

Évaluer avec le numérique

Partie 1 – De quoi parle-t-on ?

Caroline Ladage¹

Introduction

L'évaluation est un moment clé des processus d'apprentissage et elle compte pour une part importante dans la charge de travail d'un enseignant.

L'évaluation des apprentissages (en tout ou en partie) peut être facilité par le numérique. Évaluer avec le numérique c'est réaliser l'activité d'évaluation des apprentissages, en tout ou en partie, avec le numérique, en s'appuyant pour cela sur des instruments électroniques, informatiques et numériques (machines, ordinateurs, logiciels, applications, environnements informatiques...). Cette évaluation peut prendre de multiples formes, telles que l'enregistrement sonore, le questionnaire à choix multiples, la rédaction d'un texte, la réalisation d'un film, la participation à un forum, la réalisation d'une carte conceptuelle ou d'un portfolio, et bien d'autres formes encore dont Marie-Joëlle Ramage vous présentera un échantillon dans la troisième partie de ce cours.

Aujourd'hui les outils numériques utiles à l'évaluation deviennent de plus en plus faciles à utiliser et sont largement diffusés. Nous avons vu dans les autres cours de ce MOOC que l'évolution des technologies entraîne l'évolution des approches pédagogiques et des activités éducatives. Avec elles les pratiques d'évaluation évoluent aussi. On assiste en effet aujourd'hui à une multiplication des formes et des pratiques d'évaluation, des plus simples aux plus complexes.

1. Enjeux & attentes

Partant du constat que le progrès des technologies facilite les apprentissages et les évaluations des apprentissages, peut-on dire que le numérique change les pratiques d'évaluation ? Qu'il les diversifie ? Qu'il permet d'innover en la matière, d'enrichir ces pratiques, voire de les automatiser ? Peut-on dire qu'il permet d'élargir et de massifier ses publics, tout en individualisant les parcours d'apprentissage ? On est tenté de répondre par l'affirmative à l'ensemble de ces questions, tout en se gardant de voir dans le numérique la solution à tous les problèmes de l'évaluation.

Le numérique permet-il au moins en partie de répondre aux attentes d'une « bonne évaluation », dans une variété de contextes et auprès d'une variété de publics d'évaluateurs et d'évalués ?

Voilà donc, en guise d'introduction, quelques questions à considérer, quelques enjeux reconnus de longue date, qu'ils soient enjeux économiques, sociaux ou pédagogiques.

¹ Maître de conférences à Aix-Marseille Université, Faculté des Arts, Lettres, Langues & Sciences humaines, Département des sciences de l'éducation. EA 4671 Apprentissage, Didactique, Évaluation, Formation (ADEF), ENS de Lyon, IFE.

2. Comprendre les situations d'évaluation²

Comment se repérer dans la multiplicité des outils et des situations d'évaluation ? Notre objectif n'est pas de dresser un tableau exhaustif d'outils. En premier lieu, il est important de comprendre les effets de l'instrumentation des activités d'évaluation par le numérique. Pour cela, il convient de repérer quelques « modèles de l'évaluation » dans le paysage aujourd'hui très riche des cadres théoriques présentés – par des chercheurs d'horizons scientifiques divers – pour penser l'évaluation. Nous proposons d'utiliser les modèles de l'évaluation comme outils pour mieux comprendre les phénomènes d'évaluation mais aussi pour concevoir, voire inventer des pratiques nouvelles. Toute personne désireuse de s'engager dans des projets d'évaluation avec le numérique, il est indispensable de se former aux modèles de l'évaluation afin d'explorer les possibles d'une évaluation s'appuyant sur une instrumentation qui peut être à la fois électronique, informatique et numérique (distinctions précisées par Éric Bruillard dans la première partie de ce cours).

L'évaluation traduit la volonté de rendre compte, d'expliquer, de comprendre une pratique, un apprentissage. L'acte d'évaluation institue une relation éducative à partir du moment où la situation d'évaluation inclut au moins une personne qui évalue et une personne dont une ou plusieurs pratiques sont évaluées. La nature de cette relation est, chacun de nous le sait, souvent très complexe, parfois conflictuelle et souvent chargée d'émotions. Ce que la théorie va mettre en évidence, entre autres, c'est que celui ou celle qui évalue s'appuie sur un système de références, sur des connaissances et des théorisations, sur un vécu aussi, qui peuvent être de tous ordres, de nature personnelle ou institutionnelle, qui vont influencer sur le choix du dispositif ainsi que sur la nature et la qualité de tous les moments de l'évaluation, de sa programmation à son résultat en passant par son déroulement.

