

Évaluer et s'autoévaluer avec le numérique : Principes et outils

Philippe Dessus, LaRAC & LIG Inria

UE 404 – Master MEEF PE 2

INSPÉ / UGA - 2019-20



A

B

C

D



Soignez l'évaluation !

“ [...] les effets des mauvaises pratiques **sont bien plus puissants** [dans le cas de l'évaluation] que dans tout autre aspect de l'enseignement. Les élèves peuvent, avec difficulté, **s'échapper des effets d'un mauvais enseignement ; ils ne peuvent (s'ils veulent réussir dans un cours) échapper aux effets d'une mauvaise évaluation.** ” (McDonald et al. 2000, p. 8)

Théorie¹

Rappels : Définition de l'évaluation

L'évaluation en éducation est le processus par lequel **on délimite, obtient, et fournit des informations utiles permettant de juger des décisions possibles**". (Stufflebeam *et al.* 1980, p. 48)

- c'est un processus (dynamique)
- on la prépare (délimite), observe, et recueille des informations et on les synthétise pour les diffuser
- orientée vers des buts précis (on n'évalue pas pour évaluer, mais pour agir)

Types d'évaluation (1/2)

Les types d'évaluation varient en fonction du moment de l'apprentissage, du support, des processus cognitifs scrutés, du contexte, des décisions à prendre

Types d'évaluation (2/2)

”Lorsque le cuisinier goûte la soupe, c'est formatif ; lorsque les invités goûtent la soupe, c'est sommatif” (R. Stake cité par Scriven 1991 p. 169)

- évaluation *de* l'apprentissage (sommatif), **évaluer en testant**, où des inférences sont faites entre l'évaluation de réponse à un test et les compétences de l'élève (on goûte une soupe pour évaluer la qualité du cuisinier)
- évaluation *pour* l'apprentissage (formatif), **évaluer en jugeant**, où la qualité de la production est *directement* évaluée par l'élève, aidé par l'enseignant

Évaluer pour l'apprentissage : les rétroactions

Trois types de rétroactions, informantes pour l'élève (Hattie & Timperley 2007 ; voir aussi Lepareur 2016)

- Où est-ce que je vais ? (quels sont les buts ?)
- Comment est-ce que j'y vais ? (quels sont les progrès que j'ai faits vers ces buts ?)
- Que faire ensuite ? (quelles tâches dois-je réaliser pour mieux y arriver ?)

Évaluer pour l'apprentissage : le processus

Wiliam & Thompson (2007, cités par Lepareur 2016)

1. Clarifier, partager et faire comprendre les **intentions d'apprentissage et les critères de réussite**
2. Organiser de véritables **discussions, activités et tâches** qui produisent des preuves sur les apprentissages
3. Donner un **feedback** qui fait progresser les élèves
4. Inciter les élèves à être des **personnes-ressources** pour leurs pairs
5. Inciter les élèves à **s'approprier leurs apprentissages**

Comment s'y prend-on ? (Pellegrino 2013, p. 380)

