

# Construire des évaluations une mauvaise « bonne idée »

Les tests communs se généralisent à l'école, dès la 3H. Prétendant traiter tous finalement être un leurre de justice, essoufflent de plus en plus d'enseignant·e. D'autres manières de faire existent, et elles ne sont pas plus couteuses...

Raphaël Pasquini, professeur associé - Évaluation formative et certificative des apprentissages scolaires, HEP Vaud

Depuis quelques années, aux degrés primaires et secondaires, nous observons une pratique de plus en plus courante: la construction d'épreuves sommatives en équipe, appelées dans le canton de Vaud «tests communs». Pour élaborer de telles épreuves, les enseignant·e·s d'un cycle ou d'une discipline se répartissent le nombre d'évaluations à mener sur une année ou un semestre. À tour de rôle, seul·e ou accompagné·e de quelques collègues, ils préparent l'épreuve en fonction du programme et du calendrier. Une fois terminé, le test est mis en discussion auprès des collègues à travers des négociations décrites régulièrement comme fastidieuses et stressantes, puis finalisé par les concepteurs·trices. Il peut alors être passé par les élèves à une date souvent prévue (très) à l'avance et identique, à quelques jours près, pour toutes les classes. Parfois, les équipes se retrouvent ensuite pour une correction commune où elles mutualisent leurs questions, incertitudes et tentent d'homogénéiser leurs jugements. Ces pratiques ne sont pas récentes. Réduites par le passé à des situations particulières, elles tendent à se généraliser aujourd'hui alors qu'aucun texte de loi ne les prescrit, probablement pour répondre à un juridisme grandissant de notre société et de l'École, et à un souci de transparence censé garantir le traitement équitable des évalués. Plébiscitées voire parfois imposées par des conseils de direction, soutenues par certain·e·s enseignant·e·s qui y voient notamment un gain de temps, ces pratiques se fondent sur les deux croyances suivantes: les tests communs permettent de traiter tous les élèves de la même manière, et ils garantissent la paix sociale entre le corps enseignant et les parents. Si le second argument peut être entendu lorsque la famille réagit abusivement aux pressions d'une École somme toute sélective, le premier est indéfendable. En effet, l'égalité de traitement ne revient pas à tester un groupe conséquent d'individus au même moment et de la même manière, mais bien à donner à tous les élèves les mêmes enseignements, opportunités d'apprentissage et modalités d'évaluation. Dans cette acception, les épreuves sommatives, quelle que soit leur nature,

doivent donc être en rapport étroit avec ce qui a été enseigné et appris.

Or, avec des tests communs, c'est bien l'inverse qui se produit. En isolant l'évaluation sommative de l'enseignement et de l'apprentissage, et en la rendant semi-standardisée, les enseignant·e·s ne se focalisent plus que sur la pointe de l'iceberg: le «produit» final. Nous observons alors deux phénomènes interpellants qui affectent bon nombre d'élèves.

Le premier montre des enseignant·e·s qui préparent ardemment leurs classes au test commun, prenant conscience de l'écart entre ce qu'ils ont enseigné à leurs élèves et ce que leurs collègues ont prévu d'évaluer. Cet effet est appelé en recherche «*teaching to the test*» (Popham, 2001). Massivement présent avec les épreuves externes, il apparaît malheureusement dans les classes avec les épreuves communes internes. Il montre que lorsqu'on se prépare à un test de ce type, l'enseignement se réduit aux contenus de l'évaluation et a donc pour effet de rétrécir le curriculum. Par ailleurs, le sur-entraînement qui a lieu en classe produit des effets néfastes sur la motivation des élèves et augmente leur anxiété.

Le second phénomène documente une forme de standardisation des exigences, très éloignée des apprentissages effectivement réalisés par les élèves et de leur progression. En effet, les tests communs sont souvent notés à l'aide de barèmes standardisés qui ne tiennent pas compte des contenus abordés ni du temps d'apprentissage. Dans ces situations, le souci des enseignant·e·s est d'avoir des échelles linéaires ou un seuil de suffisance en pourcent défendable, avoisinant le plus souvent les deux tiers des points de l'épreuve. Parfois, il y a même pire: les notes sont posées en référence à la moyenne de la classe, dans une perspective normative. Là, l'évaluation ne porte dès lors plus sur les acquis des élèves, mais sur leur écart à la norme: avec le même score médian, un·e élève pourra réussir dans une classe aux résultats faibles, mais échouer dans une classe où les élèves sont majoritairement «bons», ce qui est contraire aux finalités de l'évaluation scolaire.

