
LES EVALUATIONS EXTERNES EN COMMUNAUTE FRANÇAISE DE BELGIQUE : QUEL SENS DONNER A LA MESURE ?

Dominique Lafontaine

Résumé

Quelles sont les origines et les modalités du dispositif d'évaluation externe mis en place en Communauté française de Belgique à partir de 1994-1995 ? Quelle(s) fonc-

tion(s) est-il supposé remplir ? L'analyse met en évidence le paradoxe d'un dispositif institutionnel rejeté à la marge du cadre législatif et dont la définition des finalités est laissée à l'appréciation des acteurs appelés à le mettre en oeuvre.

Contexte

La Communauté française de Belgique est l'un des systèmes éducatifs de l'Union européenne où la marge de manoeuvre des acteurs pédagogiques est la plus élevée. Ainsi, les décisions de passage de classe, d'orientation, les évaluations certificatives, les examens d'entrée, le choix des manuels, ont toujours eu un caractère local (i.e. ils ne résultent pas d'un mouvement historiquement daté vers plus de décentralisation) et ne font l'objet d'aucune régulation externe organisée.

La Communauté française de Belgique est aussi un système éducatif connu pour ses taux de redoublement particulièrement élevés (Crahay, 1996) et une importante disparité de performances en fonction de l'établissement fréquenté (Lafontaine, 1996). La mise en

évidence de ces indicateurs d'inégalité par le rapport sur l'état du système éducatif publié par l'OCDE au début des années '90 (OCDE, 1991) a conduit à mettre sur pied, en 1994-95, un dispositif d'évaluation externe des performances des élèves en français et en mathématique à différents paliers de la scolarité (2e, 4e, 6e années primaires, 2e année secondaire).

Pour faire bonne mesure...

A l'initiative de ce dispositif, on trouve un haut responsable de l'Administration du Ministère de l'Education, soutenu par une importante partie de l'inspection de l'enseignement primaire. Cette initiative a donc une origine individuelle plutôt qu'institutionnelle, même si elle est implicitement soutenue par le Ministre de l'Education

de l'époque (Magy, 1998). Mais elle ne fait l'objet d'aucun texte législatif.

Dans le même esprit, les sources de financement de l'opération sont peu assurées. Il s'agit de budgets de recherche, alloués annuellement à des universitaires qui auront la charge de l'opération. Enfin, les objectifs des évaluations, ainsi que leurs modalités, sont déterminés, avec plus ou moins de précision, par les membres d'une cellule de pilotage aux contours fluctuants.

Ainsi, au départ, lors de sa création en 1994, ce dispositif répond, dans la tête de son initiateur, à un double souci :

- a) sortir les enseignants de leur isolement, en leur fournissant un référentiel (les résultats d'un échantillon représentatif) auquel comparer les performances de leur classe. L'outil se veut formatif et diagnostique. Il est passé en début d'année.
- b) réduire les disparités de rendement entre écoles, dont l'importance a été soulignée par les recherches internationales (IEA, OCDE).

Dans les faits, le dispositif, fortement investi par les services d'inspection du primaire, est surtout utilisé, lors des deux premières années, comme un « levier didactique », i.e. comme un instrument permettant aux inspecteurs de faire passer leur message pédagogique (test « modèle », accompagnement de l'évaluation par un document de pistes didactiques exploité sur le terrain par les inspecteurs). Ainsi, alors que l'évaluation se veut officiellement « au service des enseignants », elle est d'abord au service

de l'inspection et mise implicitement sur l'effet en retour (*backwash effect*) de l'évaluation sur les pratiques (Anderson & Wall, 1996).

Progressivement, avec le passage de l'évaluation dans le secondaire en 1996-97, et le départ à la retraite de l'initiateur du dispositif, on assiste à un recul du rôle de l'inspection primaire. Le côté « levier didactique » cède alors à la place à une forme d'évaluation plus sommative : ce sont plutôt les chercheurs et l'inspection du secondaire qui orientent alors le mouvement. On s'inscrit davantage dans une dynamique d'évaluation du système. Les documents de pistes didactiques sont abandonnés, officiellement pour des raisons budgétaires. Des indicateurs sont élaborés, tant en termes d'équité que d'efficacité : quels pourcentages et quelles catégories d'élèves atteignent certains niveaux de compétences (utilisation des analyses IRT¹), calcul d'indices de disparités (coefficients de corrélation intraclasses), disparités de réussite en fonction de la composition des classes ou de certaines caractéristiques des élèves, ancrage avec les enquêtes internationales (Lafontaine, 1996; Lafontaine, 1997).

