

Grand angle



L'évaluation en enseignement supérieur : mesure des apprentissages ou épreuve de sélection ?

En enseignement supérieur, les notes ont un grand impact sur le parcours d'une personne étudiante : elles ouvrent et ferment des portes dans le monde académique et professionnel, et se répercutent sur l'estime de soi, le bien-être et les aspirations étudiantes.

Avoir de bonnes notes, mais surtout avoir de meilleures notes que ses pairs, est une condition nécessaire pour se démarquer dans les processus de sélection. Dans ce contexte de concurrence, l'évaluation, surtout lorsqu'elle prend la forme d'un examen sommatif à enjeu élevé, est souvent perçue comme une épreuve stressante et sans appel. Réduite à la dimension de la notation, l'évaluation est « la plupart du temps considérée comme néfaste pour l'apprentissage et la motivation » (Yerly et Berger, 2022, p. 7).

Dans ce grand angle, l'ORES explore quelques questions liées à l'évaluation normative en enseignement supérieur

- D'où vient l'évaluation normative en éducation et comment expliquer sa persistance ?
- Quelles sont les répercussions de l'évaluation normative sur les personnes étudiantes et leur réussite ?



1 Qu'est-ce que l'évaluation normative ?

Depuis des siècles, les systèmes scolaires occidentaux utilisent la notation pour catégoriser et comparer les personnes étudiantes. Au XXe siècle, avec la massification de l'éducation, l'école s'est vue octroyer un nouveau rôle : déterminer le destin social des individus, en les triant selon leur « mérite », leurs « aptitudes » ou leurs « efforts » (Romainville, 2021, p. 90).

L'évaluation normative constitue le mécanisme pour opérer ce tri : les personnes qui performant le mieux sont sélectionnées pour accéder aux disciplines ou aux programmes les plus prestigieux ou contingentés, les plus « faibles » sont éjectées du système ou reléguées à des trajectoires moins valorisées. Autrement dit, l'évaluation normative a pour fonction la hiérarchisation des personnes étudiantes à des fins de sélection.

Évaluation normative : « Mode d'évaluation où la performance d'un sujet est comparée (...) à celle des autres personnes d'un groupe de référence d'après un même instrument » (Legendre 2005, cité dans (CSE, 2018, p. 8))

En rupture avec cette approche normative, les spécialistes en sciences de l'éducation prônent depuis des décennies l'adoption de **pratiques d'évaluation centrées sur l'apprentissage**. Dans cette vision, qui est au cœur de l'approche par compétences, **l'évaluation est une appréciation qualitative et critériée, conçue pour soutenir la personne étudiante dans son processus d'apprentissage** (Côté, 2017). Elle mesure l'atteinte des cibles d'apprentissage, plutôt que de comparer les personnes étudiantes entre elles.



Malgré le consensus scientifique en faveur de l'évaluation centrée sur l'apprentissage, les pratiques d'évaluation normative demeurent très présentes en enseignement supérieur.

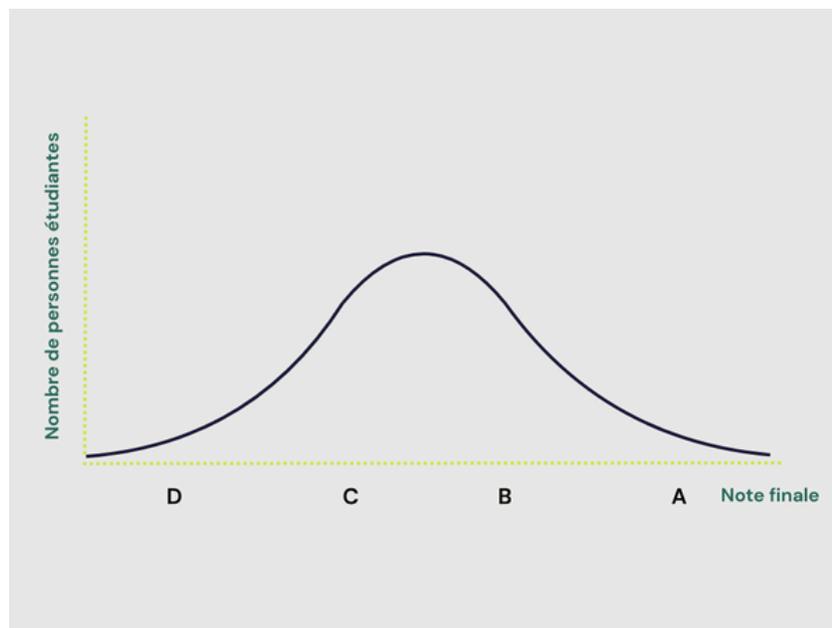
Pourquoi donc? Sur quelles croyances se fonde la persistance de l'approche normative?



