des élèves, elles ne peuvent pas servir « pour le bulletin ». Leur nature est diagnostique: il s'agit d'évaluer les acquis et les faiblesses des élèves pour les savoirs, savoir-faire et compétences contenus dans l'épreuve. Particularité appréciée des enseignants (Lafontaine & Dierendonck, 2008), le dispositif ne se contente pas de constater les lacunes des élèves; l'évaluation et l'analyse des résultats sont systématiquement suivies d'un document de Pistes didactiques, relativement étoffé, qui propose un ensemble d'outils ou de propositions concrètes pour travailler les apprentissages pour lesquels des difficultés sont constatées.

Par ailleurs, les résultats d'un échantillon représentatif de la Communauté française analysés par l'Université de Liège servent aussi à l'élaboration d'indicateurs pour le pilotage du système éducatif. Ce second objectif se rapproche de celui des évaluations internationales, mais il vient en second lieu par rapport à celui d'amélioration des apprentissages.

Qui conçoit les épreuves ?

Un groupe, composé de membres du Service du pilotage général de l'enseigne-

“ Dans les différents domaines évalués, la majorité des élèves maîtrisent bien les savoirs et les compétences les moins complexes. Dès que les opérations se complexifient, une proportion importante des élèves se trouve en difficulté et la disparité d'acquis se marque nettement, à tous les niveaux de la scolarité. ”

ment, d'inspecteurs, de chargés de missions et d'un chercheur universitaire, est chargé de la conception des épreuves. La composition du groupe diffère selon les années d'étude et les disciplines concernées.

Le contenu et la forme

Chaque épreuve annuelle porte sur un ensemble de compétences, savoir faire et savoirs relatifs à la discipline évaluée, sélectionné dans les référentiels de compétences. Tous les élèves passent le même test. Dans le secondaire, une épreuve spécifique est conçue pour les élèves de 2^e différenciée et pour les élèves de 4^e ou 5^e de l'enseignement de qualification.

L'épreuve comporte principalement des questions à choix multiple ou des questions ouvertes à réponse brève (aisées à corriger de manière standardisée), ainsi qu'un nombre limité de questions qui exigent une réponse plus élaborée.

Administration et correction

Ce sont les enseignants eux-mêmes qui administrent les tests dans leur classe et qui les corrigent, à l'aide de grilles de correction standardisées. Dans le

primaire, les corrections se font en groupe, avec le soutien de l'Inspection.

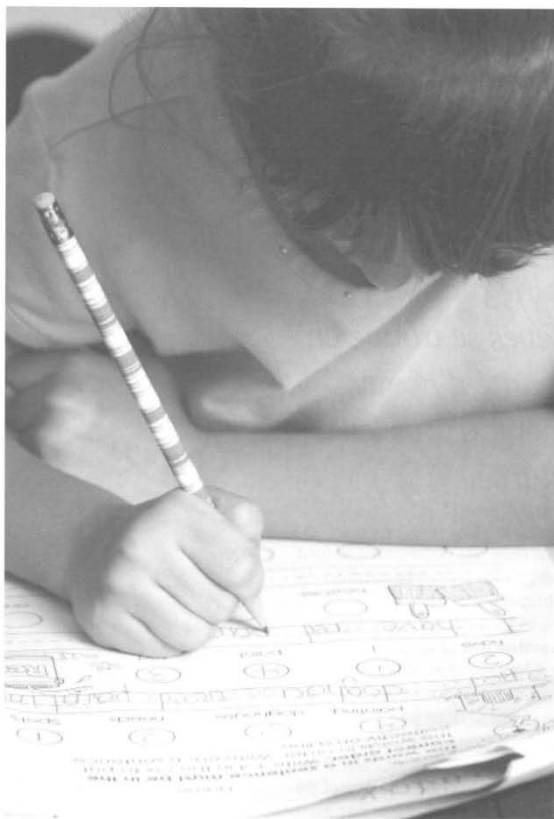
Gestion des résultats

L'enseignant, après correction, peut encoder les résultats de sa classe dans une grille informatisée. Cette grille lui donne un premier aperçu des résultats individuels de chaque élève et de l'ensemble de la classe par item et pour l'ensemble de l'épreuve, voire pour certains sous-scores.