Ce changement de cap illustre bien le statut particulier de ces évaluations. Un dispositif institutionnel existe, mais, paradoxalement, c'est aux acteurs qui le mettent en oeuvre qu'est entièrement dévolu le rôle de lui donner son sens. En amont, il n'y a pas de politique d'évaluation impulsée par les décideurs et, en aval, pas de demande citoyenne

¹ Voir Laveault et Grégoire (1997).

forte pour que l'école rende des comptes. Quand une école dysfonctionne, les parents ont toujours le choix de changer leurs enfants d'établissement, puisque le choix de celui-ci ne souffre aucune restriction (aucune forme de carte scolaire).

Des évaluations presque honteuses...

En 1997, voit le jour un texte législatif d'importance pour le système éducatif de la Communauté française de Belgique : le « Décret définissant les missions prioritaires de l'enseignement fondamental et de l'enseignement secondaire et organisant les structures propres à les atteindre ». Ce décret, qui coule dans un texte de loi la charpente du système éducatif, institue, dans un chapitre intitulé "du pilotage", trois types de commissions : une commission chargée de définir les socles de compétences, une commission des programmes et une commission responsable des outils d'évaluation. La composition des commissions est décrite avec une abondance de détails (il faut respecter tous les savants équilibres politico-administratifs desquels notre système éducatif se nourrit et dans lesquels il perd son sens).

En revanche, aucune relation n'est établie entre le dispositif de pilotage préexistant et ces nouvelles dispositions législatives. Le dispositif de pilotage se trouve ainsi en quelque sorte « oublié » dans le texte légal, mais continue néanmoins à fonctionner et à recevoir des crédits du

même Ministre de l'Education. Le texte de loi, qui institue par ailleurs plusieurs réformes d'importance, ne prévoit ni évaluation des performances des élèves, ni évaluation des effets des réformes. Il envisage donc un pilotage sans instruments, « à vue », s'en remettant sans doute implicitement à des actions de micro-pilotage orchestrées par les services d'inspection. Mais le texte de loi n'en dit rien.

Ceci confirme l'absence de politique d'évaluation dénoncée plus haut. Le paradoxe est ici à son comble, où l'on voit le texte de loi reprendre un terme dans l'air du temps (le pilotage) - toujours pour faire bonne mesure - mais en le vidant de sa substance. Qui dit pilotage dit en effet évaluation et régulation... Or ni l'un ni l'autre ne sont présents au rendez-vous. Par ailleurs, le paradoxe mis en évidence plus haut à propos des évaluations externes (création d'un cadre sans finalités définies) se répète à propos de la commission des outils d'évaluation. Cette commission « produit des batteries d'épreuve d'évaluation étalonnées et correspondant aux socles de compétences. Le Gouvernement les diffuse, à titre indicatif, auprès de tous les établissements organisés ou subventionnés par la Communauté française, selon les modalités qu'il définit » (Décret sur les missions, p.12). Quelques mois plus tard, un appel d'offre de recherche est lancé pour la construction de ces épreuves, comme s'il s'agissait d'une simple instrumentation. A nouveau, la question des finalités de ces épreuves n'est pas posée; c'est aux acteurs (services de recherche) que revien-

dra le rôle de réfléchir à cette question fondamentale.

Par ailleurs, il est important de souligner le fait que ces épreuves seront diffusées à titre indicatif dans les établissements. A titre indicatif de quoi et pour quel usage ? Nul ne le sait. On retrouve là encore le signe d'une ambiguïté fondamentale du fonctionnement du système éducatif en Communauté française de Belgique. Différents facteurs ou dysfonctionnements font sentir le besoin de resserrer les mailles d'un filet par trop distendues (trop de disparités, trop d'inéquités) - de là la création de socles de compétences, d'évaluations externes, d'épreuves étalonnées -, mais personne ne semble en mesure d'endosser la responsabilité politique de faire aboutir des mécanismes de contrôle qui sont très étrangers au fonctionnement habituel de nos institutions. Du bout des lèvres, on fait donc comme si..., espérant obtenir les effets en faisant l'économie des mesures.

Le sens de la mesure

Malgré l'intérêt des analyses effectuées et la richesse des bases de données constituées via les évaluations externes annuelles, le sentiment domine, chez les chercheurs impliqués dans ces évaluations externes (dont l'auteur de cet article fait partie), que le dispositif oublié du cadre législatif, est aujourd'hui en panne de sens institutionnel (l'opération prévue pour 1999-2000 a d'ailleurs été suspendue). Quel sens et quelles fonctions sont en effet appelées à

remplir des évaluations externes ainsi rejetées à la marge du système, et dont rien n'est officiellement attendu en termes d'effet en retour sur les politiques mises en place par le système ?