La courbe normale

La courbe normale ou courbe de Gauss (en forme de cloche) est la représentation graphique d'une loi de probabilité binomiale de résultats aléatoire. Par exemple, si l'on jette une poignée de monnaie au sol, la loi permet de prévoir que la proportion de résultats « pile » a beaucoup plus de chances de se situer entre 40% et 60% que d'être de 90% ou de 10%.

Au tournant du XXe siècle, les sciences humaines se sont emparées de cette loi pour modéliser le comportement humain, donnant naissance à la psychométrie. Cette nouvelle science, qui a inventé les tests standardisés et à la notion de QI, postule que, dans la population, les capacités intellectuelles se répartissent sur une courbe normale. Ainsi, dans le champ de l'éducation, le modèle psychométrique est à l'origine de l'évaluation normative, vue comme une façon plus « objective » et « scientifique » de mesurer les performances des étudiantes et étudiants (Bowen et Cooper, 2022; Brookhart et al., 2016).



Concrètement, lorsqu'appliquée au champ de l'évaluation en éducation, cette conception impose une certaine répartition des notes, qui devrait refléter la répartition soi-disant « naturelle » des capacités des étudiantes et des étudiants. Dans une logique normative, « une bonne évaluation devrait comporter beaucoup de notes moyennes et peu de très bonnes et de très mauvaises notes » (Huver et Springer, 2011). Techniquement, cela signifie que la réussite des uns dépend de l'échec des autres.



Un référent invisible

Dans les années 1970, le psychologue Benjamin Bloom critique l'évaluation normative en posant la question suivante : Comment les résultats d'apprentissage pourraient-ils refléter les lois de la probabilité, alors que les conditions d'apprentissage et d'enseignement diffèrent pour chaque personne?

Dans les décennies suivantes, les sciences de l'éducation prennent leurs distances avec l'approche normative, adoptant d'abord une perspective behavioriste héritée de Bloom (paradigme de l'enseignement) et évoluant vers une posture socioconstructiviste aujourd'hui majoritaire (paradigme de l'apprentissage).

Malgré tout, la courbe normale demeure le « partenaire silencieux » du système de notation, qu'il soit en lettres ou en chiffres (Brookhart et al., 2016). L'évaluation critériée est valorisée officiellement, mais les pratiques normatives (comme de comparer les résultats individuels à la moyenne du groupe, ou encore de « normaliser » les notes) persistent dans les pratiques, du primaire à l'université. **En enseignement supérieur, les mécanismes de sélection pour l'accès à des stages ou des bourses prestigieuses participent pleinement à consolider cette approche normative.**



Selon le Conseil supérieur de l'éducation, la logique de sélection scolaire qui prévaut en enseignement supérieur a des effets en cascade sur l'ensemble du système scolaire, contribuant à la « survalorisation de notes et du classement » dès l'école primaire (CSE, 2018, p. 69).

La cote de rendement au collégial (cote R)

La cote de rendement au collégial est le principal critère qui détermine l'accès à certains programmes ou établissements universitaires.

Avant les années 1980, la moyenne collégiale était utilisée par les universités pour classer et sélectionner les étudiants et étudiantes aux fins de l'admission. Par la suite, une formule de calcul plus élaborée incluant la moyenne et l'écart-type (la cote Z) a été utilisée jusqu'en 1995, jusqu'à l'adoption de la cote R par les universités. Cette nouvelle méthode de calcul a été conçue pour corriger les biais de la cote Z, en prenant en compte la force relative du groupe :

« La cote de rendement au collégial combine, pour chaque cours suivi par un étudiant, trois informations : un indicateur de la position de cet étudiant en fonction de la note obtenue dans son groupe (la cote Z au collégial), un indicateur de la force de ce groupe et un indicateur de la dispersion de ce groupe » (BCI, 2025, p. 7).