Les données d'un échantillon représentatif d'élèves pour chacune des années d'études sont collectées et analysées par le service Analyse des systèmes et des pratiques d'enseignement de l'Université de Liège. Les résultats sont présentés dans un document envoyé à l'ensemble des établissements de la Communauté française.

Lorsque l'enseignant reçoit le document « Résultats et commentaires » présentant les résultats de l'échantillon représentatif de la Communauté française, il peut :

- comparer les résultats de sa classe à la moyenne de l'échantillon, soit pour le score global à l'épreuve, soit pour certaines compétences ou questions particulières;



- comparer les résultats de sa classe avec ceux de classes qui ont un profil proche de la sienne, par exemple les classes en discrimination positive, des classes comptant une proportion importante d'élèves en retard scolaire, une proportion élevée d'enfants ne parlant pas habituellement le français à la maison... ; c'est une façon de lire les résultats de sa classe en les relativisant, tant on sait que les contextes de classe sont différents d'une école à l'autre.

Publicité et accessibilité des données

Au terme de chaque opération, le contenu des épreuves est en accès libre sur le site www.enseignement.be.

Chaque enseignant dispose des résultats de sa propre classe. Il est prévu dans le Décret de 2006 que les directions analysent les résultats de leur établissement avec l'équipe pédagogique. Il est par ailleurs attendu que les services d'Inspection et d'animation pédagogique

intègrent les résultats et les apports des évaluations externes dans le cadre de leurs missions respectives. Enfin, l'IFC (Institut de la Formation en Cours de Carrière) organise des formations à l'exploitation des pistes didactiques et des résultats aux évaluations externes non certificatives auxquelles les enseignants peuvent s'inscrire dans le cadre des journées «macro».

Il est en revanche formellement interdit d'établir tout classement des établissements sur la base des évaluations externes ou de les utiliser à des fins publicitaires. Les résultats de la classe ou d'un enfant particulier peuvent être communiqués aux parents, mais aucune comparaison avec d'autres enfants, d'autres classes, d'autres écoles n'est autorisée.

Les évaluations internationales

L'histoire des grandes évaluations internationales de rendement éducatif débute dans les années 1960, sous les auspices de l'I.E.A (Association internationale pour l'évaluation du rendement scolaire). La Communauté française a régulièrement pris part à ces enquêtes ; on trouve dans Lafontaine & Blondin (2004) un aperçu de la participation et de l'évolution des performances de nos élèves dans différents domaines.

L'enquête la mieux connue du grand public est incontestablement le programme PISA (Programme pour le Suivi des acquis des élèves de 15 ans) de l'OCDÉ (OCDÉ, 2001 ; OCDÉ, 2004 ; OCDÉ, 2007). PISA est une enquête cyclique qui a lieu tous les trois ans (depuis 2000) et qui vise à évaluer les acquis des élèves à 15 ans, à l'âge où ils s'apprennent, dans la plupart des pays, à quitter la scolarité obligatoire. L'évaluation porte sur la compréhension en lecture, les ma-

thématiques et les sciences. Lors de chacun de ces cycles, l'un des domaines est évalué plus en profondeur. En 2000, c'était le cas de la lecture, en 2003 des mathématiques, en 2006 des sciences, en 2009 de la lecture à nouveau (et ainsi de suite).

La Communauté française participe aussi à l'enquête I.E.A. Pirls (*Progress in Reading Literacy Study*) en 2006 et 2011, qui porte sur les compétences en lecture des élèves de 4^e année primaire (Lafontaine, XXX). Pour plus d'informations sur les enquêtes de l'I.E.A., on peut consulter le site www.iea.nl

Finalités

Les principaux objectifs des évaluations internationales (Lafontaine & Simon, 2009) sont :

- de mesurer le niveau d'acquis des élèves dans un ou plusieurs domaines, de décrire ce dont sont capables ces élèves (sur des échelles hiérarchisées de compétences) ;
- de tenter d'identifier des facteurs du contexte familial ou éducatif susceptibles de rendre compte des différences en termes d'efficacité et d'équité, entre pays et à l'intérieur des pays.

