

Quelle évaluation pour quelle pédagogie?

Laurent Lescouarch

En cette rentrée scolaire, la plupart des classes de France et de Navarre étaient en période d'évaluation. On a évalué les acquis, les manques, pour pouvoir concevoir des dispositifs adaptés aux besoins des élèves conformément au leitmotiv repris en chœur par tous les derniers textes officiels. Ainsi, après une circulaire spécifique en 2006¹, la circulaire de rentrée 2007² met particulièrement l'accent sur le développement de protocoles nationaux d'évaluation diagnostique à l'école et fixe explicitement des modalités nouvelles en supprimant les évaluations CE2 et sixième pour leur substituer une évaluation en fin de CE1, CM2.

Ces orientations sont à mettre en relation avec les conceptions présidant à la mise en œuvre du socle commun de connaissances et de compétences³ dans l'esprit de la loi de 2005. La dimension quasi expérimentale donnée à ces protocoles (avec un support d'exercice identique sur tout le territoire national, une passation minutée) est un gage de sérieux dans l'esprit du public et renvoie à tous les clichés scientistes. On pourrait facilement imaginer les enseignants revêtant une blouse blanche pour rentrer dans leur classe transformée pour un temps en « laboratoire », les autres modalités d'évaluation jugées dépassées et peu objectives étant symboliquement renvoyées à la préhistoire de la pédagogie.

Toutefois, avant de céder aux sirènes de la techno-pédagogie (au sens d'une conception de la pédagogie comme application de techniques « scientifiques »), il faut rappeler que ces outils se caractérisent par un problème majeur de fiabilité en relativisant considérablement la portée. En effet, lorsque les résultats à ces évaluations sont utilisés pour juger de l'efficacité des politiques éducatives en France, on oublie bien souvent de mettre en avant leurs limites. Qui sait par exemple (en dehors des enseignants les faisant passer) que les mêmes exercices ont été utilisés pendant des années dans les évaluations CE2 avec un système de roulement d'une année sur l'autre⁴ ?

Il n'est pas besoin d'être chercheur pour percevoir en quoi cette seule disposition invalide complètement toutes conclusions péremptoires à partir des

¹ MEN (2006) *Dispositif national d'évaluation diagnostique – année 2006-2007*. Circulaire n°2006-095 du 9 juin 2006 In BOEN N°24 d 15 juin 2006.

² MEN (2007) *Préparation de la rentrée 2007*. Circulaire N°2007-011 DU 9-1-2007 BOEN N° 3 du 18 janvier 2007

³ MEN (2007) *Mise en oeuvre du socle commun de connaissances: boraires et programmes d'enseignement de l'école primaire*. In BOEN Hors Série N°5 du 12 avril 2007.

⁴ Ces outils d'évaluation sont consultables au 24/09/2007 sur le site ministériel suivant : <http://cisad.adc.education.fr/eval/index.htm>

résultats de ce dispositif... La validité scientifique de telles évaluations est donc limitée à plusieurs niveaux : j'ai pu observer dans des écoles l'utilisation des évaluations de l'année précédente comme préparation à cette épreuve à la fin du CE1, des conditions de passation (temps, aide de l'adulte) pouvant varier d'une classe à l'autre. De même, la pertinence des items et des supports est toute relative (des exercices censés évaluer des compétences mathématiques était par exemple échoués systématiquement par les mauvais lecteurs pour des motifs de... lecture).

En prenant en compte leur qualité pédagogique parfois douteuse et leurs biais importants, quelle peut être la validité de telles épreuves ? Loin de refléter une vérité scientifique, elles ne constituent qu'un repère approximatif de la situation des élèves dans les apprentissages. Elles nous permettent effectivement de fonder une démarche pédagogique mais celle-ci est plus proche du repérage par les étoiles en navigation que du recours au GPS. Pourtant, c'est à partir des données statistiques issues de ces protocoles que sont élaborés les tableaux de bord de l'efficacité du système et que sont pensés tous les dispositifs de remédiation, ce qui peut être assez inquiétant.

Ces constats ne sont pas nouveaux et on pourrait considérer que c'est de l'histoire ancienne puisque les évaluations CE2 sont amenées à disparaître. Il serait commode de penser que tous les bilans successifs effectués vont nous permettre d'avoir aujourd'hui des évaluations scientifiquement fiables et indiscutables. Or, force est de constater que la réalité est bien différente puisqu'à cette rentrée, des collègues ont attiré mon attention sur la réutilisation des supports tirés des ex évaluations CE2 proposés aux élèves pour les évaluations CE1 et CM2.

De façon comique (ou dramatique), on retrouve même comme par le passé des exercices (qui ont déjà été tour à tour utilisés au niveau CE2 puis 6^{ème}) être proposés aujourd'hui pour les CE1 et CM2 la même année (Ainsi l'exercice 8 (p.4 et 5) des évaluations CE1 2007 est repris dans l'exercice 33 (p.32) de l'évaluation CM2 2007...).