La cote de rendement au collégial (suite)

Selon le Bureau de coopération universitaire, responsable de la méthode de calcul, la cote R est un « instrument de mesure juste et équitable », qui donne « aux meilleurs étudiants de tous les collèges des chances égales d'accès aux programmes universitaires les plus contingentés » (BCU, 2025, p. 20).

Le Conseil supérieur de l'éducation constate cependant que la cote R nourrit une concurrence parfois malsaine, en plus de pénaliser de façon systématique les étudiantes et les étudiants qui connaissent « un début lent ou difficile » (CSE, 2018, p. 47). Une équipe de recherche a par ailleurs montré que la cote R continue de refléter un biais en fonction de la force du groupe, et « légitime, voire accentue les inégalités sociales d'accès aux programmes contingentés de l'enseignement universitaire » (Moulin et al., 2022).

2

Quelles répercussions sur l'apprentissage et la réussite étudiante?

Collaboration VS compétition

La prédominance de la logique de classement dans l'évaluation accentue la compétition et décourage la collaboration entre les personnes étudiantes. En effet, dans une logique normative où chacun se positionne par rapport aux autres, chaque étudiante ou étudiant qui réussit rend la réussite un peu plus difficile à atteindre pour les autres. Ce climat de compétition est particulièrement exacerbé dans certains programmes d'études prestigieux, par exemple le droit ou le domaine des STIM (Canning et al., 2020; Nyström et al., 2019).

Dans un système éducatif fondé sur la sélection, « *chercher à être meilleur que les autres devient une réponse adaptée face aux contraintes structurelles du milieu académique* » (Świątkowski et Dompnier, 2017, p. 107).

Les résultats de cette « course à obstacles » ne sont qu'en apparence fondés sur le mérite. En réalité, ce système favorise les étudiantes et étudiants issus de milieux privilégiés, qui arrivent mieux préparés au cégep ou à l'université. À l'inverse, il désavantage les personnes appartenant à des groupes sous-représentés (notamment de première génération) qui vivent plus souvent un sentiment d'imposteur en enseignement supérieur, ne détiennent pas le savoir implicite ni ne bénéficient d'un soutien de leur entourage pour savoir comment bien performer dans les évaluations (Bowen et Cooper, 2022, p. 190; Campbell, 2024; Canning et al., 2020).



Anxiété et course aux notes

La culture de la compétition amplifie le stress et l'anxiété liés aux notes, particulièrement chez les personnes appartenant à des groupes historiquement sous-représentés en enseignement supérieur ou dans certaines domaines (ex. les femmes en génie) (Posselt et Lipson, 2016).

Pour contrer les effets négatifs du niveau élevé de compétition sur la santé mentale des personnes étudiantes, plusieurs programmes de médecine, au Canada et ailleurs, ont d'ailleurs remplacé leur système de notation par la mention « succès-échec » (CSE, 2018, p. 51; Nyström et al., 2019).

Le stress des examens : nécessaire?

Une idée répandue veut que les évaluations stressantes fassent partie de l'apprentissage, puisqu'elles sont une préparation aux réalités stressantes de la vie professionnelle. Pourtant, les contraintes typiques des examens à enjeux élevés (par ex. un test d'une durée déterminée dans une salle fermée sans accès à ses notes) ne correspondent pas à la réalité des milieux professionnels, académiques ou sociaux qu'intégreront les personnes étudiantes après leur formation. En ce sens, il ne s'agit pas d'évaluations authentiques.

Selon les membres du comité de travail sur les « cours défis » au collégial, un test comme l'épreuve uniforme de français (EUF), qui consiste à produire une dissertation de 900 mots en 4h30, « n'offre peut-être pas un portrait juste des capacités d'une personne étudiante » en raison du « stress intense » qu'il engendre (MES, 2024, p. 73). La Fédération étudiantes collégiale du Québec qualifie quant à elle l'EUF de contexte d'évaluation « totalement artificiel, anxiogène et dépassé » (FECQ, 2022, p. 21).