On ne peut en effet se contenter de naviguer/piloter à vue en pratiquant une évaluation protéiforme qui rencontrerait plusieurs fonctions de l'évaluation et satisferait les besoins de plusieurs catégories d'acteurs (enseignants, direction, inspection, parents, élèves, décideurs, chercheurs...). A chaque public visé correspond en effet une fonction prioritaire de l'évaluation et des modalités plus propices à les satisfaire (Farr, 1999). C'est pourquoi il est essentiel que le politique au sens large donne des indications précises sur les effets attendus de l'évaluation externe. Ce n'est pas au travers d'un même instrument d'évaluation que l'on amènera les enseignants à changer leurs pratiques d'évaluation (et en amont leurs pratiques d'enseignement), que l'on rendra compte de l'état du système éducatif, que l'on évaluera les disparités entre établissements et encore moins que l'on contribuera à les réduire.

Pour un fonctionnement efficace, une architecture d'ensemble de l'évaluation est nécessaire, qui articule des finalités avec des modalités concrètes de mise en oeuvre et les mesures d'accompagnement en amont et en aval.

Pour l'heure, la place donnée aux évaluations externes et la manière dont fonctionnent celles-ci en Communauté française de Belgique a au moins un mérite : celui de mettre le système éducatif à l'abri des dérives que l'on observe à large échelle dans

les pays où l'obsession d'évaluer pour rendre des comptes (*accountability*) a enfermé les enseignants dans une voie sans issue - enseigner pour que leurs élèves réussissent le test (*teaching for the test*) (FARR, 1999). Contre ce risque, la Communauté française de Belgique est vaccinée, et pour longtemps encore.

Il ne faudrait cependant pas en conclure que l'évaluation au sens large occupe une place secondaire dans le fonctionnement de notre système éducatif. L'évaluation interne, menée dans les établissements, et dont les enseignants ont l'entière responsabilité, est omniprésente, diffuse et imprévisible. Seuls les acteurs les plus avertis (les meilleurs élèves et/ou ceux issus des milieux les plus favorisés) sont en mesure d'en situer les enjeux et d'en déjouer les pièges. Faute de lieux et de moments précis d'évaluations externes qui prennent le risque de certifier, sur une base commune, que les élèves ont atteint les niveaux de compétences attendus à tel ou tel palier de la scolarité, tout se juge à l'aune des évaluations internes¹. Et ceci, les recherches l'ont abondamment montré (pour une synthèse, voir Crahay, 1996; Lafontaine, 1997), est une mesure dont on connaît le sens et les effets -

produire des inéquités qui sanctionnent en première ligne ceux qui n'auront pas posé les bons choix (d'école, de filières, d'options...) dans un système qui octroie à chacun cette liberté.

- Anderson, J.-C. & Wall, D. (1993). Does Washback exist ? *Applied linguistic*, Vol. 14, 2, 115-129.
- Crahay, M. (1996). *Peut-on lutter contre l'échec scolaire ?* Bruxelles : De Boeck.
- Farr, R. (1999). Putting it all together : solving the reading assessment puzzle. In S.J. Barrentine (Ed.), *Reading Assessment*. International Reading Association.
- Lafontaine, D. (1996). *Performances en lecture et contexte éducatif. Enquête internationale menée auprès d'élèves de 9 à 14 ans*. Bruxelles : De Boeck.
- Lafontaine, D. (1997). Le niveau en lecture à l'entrée du secondaire en Communauté française de Belgique : des compétences fragiles et si diverses... *Enjeux*, 41-42, 129-161.
- Laveault, D. & Gregoire, J. (1997). Les modèles de la réponse à l'item. In *Introduction aux théories des tests en sciences humaines*. Bruxelles : De Boeck.
- Magy, J. (1998). *Oser la qualité dans l'enseignement en Communauté française de Belgique*. Namur : Editions Erasme.
- OCDE (1991). *Les systèmes éducatifs en Belgique : similitudes et divergences*. Bruxelles : Ministère de l'Éducation, de la Recherche et de la Formation.

¹ Il n'existe pas de bac; l'examen d'accès à l'enseignement supérieur (dit examen de maturité) - qui était une épreuve interne - a disparu depuis quelques années. Le certificat d'études de base, certifiant la réussite des études primaires, est octroyé sur la base d'une épreuve interne. Des acteurs locaux (inspecteurs, pouvoirs organisateurs tels que la commune ou le diocèse) peuvent décider d'organiser une épreuve externe commune à plusieurs établissements, mais cela relève de leur initiative. La difficulté de ces épreuves varie sensiblement selon les endroits et les chances d'obtenir son certificat varient en conséquence.