Ces évaluations s'intéressent au système éducatif dans son ensemble. Elles s'adressent en premier lieu aux décideurs en matière de politique éducative, plutôt qu'aux acteurs du monde éducatif.

Population concernée et échantillon

Chaque système éducatif teste un échantillon représentatif d'élèves reflétant la diversité du paysage éducatif du pays. Dans PISA par exemple, la population concernée est constituée des élèves de 15 ans où qu'ils en soient dans leur scolarité (en avance, en retard, à l'heure dans leur parcours ; cela quelle que soit la filière). Dans chaque établissement, 35 élèves participent à l'évaluation ; ces élèves viennent de classes, voire d'années ou de filières différentes (Baye et al., 2009). Dans Pirls, c'est un niveau d'études qui est concerné : la 4^e année primaire et des classes entières sont sélectionnées (Lafontaine, 2009).

Contenu et formes de PISA et Pirls

Tous les élèves des pays participants passent des épreuves identiques qui ont été traduites dans les différentes langues. Les procédures de passation sont standardisées (les consignes sont les mêmes dans les différents pays). PISA et Pirls sont des évaluations papier-crayon, d'une durée de 2 heures par élève.

Il existe différentes formes parallèles de tests : dans une même classe (Pirls) ou une même école (PISA), les élèves se voient administrer des carnets différents, avec des parties communes et d'autres distinctes. L'ensemble est évidemment conçu pour pouvoir tirer des conclusions valides pour l'ensemble de l'échantillon et pour un ensemble d'items plus étoffé qu'il ne le serait si tous les élèves passaient le même carnet. Dans de telles conditions, on ne peut bien entendu pas comparer directement les pourcentages de réussite de deux élèves ayant passé des carnets différents. Cette analyse doit se faire à l'aide d'outils statistiques appropriés.

Comme dit précédemment, les performances des élèves pris individuellement ne sont pas la préoccupation première des évaluations internationales. Leur but est d'obtenir la meilleure mesure possible au niveau d'ensembles plus larges.

Contrairement à ce que beaucoup de personnes pensent, les évaluations internationales récentes, dont PISA et Pirls, recourent certes à des QCM, mais aussi, et d'une manière importante, à des questions ouvertes à réponse construite (environ 50% des questions) qui sont corrigées à l'aide de grilles de critères relativement complexes.

Les élèves complètent aussi un questionnaire de contexte. Plus de 40 questions leur sont posées à propos de leur environnement familial, de leurs loisirs, de leur motivation, de leurs attitudes, de l'établissement qu'ils fréquentent, de leur implication dans la vie scolaire etc.

Les chefs d'établissement complètent également un questionnaire de contexte fouillé portant sur les caractéristiques et le fonctionnement de l'établissement dont ils ont la responsabilité.

Administration et correction

Les épreuves sont administrées par du personnel exter-

ne aux établissements, formé à cette fin par l'Université de Liège, qui a la responsabilité de mettre en œuvre PISA et Pirls en Communauté française de Belgique.

De même, les questions ouvertes qui nécessitent une correction manuelle sont corrigées par une équipe de correcteurs dûment formés pour accomplir cette tâche. Les mêmes réponses d'élèves sont codées à plusieurs reprises par des correcteurs différents; on s'assure ainsi, au niveau international, que les réponses sont corrigées d'une manière fiable dans les différents pays.

Gestion et communication des résultats

Les enseignants et les directions ne gardent aucune trace des réponses des élèves. Les administrateurs de test emportent avec eux les carnets qui seront traités de manière anonyme à l'Université de Liège d'abord, au centre international responsable des opérations PISA (*Australian Council of Educational Research*) ensuite.

Après correction et analyse des résultats, l'Université de Liège communique un dossier aux établissements participants qui leur présente les résultats de leur établissement et ceux de l'ensemble de l'échantillon de la CFB. Les résultats individuels des élèves ne sont pas communiqués; comme on l'a vu, ils ne sont en effet pas directement comparables.