Repérer et choisir le modèle pertinent pour un projet d'évaluation constitue donc un moment clé de la relation éducative. L'évaluation ne peut être réduite à un dispositif de mesure ni linéaire ni univoque et il ne s'agit pas d'attendre d'un dispositif ou d'un outil qu'il serve à tout, ou qu'il évalue de manière constante quelles que soient les caractéristiques et quel que soit le contexte de la situation d'évaluation. Se former aux modèles de l'évaluation conduit donc à explorer les possibles d'une évaluation qui ne soit pas seulement contrôle, mais qui, tout en incluant souvent une mesure, va au-delà pour s'efforcer de rencontrer toute la complexité inhérente aux situations à évaluer.

3. Les modèles de l'évaluation

L'évaluation est un domaine majeur de la recherche en éducation et en formation ainsi que dans le champ des apprentissages. Historiquement, beaucoup de théorisations et de modèles ont été proposés par un nombre important de chercheurs issus de différentes disciplines (psychologie, sciences de l'éducation, didactique, ergologie, technologies de l'information et de la communication, sociologie, etc.). Nous ne proposons ici qu'un bref aperçu de certains de ces « modèles de l'évaluation » bien connus, qui paraissent particulièrement adéquats pour étudier et questionner les situations d'évaluation avec le numérique.

Le premier modèle qui vient à l'esprit aussi bien d'un point de vue historique que du point de vue de sa diffusion large dans les pratiques éducatives et sociales, c'est le modèle de *l'évaluation comme mesure*. Historiquement, le modèle de la mesure est présenté comme le premier cadre théorique pour penser les pratiques d'évaluation des apprentissages. Nous verrons dans la deuxième partie de ce cours comment les modèles se sont diversifiés et ont évolué en cherchant à tenir compte de la diversité des situations et des problématiques d'évaluation.

² Pour la rédaction de ces parties nous nous sommes appuyée sur les textes de Jean-Jacques Bonniol, Gérard Figari, Lucie Mottier Lopez et Michel Vial, dont les références sont reproduites à la fin de ce texte introductif.

4. Un modèle ancestral et prégnant : l'évaluation comme mesure³

C'est là le point de départ, dans la première moitié du XX^e siècle : jusque dans les années 1930 l'évaluation des apprentissages est pensée principalement en termes de mesure. Évaluation et mesure tendent à se confondre. On mesure les résultats d'un apprentissage et les progressions d'un apprentissage. On situe sur une échelle de mesure, on compare, on classe, on vérifie la conformité, on cherche des indicateurs.

a) La question de la qualité de la notation. L'usage de la mesure en éducation est apparu au début du XX^e siècle avec l'introduction de la note pour signifier les réussites ou les échecs des apprentissages. Dans un contexte de massification et de démocratisation de l'accès à la l'éducation apparaît le souci d'assurer une qualité à la notation scolaire : la note doit être objective. Or on découvre que de nombreux biais pèsent sur les notes des enseignants. C'est la « docimologie » – terme proposée par le psychologue Henri Piéron dans les années 1920 – qui va développer des savoirs nouveaux sur la subjectivité en matière de notation.

<p>Docimologie</p> <p>d'après le <i>Trésor de la langue française informatisée</i></p>	<p>PÉDAG. Science des examens et des concours, étude de la qualité et de la validité des différents systèmes de notation scolaire et de contrôle des connaissances. <i>Dans l'enseignement secondaire, la docimologie intéresse plus particulièrement le système de notation et le régime des compositions, qui constituent le mode traditionnel de contrôle des connaissances et de classement des élèves (Pédag. 1972).</i></p> <p>Étymol. et Hist. 1960 (G. PIRE, <i>La Notation à l'école. Qu'est-ce que la docimologie ?</i> École nouvelle fr.). Composé du gr. δοκιμή « épreuve » et de l'élément suff. -logie*. Terme proposé par le psychologue fr. H. Piéron apr. 1922 (cf. H. PIÉRON, <i>Examens et docimologie</i>, 1963 ds GILB. : J'ai proposé le terme <i>docimologie</i> [...] j'entrepris la première recherche de <i>docimologie</i> [en] juin 1922).</p>
---	--

Les questions qui se posaient alors portaient sur la difficulté de déterminer les qualités requises d'un bon instrument de recueil d'information sur les apprentissages et d'identifier les sources potentielles d'erreur de mesure. Comme le souligne Lucie Mottier Lopez (2013), « les notations des enseignants étant alors comparées à une mesure ».

