Wide angle

Assessment in Higher Education:
A Measure of Learning or a Selection Process?

In higher education, grades have a major impact on a student's pathway: they open and close doors in the academic and professional world, and have repercussions on self-esteem, well-being and student aspirations.

Getting good grades, but more importantly, getting better grades than one's peers, is a necessary condition for standing out in a selection processes. In this competitive context, assessment, especially when it takes the form of a high-stakes summative exam, is often perceived as a stressful, no-holds-barred ordeal. Reduced to the dimension of grading, assessment is "most often seen as detrimental to learning and motivation" (Yerly & Berger, 2022, p. 7).

In this issue of Wide Angle, ORES explores some questions related to normative assessment in higher education:

- Where does normative assessment in education come from, and how can we explain its longevity?
- What impact does normative assessment have on students and their success?

Observatoire sur la réussite en enseignement supérieur

ORES

1 What is Normative Assessment?

For centuries, Western school systems have used grading to categorize and compare students. In the 20th century, with the massification of education, schools were given a new role: to determine the social destiny of individuals, sorting them according to their "merit," "aptitudes" or "efforts" (Romainville, 2021, p. 90).

Normative assessment is the mechanism through which categorization takes place: those who perform best are selected for access to the most prestigious or restricted disciplines or programs, while the "weakest" are ejected from the system or relegated to less valued pathways. In other words, the function of normative assessment is to rank students for selection purposes.

> Normative assessment: "A mode of assessment in which a subject's performance is compared (...) to that of others in a reference group on the basis of the same instrument" (Legendre 2005, cited in (CSE, 2018, p. 8)).

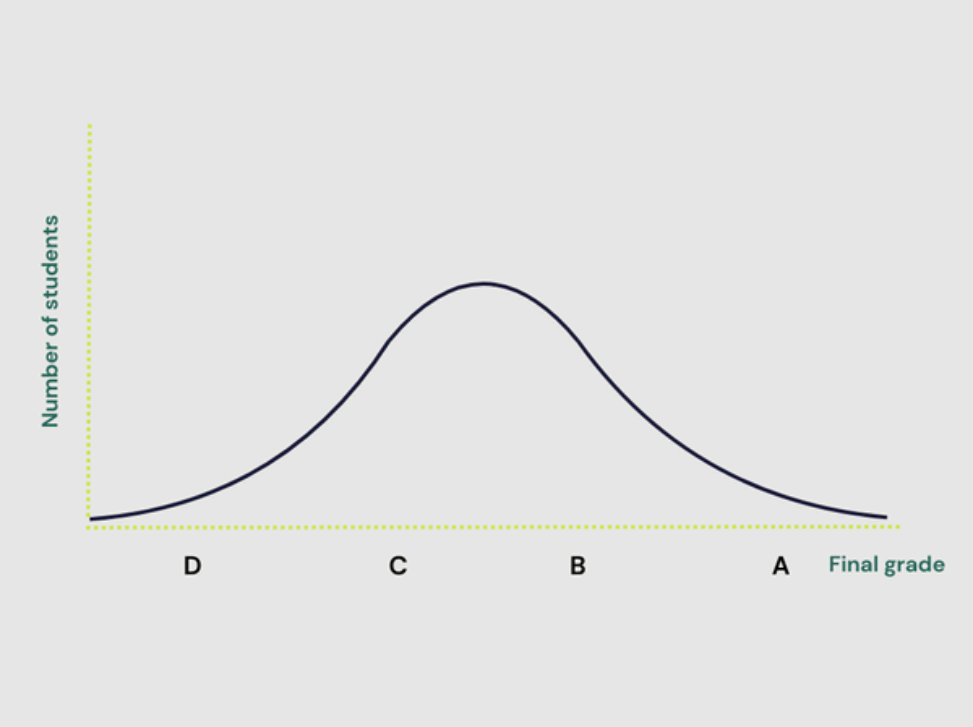
Breaking with this normative approach, specialists in educational sciences have for decades been advocating the adoption of **assessment practices centred on learning**. In this vision, which lies at the heart of the competency-based approach, **assessment is a qualitative, criterion-referenced evaluation designed to support students throughout the learning process** (Côté, 2017). It measures the achievement of learning targets, rather than comparing students to one another

Despite the scientific consensus in favour of learning-centred assessment, normative assessment practices remain very present in higher education.
Why is this so?
On what beliefs is the longevity of the normative approach based?

The Normal Curve

The normal curve or Gauss curve (bell-shaped) is the graphical representation of a binomial probability distribution of random results. For example, if you throw a handful of coins on the ground, such distribution predicts that the proportion of “tails” results is much more likely to be between 40% and 60% rather than 90% or 10%.