Les notes et la motivation

L'évaluation et la notation engendrent des états émotifs très différenciés chez les personnes étudiantes (Stiggins, 2008, p. 240). Pour ceux et celles qui « performant » bien, le système d'évaluation peut agir comme moteur de succès : il génère de la confiance, alimente le désir de performer, de mettre des efforts, etc. Cela dit, les bons résultats peuvent aussi accentuer la pression de performance. Au collégial, les personnes ayant obtenu de bons résultats à la première session ressentiraient d'ailleurs davantage « d'anxiété liée aux notes par la suite » (Benlakehal, 2023, p. 53).



À l'inverse, pour les personnes qui ont de moins bons résultats, l'évaluation entretient le **cercle vicieux de l'échec et du décrochage**, et agit comme une « prophétie auto-réalisatrice ». Des recherches montrent d'ailleurs que dès l'âge primaire, « beaucoup d'enfants commencent à considérer leur niveau de capacités scolaires comme une composante stable et définitive, quels que soient les efforts qu'ils fournissent » (Prokofieva et al., 2017).

Plusieurs études ont examiné les liens entre le type de rétroaction offert par l'évaluation et la motivation, montrant que la note en elle-même constitue plutôt une source de motivation extrinsèque, donc moins favorable à l'apprentissage et à la réussite (Gorichanaz, 2024, p. 2021; Koenka et al., 2021). La motivation par les notes est une « motivation atrophiée » (Elbow, 1997), qui n'amène pas les personnes étudiantes à s'engager en profondeur dans leur apprentissage ni à développer leurs compétences métacognitives.

Les stratégies d'apprentissage dans un contexte normatif

Les modalités d'évaluation façonnent les stratégies d'apprentissage des personnes étudiantes, et elles « conditionnent » également les pratiques enseignantes (Barbeau et al., 2021; Morrissette et Legendre, 2014, p. 236).

Plusieurs personnes enseignantes utilisent la notation comme levier pour induire certaines stratégies d'apprentissage, croyant par exemple que « l'attribution de notes est le meilleur moyen d'amener l'étudiant à étudier » (Bélanger et Tremblay, 2012, p. 76). L'importance accordée aux notes et la prédominance de la logique de classement dans l'évaluation renforce un « rapport utilitaire au savoir » chez les personnes étudiantes (CSE, 2018, p. 54) :

- Les stratégies pour obtenir de bonnes notes amènent les personnes étudiantes à négliger ce qui ne s'évalue pas en chiffres. Par exemple : choisir de ne pas s'engager dans des activités d'apprentissage ou des évaluations qui ne sont pas notées (qui ne « comptent pas ») ou dont la pondération est faible (Segueda, 2023).
- Parallèlement, la place prépondérante des notes dans un système nourri aux chiffres fait en sorte que l'enseignement a tendance à devenir une préparation aux examens, comme l'épreuve uniforme de français en est une illustration (MES, 2024, p. 73).
- La course aux notes incite à la tricherie et au plagiat. Les risques pour l'intégrité des évaluations constituent d'ailleurs la crainte principale exprimée par les acteurs et actrices de l'enseignement supérieur à l'égard de l'irruption des outils d'intelligence artificielle générative en enseignement supérieur (CSE et CEST, 2024, p. 31).



L'évaluation en enseignement supérieur : mesure des apprentissages ou épreuve de sélection ?

- Les examens sommatifs à enjeux élevés encouragent le bachotage, c'est-à-dire la préparation superficielle et intensive à la veille d'un examen, communément appelé le « bourrage de crâne » (CSE, 2018, p. 72; French et al., 2024).
- L'importance disproportionnée de la note dans le système d'évaluation décourage la prise de risques chez les personnes étudiantes. Par exemple, ne pas prendre un cours qui nous intéresse ou nous serait utile dans notre cheminement mais qui est perçu comme « plus difficile », pour éviter de faire baisser sa moyenne (Blum et Kohn, 2020, p. 13).