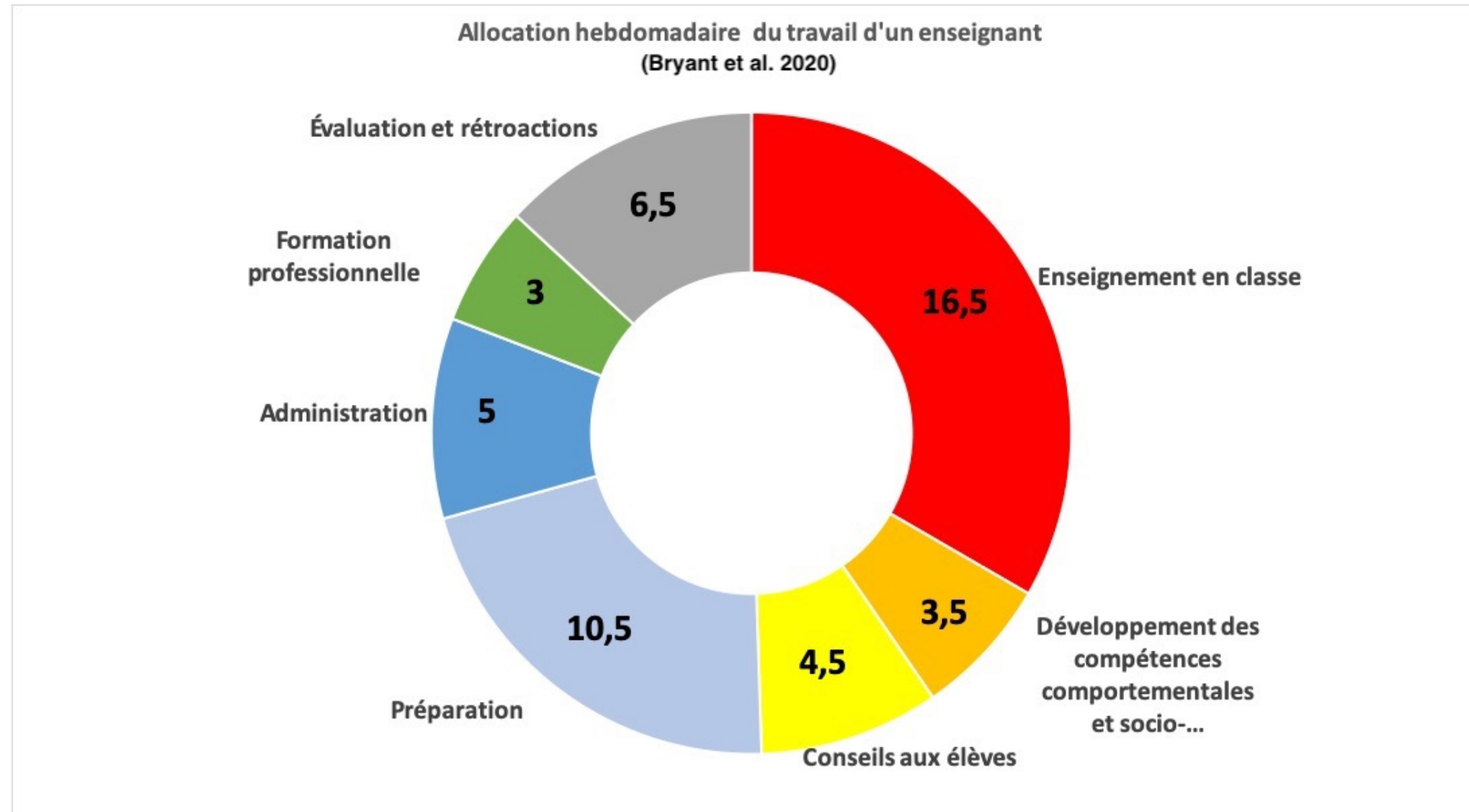
Trois moments interconnectés

- Quelles **connaissances et compétences précises** voulez-vous voir acquises par les élèves et *comment* voulez-vous qui les acquièrent ?
- Quelles **preuves** d'acquisition de ces connaissances et compétences accepterez-vous comme valables ? Comment les analyserez-vous et interpréterez-vous ?
- Quelles **tâches** les élèves *réaliseront* pour communiquer leurs connaissances et compétences ?

L'évaluation dans l'enseignement (1/3)

Un enseignant travaille, en moyenne, environ 50 h par semaine (Bryant *et al.* 2020)
La moitié de ce temps se passe en présence directe des élèves

L'évaluation dans l'enseignement (2/3)



L'évaluation dans l'enseignement (3/3)

Le numérique permet de réallouer 20-30 % du temps hebdomadaire (Bryant et al. 2020) :

- préparation : gain 5 h
- **évaluation/rétroactions : gain 3 h**
- administration : gain 2,5 h
- enseignement en classe : gain 2 h
- formation professionnelle : gain 0,5 h