Publicité et accessibilité des données

Le contenu des épreuves internationales est pour l'essentiel gardé sous embargo; il est en effet appelé à être réutilisé lors des opérations suivantes pour mesurer les évolutions dans le temps. C'est là l'une des grandes forces des évaluations internationales - fournir une estimation valide des progrès des acquis à moyen

et à long terme - qui a pour prix la mise sous embargo des épreuves. Un certain nombre de tâches et de questions d'évaluation sont néanmoins rendues publiques, à titre illustratif.

L'une des autres caractéristiques des évaluations internationales et de PISA en particulier est le battage médiatique qui entoure, lors de chaque cycle, la communication des résultats. Conférence de presse internationale en différents endroits du monde, conférences de presse nationales ou régionales assurent une visibilité importante aux résultats.

L'OCDÉ fait clairement de la communication des résultats un enjeu d'importance. Elle ambitionne effectivement de sensibiliser les décideurs et responsables des politiques éducatives aux résultats de PISA. Plusieurs rapports, du plus technique (*Technical Report*) au plus accessible (*Executive summary*), sont disponibles en version papier ou téléchargeables sur le site web (www.oecd.pisa.org); les bases de données internationales sont également en accès libre et gratuit pour des analyses secondaires.

À l'échelle de la CFB, les chercheurs de l'Université de Liège (aSPe dirigé par D. Lafontaine) en charge de PISA s'attachent aussi à diffuser les résultats de PISA, en veillant à une interprétation rigoureuse des résultats, située dans son contexte, et surtout à ne pas faire de PISA l'instrument ultime à l'aune duquel notre système pourrait être évalué et en insistant sur la nécessaire complémentarité des approches « communautaires » et internationales.

Des finalités distinctes

Comme l'a montré cette rapide comparaison des évaluations internationales et des épreuves externes non certificatives, pas mal de différences séparent les deux approches. Alors que le

contenu et la forme des épreuves sont plus semblables qu'on ne pourrait le penser a priori, ce sont surtout les finalités des deux types d'évaluation et leur destinataire privilégié qui les distinguent. Alors que les évaluations internationales portent sur le système et s'adressent en priorité aux décideurs, les évaluations externes « nationales » visent d'abord les enseignants et s'inscrivent dans une perspective d'amélioration des apprentissages. Il n'empêche que l'on peut aussi, au départ des évaluations externes non certifiantes, tirer un certain nombre de constats au niveau de l'ensemble des élèves ; la qualité des épreuves et le caractère représentatif de l'échantillon le permettent en effet. On ne peut toutefois pas se prononcer, au départ des évaluations externes, sur l'évolution dans le temps des acquis des élèves, les épreuves étant différentes d'une opération à l'autre.

Les constats que l'on peut tirer de ces deux types d'évaluation présentent pas mal de convergences. Ainsi, toutes deux mettent en évidence des résultats moyens qui pourraient apparaître comme satisfaisants à première vue ; mais ceux-ci vont de pair avec la présence d'un nombre relativement important d'élèves en grande difficulté (les évaluations internationales nous indiquent que cette proportion d'élèves en difficulté est plus élevée que dans la plupart des autres systèmes éducatifs de l'OCDE). Dispersion importante des acquis, résultats significativement liés à l'origine sociale des élèves et à l'établissement fréquenté font également partie des constats convergents.

En termes plus qualitatifs ou diagnostiques, le tableau qui se dégage des deux types d'évaluation quant à ce que maîtrisent les élèves de la CFB et quelles sont leurs lacunes est assez largement semblable : dans les différents domaines évalués – lecture, mathématiques, sciences – la majorité des élèves maîtrisent bien les savoirs et les compétences les moins complexes. Ils sont capables d'effectuer les tâches cognitivement les moins coûteuses – restituer une information explicite, effectuer une opération routinière, par exemple. Dès que les opérations se complexifient, une proportion importante des élèves se trouve en difficulté et la disparité d'acquis se marque nettement, à tous les niveaux de la scolarité.