Pour poursuivre la réflexion

- Relire le rapport du Conseil supérieur de l'éducation (2018) Évaluer pour que ça compte vraiment, un état des lieux toujours d'actualité sur les pratiques d'évaluation des apprentissages à tous les niveaux du système éducatif.
- Prendre connaissance des résultats d'une étude de l'Observatoire sur la santé mentale étudiante en enseignement supérieur (OSMÉES) portant sur les décès par suicide chez les personnes étudiantes, mettant en évidence « la culture de performance » et « certaines pratiques d'évaluation qui contribuent aux enjeux d'anxiété, de compétition et [d]exclusion ».
- Consulter l'avis de la Fédération étudiante collégiale du Québec (FECQ) (2021) au sujet des pratiques d'admission universitaire.
- Visionner un webinaire de l'AQPC animé par Caroline Cormier et Bruno Voisard sur « La dénotation à des fins d'équité ».
- Joindre une communauté de pratique où l'on peut découvrir les pratiques alternatives de notation (PAN) et partager avec des « profs qui ont mis en place des systèmes de PAN dans leurs cours » (CdP-PAN)



Références

Barbeau, N., Frenette, É. et Hébert, M.-H. (2021). Et si les stratégies d'apprentissage des étudiants et leurs perceptions envers l'évaluation des apprentissages avaient un lien avec l'ajustement académique dans un contexte de persévérance aux études universitaires? *Revue internationale de pédagogie de l'enseignement supérieur*, 37(2). <https://doi.org/10.4000/ripes.2807>

BCI. (2025). La cote de rendement au collégial: ce qu'elle est, ce qu'elle fait. Bureau de coopération interuniversitaire. https://www.bci-qc.ca/wp-content/uploads/2025/02/CRC-Ce-quelle-est_Ce-quelle-fait-BCI-21janvier2025-4.pdf

Bélanger, D.-C. et Tremblay, K. (2012). Portrait actualisé des croyances et des pratiques en évaluation des apprentissages au collégial. <https://core.ac.uk/download/pdf/52989625.pdf>

Benlakehal, A. (2023, 22 février). Anxiété liée aux évaluations, symptômes dépressifs et rendement scolaire: exploration du rôle modérateur de l'adaptation au cégep. <https://papyrus.bib.umontreal.ca/xmlui/handle/1866/27994>

Blum, S. D. et Kohn, A. (2020). *Ungrading: why rating students undermines learning (and what to do instead)*. West Virginia University press.

Bowen, R. S. et Cooper, M. M. (2022). Grading on a Curve as a Systemic Issue of Equity in Chemistry Education. *Journal of Chemical Education*, 99(1), 185-194. <https://doi.org/10.1021/acs.jchemed.1c00369>

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T. et Welsh, M. E. (2016). A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86(4), 803-848.



Références

Campbell, P. I. (2024, 2 octobre). Decolonising the curriculum hasn't closed the gap between Black and white students – here's what might. The Conversation. <http://theconversation.com/decolonising-the-curriculum-hasnt-closed-the-gap-between-black-and-white-students-heres-what-might-238728>

Canning, E. A., LaCrosse, J., Kroeper, K. M. et Murphy, M. C. (2020). Feeling Like an Imposter: The Effect of Perceived Classroom Competition on the Daily Psychological Experiences of First-Generation College Students. *Social Psychological and Personality Science*, 11(5), 647-657. <https://doi.org/10.1177/1948550619882032>

Côté, F. (2017). L'évaluation des apprentissages au collégial: un réseau de concepts pour guider les pratiques évaluatives. *Pédagogie collégiale*, 30(4). <https://images.sdm.qc.ca/fichiers/Public/2017/B777763.pdf>

CSE. (2018). Rapport sur l'état et les besoins de l'éducation 2016-2018 – Évaluer pour que ça compte vraiment. Conseil supérieur de l'éducation.

CSE et CEST. (2024). Intelligence artificielle générative en enseignement supérieur: enjeux pédagogiques et éthiques. Conseil supérieur de l'éducation et Commission de l'éthique en science et en technologie. <https://www.cse.gouv.qc.ca/publications/ia-enseignement-sup-50-0566/>

Diédhiou, S., Duong Thi, D. et Robichaud, A. (2022). S'adapter à l'éthique de la différenciation en évaluation des apprentissages: une analyse de l'expérience des enseignants formés à l'étranger au Québec. *Mesure et évaluation en éducation*, 45(2), 37-67. <https://doi.org/10.7202/1101445ar>

Elbow, P. (1997). Grading student writing: Making it simpler, fairer, clearer. *New Directions for Teaching & Learning*, 1997(69), 127. <https://doi.org/10.1002/tl.6911>



Références

FECQ. (2022). Mémoire sur l'enseignement de la langue et l'Épreuve uniforme de français (EUF). https://docs.fecq.org/FECQ/M%C3%A9moires%20et%20avis/2022-2023/Memoire-EUF_118eCo_La-Malbaie.pdf