Numérique vs. papier-crayon ? 👍

- peut recueillir des **processus** (durées, hésitations, etc.) analysables par la suite
- traitement **automatique**, donc moins de "fatigue évaluative" côté enseignant, plus d'entraînement côté élèves
- attribue des niveaux de manière plus **fiable et rapide**, reposant sur des statistiques
- peut proposer des **rétroactions** (semi-)automatiquement juste après la production
- offre des **visualisations synthétiques** pour suivre les progrès des élèves
- s'adapte aisément **au nombre** (2 fois plus d'évaluations < 2 fois plus de travail)
- peut être considéré comme **moins sujet aux stéréotypes** ou aux effets d'ancrage

Numérique vs. papier-crayon 🖐️

- données plus aisément **piratables, modifiables**
- difficile de ne recourir qu'au libre et gratuit, donc **aspiration de données personnelles** difficile à contourner
- **plus de données** ≠ meilleure évaluation

⚠ Attention aux fausses bonnes idées !

- fausse impression de **personnalisation** : les parcours automatisés peuvent enfermer l'élève dans des "profils"
- fausse impression d'**objectivité** : *rubish in => rubish out* (si on scrute les mauvaises choses, on obtient une mauvaise évaluation)
- fausse impression d'**exhaustivité** : avoir beaucoup de données sur l'apprentissage nécessite beaucoup de temps pour l'analyser

Prati²que

Quelques types de contextes d'évaluation

- Que produit-on ?
 - productions **contraintes** (questionnaires), plus aisées à analyser automatiquement
 - productions **libres**, plus difficiles à analyser
- Qui évalue ?
 - l'élève (auto-évaluation), les pairs, l'enseignant

Choix des outils en fonction de la tâche

La nature de la tâche influe sur le choix de l'outil (Jonassen 2014). Différence entre

- **Problèmes bien structurés** (on connaît bien l'état initial, l'état final, et les étapes pour passer de l'un à l'autre)
 - évaluer les réponses correctes
 - classer les problèmes et les proposer
- **Problèmes mal structurés** (les autres)
 - simulations
 - environnements libres, jeux sérieux, portfolios
- **Observer les élèves**
 - Réaliser des schèmes de codage

Choix des outils : Protection des données

Le critère de la gratuité ne peut être le seul critère à prendre en compte. Comme on stocke des données personnelles, le respect du RGPD (Règlement général sur la protection des données) est essentiel. Nécessité

- de montrer que le traitement réalisé est **nécessaire** à la mission de l'enseignant
- l'inscrire sur le registre des activités de traitement (voir DASEN)
- d'avoir l'**autorisation parentale**, mentionner les droits d'opposition, accès, rectification, effacement, etc.
- s'assurer que les **conditions de service** du site utilisé permettent de garder la maîtrise des données (pas d'autre utilisation)
- de garanties en termes de **sécurité**

📖 Bouteloup 2020 [Logiciels éducatifs, aspirateurs à données personnelles ?](#)

📖 Canopé 2018 [Les données à caractère personnel](#)

Évaluation de productions contraintes (questionnaires)

Il existe une large palette d'outils, déjà réalisés ou à réaliser, permettant l'évaluation *via* questionnaires

- en ligne : [Learning Apps](#)
- en local, pouvant ensuite être mis en ligne : [eXe Learning](#)
-  Dessus (2015). [Ressources : générateurs de cours interactifs](#)

Évaluation de productions libres

Notamment utiles pour analyser l'apprentissage dans des tâches non structurées

- **Productions textuelles** : traitement automatique de la langue, textométrie,
 - [Anatext d'O. Kraif](#) pour repérer les patterns de mots dans les textes,
 - [Anagraph de J. Riou](#) pour évaluer la part de déchiffrable dans les textes à lire
 -  (voir Rinck & Dessus, 2020, [la textométrie pour évaluer les productions des élèves](#) pour plus d'informations
- **Productions orales ou vidéo** : outils d'annotation (e.g., [ANVIL](#)), nécessitent un codage humain

