

De nombreux enseignants, confrontés à des questions multiples et complexes, demandent aujourd'hui des textes pour penser leurs pratiques d'évaluation sommative. Cette chronique poursuit ce but. Quatre questions récurrentes issues du terrain sont traitées tour à tour. Chacune d'entre elles donne à voir la collaboration entre un formateur et des enseignants travaillant au développement de leurs pratiques évaluatives sommatives, à l'aide d'outils d'analyse.

Une synthèse clôturera la réflexion. Les approches pour penser cette problématique étant variées, ces chroniques visent moins à dire comment évaluer qu'à donner quelques pistes pour réfléchir à ses pratiques, dans une perspective critique et compréhensive.

Troisième question vive de la pratique

Raphaël Pasquini, UER EN, HEP Vaud – raphael.pasquini@hepl.ch

J'utilise des critères mais cela me demande un temps de conception considérable: critères, points, quels outils choisir et quand?

Dans les deux premières chroniques, nous avons pointé l'importance, en évaluation sommative, d'explicitier les objectifs évalués et de les mettre en cohérence avec des tâches évaluatives, du point de vue des habiletés cognitives et des contenus. Mais quels outils d'évaluation exploiter pour rendre visible la valeur, le «poids» de l'apprentissage évalué: des points, des critères, les deux? Quelles sont les forces et les faiblesses de chacun d'entre eux? Passons par l'exemple.

La situation se déroule dans une classe de 10e. Des élèves vont être évalués sur la présentation d'un exposé. L'enseignante leur donne les critères d'évaluation et les points qui y sont associés sur un document. On peut y lire:

- Les élèves préparent un dossier sur la thématique choisie (à montrer en début d'exposé): 2 pts.
- La construction de l'exposé permet de repérer une progression. Présence d'un fil rouge: 2 pts.
- Utilisation d'un vocabulaire riche et cohérent: 1 pt.
- Les élèves se partagent la présentation qui sera fluide et claire. Le ton est vivant. Le rythme permet aux auditeurs de suivre le propos: 2 pts.
- Un support de photos et d'extraits musicaux: 2 pts.
- Les élèves démontrent leur capacité à répondre aux questions à la fin de la présentation: 1 pt.

Un élève avance devant le groupe. Il raconte succinctement et avec enthousiasme le portrait d'un rappeur

qu'il a réalisé, il passe un morceau de musique sur lequel il danse, visiblement heureux. Puis, il répond correctement à quelques questions simples de camarades. Il obtiendra 8 points sur 10 et une note de 5.

C'est donc de musique dont il s'agit. Et pourtant. Lorsque des enseignant-e-s en formation doivent identifier la discipline sur laquelle porte cette évaluation sur la seule base des critères, ils proposent le français, l'histoire, la géographie, l'économie et parfois la musique. À travers cet exercice, nous constatons ensemble le flou qui règne autour de cette grille, alors que, paradoxalement, le rôle des critères devrait notamment être de clarifier les dimensions disciplinaires évaluées. Malaise. Mais ce n'est pas tout... Nous poursuivons l'exercice en analysant l'attribution des points. Les enseignant-e-s constatent, par exemple, que la préparation du dossier vaut autant que l'utilisation de supports ou que la fluidité du ton lors de la présentation. Ils relèvent également que des points accordés au support écrit en côtoient d'autres donnés à l'oralité. Or, nous observons que ces différents éléments de la production ne se valent pas, ni en termes de contenus, ni en termes de complexité de l'apprentissage. Mais, plus interpellant, nous relevons l'absence d'objets de savoir musicaux, le lien fort distendu au Plan d'études romand et la facilité pour l'élève d'obtenir une très bonne note sans faire preuve d'apprentissages significatifs¹. Le malaise devient très gênant.

Cette situation révèle deux problèmes dont font état beaucoup d'enseignant-e-s: rendre transparente l'at-

tribution des points et formuler des critères d'évaluation. Reprenons-les dans l'ordre.

Attribuer des points est la manière la plus courante de concrétiser le «poids» donné à l'apprentissage évalué (Marzano, 2002). Aux dires des enseignant-e-s, c'est la plus simple, la plus rapide. Ils ont raison. Mais malgré ces avantages, le point possède au moins une limite importante: il ne peut rendre compte d'habiletés complexes de manière univoque (combien de points peut valoir la production d'une hypothèse historique par exemple?).