French, S., Dickerson, A. et Mulder, R. A. (2024). A review of the benefits and drawbacks of high-stakes final examinations in higher education. *Higher Education*, 88(3), 893-918. <https://doi.org/10.1007/s10734-023-01148-z>

Gorichanaz, T. (2024). "It made me feel like it was okay to be wrong": Student experiences with ungrading. *Active Learning in Higher Education*, 25(1), 67-80. <https://doi.org/10.1177/14697874221093640>

Huver, E. et Springer, C. (2011). Chapitre 1. L'évaluation, un champ en tensions. *Langues et didactique*, 17-69.

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E. et Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement: a case for written feedback. *Educational Psychology*, 41(7), 922-947. <https://doi.org/10.1080/01443410.2019.1659939>

MES. (2024). Regards croisés sur les conditions de réussite éducative des premiers cours de littérature et de philosophie au cégep ([Rapport du groupe de travail mis en place dans le cadre de la mesure 3.5 du Plan d'action pour la réussite en enseignement supérieur (PARES) 2021-2026]).

Morrissette, J. (2021). Un pluralisme disciplinaire et méthodologique comme perspective d'avenir du domaine. Dans *40 ans de mesure et d'évaluation*. Presses de l'Université du Québec.



Références

Morrisette, J. et Legendre, M.-F. (2014). Défis et enjeux de l'approche par compétences dans un contexte d'obligation de résultats. Dans J. Morrisette et M.-F. Legendre-Bergeron (dir.), *Enseigner et évaluer: regards sur les enjeux éthiques et sociopolitiques* (p. 211-245). Presses de l'Université Laval.

Moulin, S., Laplante, B., Lépine, M., Blain, M., Kamanzi, P. et Duffy, C. (2022). Gouverner la sélection scolaire par un instrument: le cas de la « cote de rendement au collégial » des universités québécoises. *Lien social et Politiques*, (89), 16-34. <https://doi.org/10.7202/1094546ar>

Nyström, A.-S., Jackson, C. et Salminen Karlsson, M. (2019). What counts as success? Constructions of achievement in prestigious higher education programmes. *Research Papers in Education*, 34(4), 465-482. <https://doi.org/10.1080/02671522.2018.1452964>

Posselt, J. R. et Lipson, S. K. (2016). Competition, Anxiety, and Depression in the College Classroom: Variations by Student Identity and Field of Study. *Journal of College Student Development*, 57(8), 973-989. <https://doi.org/10.1353/csd.2016.0094>

Prokofieva, V., Brandt-Pomares, P., Velay, J.-L., Hérold, J.-F. et Kostromina, S. (2017). Stress de l'évaluation scolaire: un nouveau regard sur un problème ancien. *Recherches & éducations*, (18). <https://doi.org/10.4000/rechercheseducations.4657>

Romainville, M. (2021). Ce n'est pas en pesant le cochon qu'on le fait grossir ou pourquoi l'évaluation formative peine à se faire une place au soleil. Dans *40 ans de mesure et d'évaluation* (p. 81-94). Presses de l'Université du Québec.

Segueda, S. (2023). Évaluation des apprentissages en contexte universitaire: Quand les étudiants finissent par adopter la posture du « cabri mort » face à la quantité de travaux évalués. *e-JIREF*, 9(2), 27-46. <https://doi.org/10.48782/e-jiref-9-2-27>



Références

Stiggins, R. (2008). Correcting "Errors of Measurement" That Sabotage Student Learning. Dans *The Future of Assessment*. Routledge.

Świątkowski, W. et Dompnier, B. (2017). La compétition dans le système universitaire: perspectives individuelles et sociétales. Dans F. Butera et C. Staerklé (dir.), *Conflits constructifs, conflits destructifs. Regards psychosociaux* (Antipodes, p. 101-117). <https://www.antipodes.ch/produit/conflits-constructifs-conflits-destructifs/>

Yerly, G. et Berger, J.-L. (2022). Des pratiques d'évaluation constructives à l'université? Sous quelles conditions un examen peut-il soutenir les processus d'autorégulation de l'apprentissage. *La Revue LEeE*, (6). <https://doi.org/10.48325/rleee.006.06>