Lecture & Maths

- **TACIT** : test adaptatif de compréhension en lecture, évaluant précisément le niveau des élèves
- **Calcul@TICE** : banque d'exercices de calcul mental

Autres outils évaluatifs

- Recommandations de lectures : [Babelio Jeunesse](#)
- Pour observer la classe : [LessonNote](#)-iPad
- Pour analyser les relations entre élèves : [Sociometry made simple](#)
- Pour suivre l'acquisition de compétences/connaissances : [Sacoche](#)
- Portfolio : tout outil de blog/création multimédia peut être utilisable, comme [Didapages](#)
- Évaluation par les pairs : [PeerGrade](#)

Le futur de l'évaluation ? (1/2)

L'intelligence artificielle ?

⚠ La plupart du temps, un **simple argument commercial**, pas une assurance d'évaluation fiable, pertinente

📖 De Lièvre et al. 2019

Le futur de l'évaluation : 5 principes (Jisc 2020)

- **authenticité** (évaluations de compétences de la vie réelle et non jetables)
- **accessibilité** (évaluations pour tous, accessible, utilisant les normes)
- **raisonnablement automatisé** (les classes sont *aussi* peuplées d'humains)
- **continu** (ne pas se fier qu'aux "contrôles finaux" pour évaluer les élèves)
- **sécurisé** (à la fois du côté des élèves — leur travail leur appartient — que des enseignants — tricherie)

Résumé

Le numérique peut aider :

- à créer des **"effets de test"**, solides déclencheurs d'apprentissage
- à évaluer des **activités, comportements, attitudes très divers**
- sous des **formats divers** (texte, image, vidéo)

⚠ Les classes sont peuplées **d'humains** (enseignants/pairs), donc **ne pas laisser tout le processus évaluatif au numérique**

Merci de votre attention !
des questions ?



Références

- Bryant, J., Heitz, C., Sanghvi, S., & Wagle, D. (2020). *How artificial intelligence will impact K-12 teachers*. New York: McKinsey & Company
- de Lièvre, B., Temperman, G., & Di Emidio, S. (2019). *Intelligence artificielle et éducation : quelle complémentarité ?* *Échos de Picardie, Automne 48–55.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. doi:10.3102/003465430298487
- Jisc. (2020). *The future of assessment: Five principles, five targets for 2025. Bristol: Jisc.
- Jonassen, D. H. (2014). *Assessing problem solving*. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology (4th ed., pp. 269–288). New York: Springer.
- Lepareur, C. (2016). *L'évaluation dans les enseignements scientifiques fondés sur l'investigation : Effets de différentes modalités d'évaluation formative sur l'autorégulation des apprentissages*. (Doctorat de sciences de l'éducation). Univ. Grenoble Alpes, Grenoble.
- McDonald, R., Boud, D., Francis, J., & Gonczi, A. (2000). New perspectives on assessment. *Studies in Technical and Vocational Education*, 4.
- Pellegrino, J. W. (2013). Measuring what matters. Technology and the design of assessments that support learning. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 377–387). New York: Routledge.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). London: SAGE.
- Stufflebeam, D. L., Foley, W. J., Gephart, W. J., Guba, E. G., Hammond, R. L., Merriman, H. O., & Provus, M. M. (1980). *L'évaluation en éducation et la prise de décision*. Ottawa: N.H.P.
- William, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah: Erlbaum.

Réfs

{Jonassen, 2014 #22115}
{McDonald, 2000 #17115}
{Stufflebeam, 1980 #19937}
{Pellegrino, 2013 #22116}
{Scriven, 1991 #22117}
{de Lièvre, 2019 #22118}
{Jisc, 2020 #22107}
{Stufflebeam, 1980 #19937}
{Hattie, 2007 #15067}
{Lepareur, 2016 #22119}
{Wiliam, 2007 #22120}
{Bryant, 2020 #22121}