Si, dans PISA, la CFB de Belgique apparaît comme l'un des systèmes éducatifs où l'écart entre les élèves les moins et les plus performants est le plus important de l'OCDE, ceci tient en grande partie aux caractéris-

tiques particulières de l'échantillon PISA. Pour rappel, PISA porte sur les élèves de 15 ans et va les chercher là où ils sont, répartis qu'ils sont entre filières et années d'étude. Vu les taux élevés de redoublement dans notre communauté (parmi les plus élevés des pays de l'OCDE), le tableau qui se dégage de PISA en matière d'inégalités d'acquis est particulièrement défavorable. Il s'inscrit toutefois dans la continuité de ce que mettent en évidence les autres évaluations, en jouant simplement les miroirs grossissants.

Si les évaluations internationales et nationales en CFB ne sont pas, de notre point de vue, en importante tension, nous voulons néanmoins, au moment de conclure, insister sur leur nécessaire complémentarité (Lafontaine, Soussi & Niddegger, 2009). Si nous ne disposons que des évaluations externes, même si celles-ci s'articulent aux référentiels en vigueur, il serait malaisé de se prononcer sur la qualité de niveau atteint : compte tenu de l'âge des élèves, est-ce suffisant, excellent, insatisfaisant ? Disposer d'un point de comparaison international – sans y être inféodé – est à cet égard capital. A l'opposé, si nous ne disposons que des évaluations internationales, il serait en revanche malaisé de se prononcer sur les acquis de nos élèves, compte tenu des attentes particulières que fixe notre système pour tous les élèves (dans le primaire et au 1^{er} degré) et pour certaines catégories d'élèves (selon la filière fréquentée ou l'année d'études). C'est en articulant les informations issues de ces deux sources complémentaires et en gardant toujours à l'esprit les caractéristiques propres du dispositif qui les a produites que l'on peut, à notre sens, tirer les enseignements les plus riches en vue d'améliorer l'efficacité et l'équité de notre système éducatif. ■

Sources

- Baye, A., Fagnant, A., Hindryckx, G., Lafontaine, D., Matoul, A., Quittre, V. (2009). *Les compétences des jeunes de 15 ans en Communauté française en sciences, en mathématiques et en lecture. Résultats de l'enquête PISA 2006*. Université de Liège : Cahiers des Sciences de l'Éducation, 29-30, 245 pages.
- Lafontaine, A. (2009). *PIRLS – Progress in International Reading Literacy Study. Rapport final*. Liège : Université de Liège, aSPe. Document inédit.
- Lafontaine, D. et Blondin, C. (2004). *Regards sur les acquis des élèves en Communauté française Apports des enquêtes de l'I.E.A. de PISA et des évaluations externes*. Bruxelles : De Boeck.
- Lafontaine, D., & Simon, M. (2009). L'évaluation des systèmes éducatifs. *Mesure et évaluation en éducation*, 31,3, 95-125.
- Lafontaine, D., Soussi, A., & Niddegger, C. (2009). Évaluations internationales et épreuves nationales : tensions et utilité pour les pratiques de classes. In L. Mottier-Lopez et M. Crahay (Eds), *Évaluations en tension*. Bruxelles : De Boeck. 61-81.
- Lafontaine, D., & Dierendonck, C. (2008). Comment les évaluations externes des acquis des élèves sont-elles perçues par les enseignants en Communauté française de Belgique et en Suisse romande ? In L. Mottier Lopez, Y.-E. Dizerens, G. Marcoux & A. Perréard Vité (Eds). *Entre la régulation des apprentissages et le pilotage des systèmes : évaluations en tension. Actes du 20^e colloque de l'ADMEE-Europe*, Université de Genève.
- OCDE (2001). *Connaissances et compétences : des atouts pour la vie. Premiers résultats de PISA 2000. Enseignement et compétences*. Paris : OCDE.
- OCDE (2004). *Apprendre aujourd'hui, réussir demain. Premiers résultats de PISA 2003*. Paris : OCDE.
- OCDE (2007). *Les compétences en sciences, un atout pour réussir. Volume 1 : Analyse des résultats*. Paris : OCDE.