# sommatives communes: ?

les élèves de la même manière et rassurer ainsi les parents, ils s'avèrent e-s et soustraient à l'évaluation son rapport avec l'apprentissage.

Notons que ces deux phénomènes de standardisation des barèmes et de normativité peuvent se cumuler. Il devient alors bien délicat de donner le moindre sens à la note obtenue – pour l'élève et pour l'enseignant-e (Crahay, 2007).

Si l'on tient malgré tout à conserver les tests communs, notre expérience en formation continue des enseignant-e-s montre qu'il est possible de changer de perspective tout en convergeant vers une plus grande égalité de traitement, et en valorisant l'apprentissage (Pasquini, 2017). Voici une proposition qui a été mise en œuvre avec des équipes au primaire et au secondaire dans diverses disciplines.

Tout d'abord, le groupe d'enseignant-e-s se rencontre en début de semestre ou de séquence pour co-construire son déroulement: il s'accorde sur les objectifs du plan d'études, les tâches emblématiques de la discipline, les situations incontournables à travailler. Les critères importants émergent alors. Ils sont communiqués aux élèves tout au long de l'enseignement, dans un souci de transparence et de compréhension des attentes (Harlen, 2005). Et c'est seulement ensuite que l'équipe élabore l'évaluation finale, souvent de manière plus économique puisqu'un travail en amont existe déjà. Dans cette perspective, les barèmes sont fixés avant la passation de l'épreuve au regard des critères, en fonction de ce qui a été enseigné et de ce qui est attendu en termes d'exigences. Les barèmes peuvent donc varier en fonction de la situation d'évaluation, de la discipline, de la complexité de l'apprentissage, du temps d'enseignement. L'évaluation est ici toujours commune, mais elle s'inscrit davantage dans une logique de cohérence avec les processus d'enseignement et d'apprentissage, comme ce devrait toujours être le cas (Anderson, 2002). À ce jour, nous observons que la démarche ne provoque pas plus de remous dans les familles, développe les compétences évaluatives et collaboratives des enseignant-e-s. Elle prend aussi tout son sens dans des contextes où les tests communs n'ont pas voix au chapitre, ce que les enseignant-e-s reconnaissent et apprécient.

Avec l'indispensable soutien des directions d'établissement, le changement est donc possible, pour favoriser un travail d'équipe centré sur les apprentissages des élèves. Ce changement serait d'autant plus souhaitable que de plus en plus d'enseignant-e-s se plaignent des tests communs isolés, conscients que les informations qu'ils délivrent sur les apprentissages des élèves sont très peu significatives. Au lieu de courir derrière leur agenda d'évaluation (une fois qu'un test est passé, il s'agit de penser au suivant ou d'y préparer ses élèves), ils demandent à juste titre d'enseigner davantage, de mieux faire apprendre, et de collaborer de manière constructive. Dans une école où le parcours des élèves dépend uniquement des notes, il devient urgent d'écouter ces revendications qui demandent à penser différemment la cohérence de l'évaluation sommative. Pour permettre aux enseignant-e-s de reprendre la main sur leur profession, et pour traiter les élèves comme ils le méritent et comme la loi l'exige. •

## Pour aller plus loin

- Anderson, L. W. (2002). «Curricular Alignment: A Re-Examination». *Theory into Practice*, 41 (4), 255-260.
- Crahay, M. (2007, 3e Ed.). *Peut-on lutter contre l'échec scolaire?* Bruxelles: De Boeck.
- Harlen, W. (2005). «Teacher's summative practices and assessment for learning – tensions and synergies». *The Curriculum Journal*, (16), 207-223.
- Pasquini, R. (2017a). *Évaluation: première question vive de la pratique*. *Educateur*, 4, 38-39.
- Pasquini, R. (2017b). *Évaluation: deuxième question vive de la pratique*. *Educateur*, 5, 36-37.
- Pasquini, R. (2017c). *Évaluation: troisième question vive de la pratique*. *Educateur*, 6, 34-35.
- Pasquini, R. (2017d). *Évaluation: quatrième question vive de la pratique*. *Educateur*, 7, 34-35.
- Pasquini, R. (2017e). *En guise de conclusion provisoire: quelques constats pour une évaluation sommative cohérent*. *Educateur*, 8, 38-39.
- Popham, W. J. (2001). «Teaching to the test?» *Educational Leadership*, 58 (6), 16-20.