Le travail que nous menons alors avec les enseignant-e-s vise à utiliser les points relativement à des tâches simples (des items). L'exemple le plus explicite peut-être est celui de la question appelant une réponse courte. Lorsque l'on demande à un élève d'identifier le thème d'un texte, il semble évident d'attribuer un point à la question, car la réponse est soit correcte, soit incorrecte. Pourtant, il n'est pas rare que de telles questions soient plus fortement pondérées. Dans le même souci, d'autres problèmes de pondération sont relevés avec les enseignant-e-s: des questions de complexité différente sont pondérées à l'aide de la même quantité de points; ou des questions complexes sont moins fortement pondérées que des questions simples. Un des moyens de faire correspondre le nombre de points est alors de référer chaque question, chaque item, à l'habileté cognitive et au contenu qui lui correspond. Là, toujours avec les enseignant-e-s, nous convoquons à nouveau la taxonomie dont nous avons parlé le mois dernier. Ainsi, sur le même contenu, une question de classement pourrait comporter le même nombre de points qu'une autre portant sur la comparaison, les deux traitant de compréhension.

Ce travail, important lorsque l'on recherche une validité dans l'évaluation sommative, est un premier pas, mais il n'est pas encore à même de répondre aux questions de la pratique relatives aux critères. En effet, bon nombre d'enseignant-e-s les utilisent mais disent avoir des difficultés à les produire (Mottier Lopez & Pasquini, 2017). Lorsqu'on leur demande quand ils ont recours à des critères d'évaluation, voici leurs réponses: d'abord l'habitude, la tradition: «on a toujours utilisé des critères lorsqu'il s'agit d'évaluer des productions écrites ou des exposés.» Puis émerge souvent une réflexion intéressante: «la situation nous oblige à passer par des critères.» Nous essayons de clarifier ce que cela sous-tend. Ce qui ressort, c'est qu'il existe des situations d'évaluation sommative, toutes disciplines confondues, qui nécessitent, de par leur complexité, l'utilisation de critères dont le but est de donner à voir les différentes dimensions de la production. Ces critères se réfèrent aux objectifs et aux tâches complexes, sont peu nombreux, ne se recoupent pas, mettent l'accent sur des apprentissages clés enseignés.

En concevant des évaluations critériées en référence à leur pratique et à ces quelques principes, les enseignant-e-s constatent qu'il devient plus aisé de leur assigner des points. Nous mettons en évidence qu'à ce titre, cette phase de l'élaboration de l'épreuve est aussi fondamentale que les autres, et que mettre des points,

dans cette perspective, est tout sauf un exercice mécanique. Mais les équipes relèvent à raison que c'est un travail chronophage, exigeant, puisqu'il demande sans cesse de se référer aux objets de savoir évalués, mais indispensable si l'on souhaite situer les élèves par rapport aux objectifs visés.

Alors, des points, des critères, les deux? Vu de la pratique, dans un premier temps, il semble fondamental d'analyser la complexité de chaque objectif, de chaque tâche pour voir s'ils nécessitent des critères. Si tel n'est pas le cas, on sera attentif à la relation entre chaque objectif, chaque item et le nombre de points que l'on va lui associer. Si la complexité des apprentissages évalués appelle des critères, on se centrera sur leurs caractéristiques, et ensuite sur les points qu'ils supportent chacun, toujours au regard de leur complexité².

Dans notre situation musicale, ce travail a permis d'explicitier l'objectif évalué: comparer et analyser des œuvres artistiques en analysant le sujet, le thème, la technique, la forme et le message d'une œuvre; en exerçant une démarche critique face aux phénomènes culturels actuels, en recourant à un vocabulaire spécifique. Autant de dimensions permettant de formuler des critères en rapport avec des contenus disciplinaires et des apprentissages ciblés. Un travail qui nous amènera, le mois prochain, à aborder l'épineuse question de la construction de la note.

¹ Nous retrouvons là l'«effet du hibou» décrit dans la première chronique Educateur 4/2017.

² Il est également possible de construire la note sur la seule base des critères. Mais cette pratique demande une plus grande explicitation.

Pour approfondir

Marzano, R. J. (2002). «A comparison of Selected Methods of Scoring Classroom Assessments». *Applied Measurement in Education*, 15 (3), 249-267.

Mottier Lopez, L. & Pasquini, R. (2017). «Professionnal controversies between teachers about their summative assessment practices». *Assessment in Education: Principles, Policy & Practice*, 24 (2), 228-249.