At the turn of the 20th century, human sciences adopted this law to model human behavior, giving rise to psychometrics. This new science, which invented standardized tests and the notion of IQ, assumes that, in the population, intellectual abilities are distributed along a normal curve. Thus, in the field of education, the psychometric model is at the origin of normative assessment, seen as a more “objective” and “scientific” way of measuring student performance (Bowen & Cooper, 2022; Brookhart & al., 2016).



Concretely, when applied to the field of educational assessment, this conception imposes a certain distribution of grades, which should reflect the supposedly “natural” distribution of students' abilities. In a normative logic, “a good method of assessment should result in several average grades and few very good and very bad grades” (Huver & Springer, 2011). Technically, this means that the success of some depends on the failure of others.

An Invisible Reference Point

In the 1970s, psychologist Benjamin Bloom criticized normative assessment by posing the following question: How can learning outcomes reflect the laws of probability, when learning and teaching conditions differ for each individual?

In the decades that followed, educational sciences distanced themselves from the normative approach, first adopting a behaviourist perspective inherited from Bloom (teaching paradigm) and evolving toward the socio-constructivist stance that is now dominant (learning paradigm).

Despite this, the normal curve remains the "silent partner" of the grading system, whether in letters or numbers (Brookhart et al., 2016). While criterion-referenced assessment is officially valued, normative practices (such as comparing individual results to the group average or "normalizing" grades) persist, from primary school to university. **In higher education, selection mechanisms for prestigious internships and scholarships play a key role in consolidating this normative approach.**

According to the Conseil supérieur de l'éducation, the logic for academic selection that prevails in higher education has cascading effects across the entire school system, contributing to the "overvaluation of grades and ranking" from elementary school onwards (CSE, 2018, p. 69).

The College Performance Score (R score)

The R score is the main criterion that determines access to certain university programs or institutions.

Prior to the 1980s, the college average was used by universities to rank and select students for admission. Subsequently, a more elaborate calculation formula including the mean and standard deviation (the Z-score) was used until 1995, when universities adopted the R score. This new calculation method was designed to correct the biases introduced with the Z-score, by taking into account the relative strength of the group:

> *"The R score combines, for each course taken by a student, three pieces of information: an indicator of that student's position based on the grade obtained in their group (the college Z-score), an indicator of the strength of that group and an indicator of the dispersion of that group" (BCI, 2025, p. 7).*

The College Performance Score (continued)
According to the Bureau de coopération universitaire, which is responsible for this calculation method, the R score is a "fair and equitable measurement tool" that gives "the best students from all colleges an equal opportunity to access the most restricted university programs" (BCI, 2025, p. 20).

However, the Conseil supérieur de l'éducation notes that the R score feeds a sometimes unhealthy competition, in addition to systematically penalizing students who are experiencing "a slow or difficult start" (CSE, 2018, p. 47). A research team has further shown that the R score continues to reflect a bias based on group strength, and "legitimizes and even accentuates social inequalities in access to restricted university education programs" (Moulin et al., 2022).

2 What are the Repercussions on Learning and Student Success?

Collaboration VS. Competition

The predominance of the ranking logic in assessment exacerbates competition and discourages collaboration among students. Indeed, in a normative logic where everyone positions themselves in relation to others, each successful student makes success a little harder to achieve for the others. This climate of competition is particularly exacerbated in certain prestigious study programs, for example law or STEM fields (Canning et al., 2020; Nyström et al., 2019).

> In an education system based on selection, *"seeking to be better than others becomes an adapted response to the structural constraints of the academic environment"* (Świątkowski & Dompnier, 2017, p. 107).

The results of this "obstacle course" are seemingly merit-based. In reality, this system favours students from privileged backgrounds, who are better prepared for CEGEP or university. Conversely, it disadvantages people from underrepresented groups (particularly first-generation students), who are more likely to experience a sense of imposterism in higher education, lack the implicit knowledge and don't have support from those around them to know how to perform well in assessments (Bowen & Cooper, 2022, p. 190; Campbell, 2024; Canning et al., 2020).

Anxiety and the Race for Grades

The culture of competition amplifies grade-related stress and anxiety, particularly among individuals from historically underrepresented groups in higher education or certain fields (e.g., women in engineering) (Posselt & Lipson, 2016).

To counter the negative effects of high levels of competition on students' mental health, several medicine programs in Canada and elsewhere have actually replaced their grading system with a "pass-fail" designation (CSE, 2018, p. 51; Nyström et al., 2019).

Are Stressful Assessments Necessary?

A common idea is that stressful assessments are part of learning, since they are a preparation for the stressful realities of working life. However, the typical constraints of high-stakes exams (e.g., a time-limited examination in a closed room with no access to one's notes) do not correspond to the reality of the professional, academic or social environments that students will enter after their training. In this sense, they are not authentic assessments.