Cela ne serait pas plus problématique que ça si ces évaluations ne prétendaient pas à une forme de scientificité⁵ et ne constituaient pas la base de la normalisation pédagogique en cours dans la culture du résultat. Le HCE invite ainsi dans son rapport à une utilisation systématique de ces évaluations par les enseignants pour améliorer les résultats des élèves. Dans certaines circonscriptions, les résultats de la classe font donc partie intégrante en cette rentrée des éléments obligatoires à donner à l'inspecteur dans les documents préparatoires à l'inspection. En cette période de valorisation du mérite individuel, quel sort sera réservé à ceux qui ont des résultats faibles dans leur classe ? Seront-ils aidés ou culpabilisés par rapport aux résultats de leurs élèves ? La réponse risque d'être à géométrie variable.

En outre, du fait que dans de nombreuses classes, toute évaluation de fin d'année influence l'organisation des activités durant la période précédente (il n'y a pas que les enfants qui sont évalués à travers les évaluations....), la qualité et la fiabilité de ces outils devient un enjeu considérable. Les items fonctionnent comme un programme bis alors même que ceux-ci ne sont pas tous représentatifs des programmes réels et que l'outil n'est pas fiable.

⁵ Il est notable que les protocoles insistent dans leur introduction sur les limites de ces outils mais, comme pour les sondages, les précautions méthodologiques d'interprétation sont rarement relayées.

De plus, quand bien même ces limites méthodologiques seraient résolues par la construction de protocoles plus pertinents, le principe sous jacent consistant à s'appuyer exclusivement sur ce type d'évaluations pour penser les pratiques pédagogiques est réducteur et discutable. En premier lieu, contrairement à l'idée des cycles permettant une différenciation des rythmes et modalités d'apprentissage sur 3 ans, le recours systématique chaque année à des évaluations diagnostiques réintroduit (sans le dire explicitement) une logique de programmation annuelle fermée des contenus d'apprentissage.

Cela joue également sur la forme des aides aux élèves en difficulté. Dans les Programmes Personnalisés de réussite Educative (PPRE), les objectifs des projets de différenciation sont bien souvent construits à partir des items ratés. Il se met donc en place tout une technologie de la remédiation à partir de ces supports participant du projet post-behavioriste consistant à vouloir déduire les pistes de remédiation du comportement observable et limitant les actions à un soutien scolaire sous forme de reprise.

En tant qu'ancien praticien des aides aux élèves en difficulté, je ne peux m'inscrire qu'en faux par rapport à ces pratiques. En effet, une fois que je sais par un test ou épreuve que tel enfant ne maîtrise pas le son « in », qu'est ce que je sais de sa difficulté pour y remédier ? ... Rien.

Cette information est nécessaire mais pas suffisante: Est ce que c'est parce qu'il ne discrimine pas les sons ?...Parce qu'il pense que ça ne sert pas pour lire et ne le mémorise pas ?... Parce qu'il n'a aucun projet d'apprendre à lire.... ? Le seul moyen de le savoir est de dépasser ces constats pour entrer dans la démarche compréhensive du pédagogue.

Comprendre le motif de la difficulté suppose en effet en complément le recours à d'autres types d'évaluations pour dépasser la simple gestion mécanique des résultats au test. Il faut se méfier de la logique de réponse stéréotypée héritée de la pédagogie par objectifs sur le mode : « *Il ne connaît pas le son [in], vous lui en remettrez deux tranches...* ». Une compréhension fine de la source de la difficulté pourrait a contrario permettre d'envisager des aides qui ne sont pas toutes nécessairement scolaires.

Je crois donc qu'en complément d'un dispositif diagnostique amélioré permettant de repérer les difficultés, il nous faut réhabiliter des modes d'évaluation apparentés à la démarche clinique même s'ils ne sont plus toujours très bien perçus dans l'institution. L'observation critériée reste un outil essentiel de l'enseignant permettant de s'appuyer sur le ressenti subjectif (mais en le mettant à distance), à partir de critères explicites, une formulation d'hypothèses confrontées ensuite à des observables. Cela revient à privilégier la pédagogie et sa dimension de tâtonnement expérimental (en mettant à distance la tentation de constituer une technopédagogie applicative dont les réponses aux problèmes posés seraient systématiquement contenues au préalable dans le diagnostic).

Il ne s'agit donc pas de se positionner contre ces évaluations diagnostiques mais de questionner leurs modalités et leur portée. En tant qu'enseignant ou formateur d'enseignant, il nous faut donc repenser cette question et revendiquer une utilisation de toute la gamme des modalités de recueil d'information en ne se contentant pas des seules évaluations inspirées des logiques de tests. On confond bien trop souvent évaluation et mesure comme nous le rappellent fréquemment les collègues psychologues à propos des usages abusifs des tests de QI.

Toute la difficulté est donc d'arriver à se détacher d'une pensée héritée des sciences expérimentales pour intégrer les méthodologies des sciences humaines d'inspiration qualitative dans les pratiques pédagogiques.

Pour l'instant, les enseignants sont formés essentiellement à construire et utiliser des outils de mesure des résultats des élèves. Il serait intéressant que la formation intègre également l'apprentissage de l'observation, l'entretien d'explicitation et l'analyse des réactions de l'élève. C'est à cette condition que l'exigence d'évaluation intégrée à la démarche du socle commun de connaissances et de compétences pourrait prendre sens et constituer un réel progrès.

Pour le pédagogue habitué à la complexité de l'acte d'enseignement, ces affirmations peuvent avoir une dimension d'évidence et donner le sentiment d'enfoncer les portes ouvertes, cependant le problème est qu'elles sont justement en train de se refermer...