Ce cadre de référence va aussi déboucher sur une science de la mesure appelée **édumétrie**. On y porte l'attention sur des points de différenciation pour centrer la *métrie* sur certaines spécificités de la mesure en éducation (Mottier Lopez, 2013) :

- les élèves et leurs traits distinctifs ;
- les objectifs et les domaines d'enseignements ;
- les conditions d'apprentissage ;
- les niveaux successifs d'un apprentissage.

b) Un outil emblématique de l'évaluation comme mesure : le QCM. Il existe un très grand nombre de types de mesure qui constituent autant d'objets de l'évaluation. L'outil le plus emblématique de la mesure en éducation est le QCM, le *questionnaire à choix multiples*. Notons qu'aujourd'hui l'évaluation avec le numérique est souvent associée spontanément aux QCM et aux tests automatisés. Son principe : proposer plusieurs réponses pour chacune des questions proposées (une ou plusieurs propositions sont correctes, les autres sont erronées – ce sont les « distracteurs »). Il existe un grand nombre de variantes de tests de ce type (dont parfois seul le nom varie pour présenter un même type de structure), tels que les tests automatisés, les *quiz*, les exercices. Avec le développement des

³Nous suivons dans la présentation historique Lucie Mottier Lopez (2013) dans son article « De la mesure à l'évaluation collaborative en éducation ».

technologies, ces tests se sont automatisés, la machine et le traitement informatique vérifient si la réponse donnée est juste, un feedback peut être donné et l'apprenant peut être orienté vers des questions spécifiques en fonction de ses réponses.

Le numérique semble avoir libéré le QCM d'un certain nombre de biais et de limites qui lui étaient reprochés de longue date. Non seulement avec le numérique on assiste à une multiplicité de formes et de fonctionnalités, mais il apparaît aussi comme un outil d'auto-évaluation (autocontrôle) inédit et souvent préféré à l'évaluation humaine. Comme l'a suggéré Cédric Fluckiger dans le cours sur les théories de l'apprentissage, le numérique, avec QCM et *quiz* automatisés de tous genres, joue ce rôle du tiers qui ne se fatigue jamais, qui n'est pas perçu comme menaçant et dont on ne craint pas les jugements. Ajoutons à cela que le numérique offre des situations d'évaluation là où il n'y en avait guère auparavant, comme par exemple dans les situations d'autodidaxie.

c) Des outils de mesure performants et complexes. L'informatique, l'électronique et le numérique contribuent à une sophistication de l'instrumentation des techniques de mesure. Ils permettent l'introduction de situations de mesure inédites, dont nous citerons trois exemples. Tout d'abord l'analyse des *traces* produites dans les environnements numériques d'apprentissage. Les activités humaines dans les environnements numériques laissent des traces en quantités importantes. L'étude et l'analyse de ces données (par les chercheurs, les concepteurs et les enseignants) ont pour objectifs tantôt d'ajuster l'intervention éducative, tantôt de formaliser les expériences d'enseignement pour mieux comprendre les pratiques quotidiennes des apprenants et le déploiement de l'activité dans l'environnement numérique. Ces traces peuvent aussi servir de matériau pour l'évaluation, qui est basée alors, par exemple, sur la fréquentation et la participation aux discussions sur les forums, sur le degré de consultation des ressources disponibles ou encore sur le parcours réalisé dans un environnement numérique donné. Un deuxième exemple de situation d'évaluation que le numérique a rendu possible est celle présente dans les environnements de simulation. Citons à titre d'exemple les simulateurs de vol en avion ou en hélicoptère ou les simulateurs d'actes médicaux. Enfin un troisième exemple de situation inédite de mesure des activités est celui des jeux sérieux (*serious games*) dans lesquels la mesure peut être présente de façon diversifiée et être intégrée dans des situations d'évaluation développées et complexes. L'évaluation comme mesure prend une place importante dans ces environnements interactifs et changeants qui imitent le monde réel en suivant toutefois le plus souvent des scénarios programmés à l'avance.