According to members of the comité de travail sur les « cours défis » au collégial, an examination such as the Ministerial Examination of College French or épreuve uniforme de français (EUF), which consists in producing a 900-word essay in 4.5 hours, "may not offer an accurate picture of a student's abilities" because of the "intense stress" it generates (MES, 2024, p. 73). The Fédération étudiante collégiale du Québec describes the EUF as a "totally artificial, anxiety-provoking and outdated" assessment context (FECQ, 2022, p. 21).

Grades and Motivation

Assessment and grading generate highly differentiated emotional states in students (Stiggins, 2008, p. 240). For those who "perform" well, the assessment system can act as a driver for success: it generates confidence, fuels the desire to perform, to put in effort, and so on. With that being said, good results can also accentuate the pressure to perform. At the college level, people who obtained good results in their first semester are more likely to experience "grade-related anxiety later on" (Benlakehal, 2023, p. 53).

Conversely, for those with poorer results, assessment maintains the **vicious circle of failure and dropping out**, acting as a "self-fulfilling prophecy." Indeed, research shows that as early as primary school age, "many children begin to see their level of academic ability as a stable and definitive component, no matter how hard they try" (Prokofieva et al., 2017).

A number of studies have examined the link between the type of feedback provided through assessment and motivation, showing that grades themselves tend to be a source of extrinsic motivation, which is less conducive to learning and success (Gorichanaz, 2024, p. 2021; Koenka et al., 2021). Motivation through grades is "atrophied motivation" (Elbow, 1997), which does not lead students to engage deeply in their learning or to develop their metacognitive skills.

Learning Strategies in a Normative Context

The terms of assessment shape students' learning strategies, and they also "condition" teaching practices (Barbeau et al., 2021; Morrissette & Legendre, 2014, p. 236).

Many teachers use grading as a lever to induce certain learning strategies, believing, for instance, that "assigning grades is the best way to get students to study" (Bélanger & Tremblay, 2012, p. 76). The importance attached to grades and the predominance of the grading logic in assessment reinforce a "utilitarian relationship to knowledge" among students (CSE, 2018, p. 54):

- Strategies for getting good grades lead students to **neglect what cannot be evaluated in numbers**. For example: choosing not to engage in learning activities or assessments that are not graded (that don't "count") or whose weighting is low (Segueda, 2023).

- At the same time, the prominent place of grades in a system fuelled by numbers means that **teaching tends to become preparation for examinations**, as exemplified by the Ministerial Examination of College French (MES, 2024, p. 73).

- The race for grades encourages **cheating and plagiarism**. Moreover, in the wake of the emergence of generative artificial intelligence tools in higher education, the risk that the integrity of assessments may be impeded represents the main fear expressed by higher education stakeholders (CSE & CEST, 2024, p. 31).

- High-stakes summative examinations encourage **cramming**; i.e., superficial and intensive preparation on the eve of an exam, commonly known as “brain dump preparation” (CSE, 2018, p. 72; French et al., 2024).

- The disproportionate importance of grades in the assessment system **discourages risk-taking** among students. For example, not taking a course that interests us or would be useful for our pathway, but which is perceived as “more difficult,” to avoid lowering one's GPA (Blum & Kohn, 2020, p. 13).

For Further Reflection

- Review the report published by the Conseil supérieur de l'éducation (2018) Évaluer pour que ça compte vraiment, a still topical review of learning assessment practices at all levels of the education system.

- Learn more about the results of a study by the Observatoire sur la santé mentale étudiante en enseignement supérieur (OSMÉES) on suicide deaths among students, highlighting “the performance culture” and “certain assessment practices that contribute to issues of anxiety, competition and exclusion.”

- Read the statement by the Fédération étudiante collégiale du Québec (FECQ) (2021) on university admissions practices.

- Watch the AQPC webinar hosted by Caroline Cormier and Bruno Voisard entitled “La dénotation à des fins d'équité.”

- Join a community of practice where you can discover alternative rating practices (NAPs) and exchange with “teachers who have implemented NAP systems in their courses” (CdP-PAN).

References

Barbeau, N., Frenette, É. & 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. & 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. & Kohn, A. (2020). Ungrading: why rating students undermines learning (and what to do instead). West Virginia University press.

Bowen, R. S. & 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. & 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.

References

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., LaCosse, J., Kroeper, K. M. & 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 & 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. & 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

Observatoire sur la réussite en enseignement supérieur

ORES

References

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. & 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. & 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. & 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.

References

Morrissette, J. & Legendre, M.-F. (2014). Défis et enjeux de l'approche par compétences dans un contexte d'obligation de résultats. Dans J. Morrissette 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. & 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. & 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. & 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. & 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

References

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

Świątkowski, W. & 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. & 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

Observatoire sur la réussite en enseignement supérieur