Variations autour du QCM

Le numérique permet une variété importante de configurations de situations d'évaluation. On peut noter un regain d'intérêt du modèle de la mesure depuis cette nouvelle instrumentation, qui invite à repenser les activités d'évaluation et à explorer de nouvelles possibilités. Concevoir un QCM ou un test d'évaluation avec le numérique permet en effet de varier les configurations en jouant par exemple sur les fonctionnalités suivantes :

- rétroactions (immédiateté ; personnalisation avec les tests adaptatifs) ;
- présentations multimédias (illustrations, animations, audio, vidéo) ;
- barèmes et systèmes de notation (avec possibilité d'indiquer le degré de certitude de la réponse à la question) ;
- différenciation des temporalités (varier les durées de passation) ;
- ...

Les QCM peuvent figurer dans un dispositif d'évaluation à plusieurs titres :

- tests diagnostiques ;
- enquêtes, sondages ;
- exercices formatifs ;
- auto-évaluations ;
- examens sommatifs ;
- ...

d) Limites du modèle de la mesure. La mesure apparaît incapable de rendre pleinement compte de la complexité des expériences du quotidien et suscite nombre de résistances et de mises en garde. Ainsi peut-on souligner que la mesure elle-même n'est pas l'évaluation : la mesure correspond au résultat, alors que l'évaluation consiste à interpréter ce résultat pour lui attribuer une signification sociale.

Dans une approche non radicale, le fait de différencier mesure et évaluation ne signifie pas exclure la mesure de l'évaluation. L'évaluation a en effet besoin de procédures et d'instruments rigoureux et fiables ; la mesure représente un instrument (une méthode avec ses propres règles et procédures) pouvant être utilisée parmi d'autres instruments possibles au regard des buts et intentions visées par les évaluateurs. Mais tout en pouvant englober la mesure, l'évaluation implique, quant à elle, un ensemble « d'opérations » que la mesure n'inclut pas, dont l'interprétation des résultats au regard d'un référentiel (qu'il s'agit d'établir), la communication des résultats et des interprétations aux différents partenaires concernés, les prises de décision en fonction des buts visés, des contextes sociaux, des moyens à disposition, et plus généralement des contraintes et des possibles. (Mottier Lopez, 2013)

Le numérique peut-il changer la donne ? Le développement du numérique ne pousse-t-il pas à questionner à nouveaux frais la place de la mesure dans le champ de l'évaluation en éducation et en formation ? Avec la diversification et la vulgarisation des outils de mesure suscitées par le numérique, peut-on dire qu'il y a aujourd'hui un changement de perspective ? Voilà quelques questions ouvertes, actuellement à l'étude.

Si on reprend l'exemple du QCM, qui est l'une des formes les plus anciennes et les plus répandues d'outil d'évaluation, les recherches ont pu montrer que sa forme numérique semble bien l'avoir libéré d'un certain nombre de biais et de limites, à condition de bien en saisir les rouages, aussi bien techniques que du point de vue des discours sur ces techniques (les modèles de référence de ceux qui les mettent en œuvre et qui conditionnent le déroulement des situations d'évaluation).

Cette partie introductive s'arrête sur le constat de limites du modèle de l'évaluation comme mesure, qui n'en reste pas moins prégnant à plus d'un titre. Ce modèle initial ouvre vers la prise en considération d'autres modèles de l'évaluation des apprentissages, modèles que nous proposons de découvrir dans la deuxième partie de ce cours.

Bibliographie

Bonniol, J.-J. & Vial, M. (1997). *Les modèles de l'évaluation*. Bruxelles : De Boeck.

Mottier Lopez, L. (2013). De la mesure à l'évaluation collaborative en éducation. *Revue française d'administration publique*. 4(148), 939-952.

Mottier Lopez, L. & Figari, G. (dir.). (2012). *Modélisations de l'évaluation en éducation*. Bruxelles : De Boeck.

Vial, M. (2012). *Se repérer dans les modèles de l'évaluation*. Bruxelles : De Boeck.