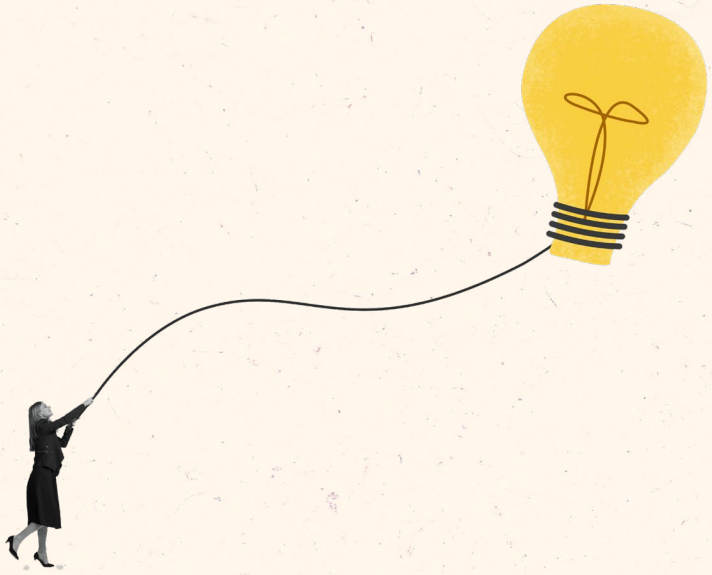
Alternative Grading Practices

A better way to support and attest to learning

Bruno Voisard, Caroline Cormier
and François Arseneault-Hubert

Source: iStock/Anton Vierietin

Bruno's confession: For 20 years, I gave marks for every summative evaluation, a practice I could no longer contemplate today. What happened to me? You could say I've been struck by alternative grading practices (AGP). All right, this may sound strange, and I can see how someone might think: "It looks like he's joined a cult." From my point of view, I'd say I've come across ideas that have brought about a change in my approach to evaluation. An article in Chemical and Engineering News (Arnaud, 2021) that focused on specifications grading (Nilson, 2015) was my way into the world of AGPs. I then discovered a variety of these practices and experimented with several of them in the classroom. My only disappointment? Not learning about them sooner!



Source: iStock/Anton Vierietin

The usual way of calculating a final grade, by adding up the marks from each summative evaluation, penalizes errors during the course of learning and undermines students' intrinsic motivation (Koenka *et al.*, 2021). The famous question "Does it count?" is a case in point. What's more, the usual calculation of the final grade does not adequately reflect the degree of attainment of terminal objectives, since it penalizes errors even when we have evidence that learning has been achieved after a summative evaluation. In short, this calculation does not adequately achieve either of the broad goals of evaluation, which are to **support learning** throughout the session and to **attest to student achievement** at the end of the session (Crooks, 1988; CSE, 2018). By moving away from the summation of grades, AGPs aim to better achieve the broad goals of evaluation.

AGP features and benefits

The term *alternative grading practices* refers to the way in which work is graded. It's not the same as evaluation, which has a broader meaning, but grading has a major influence on the way we evaluate. Unlike conventional grading, AGPs do not involve assigning a grade to each summative evaluation. As they are very distinct from the grading practices generally applied at the college level (Bélanger & Tremblay, 2012; CSE, 2018; Howe & Ménard, 1993; Leroux, 2009), they are referred to as *alternative* practices in the professional and research literature (see, for example, Townsley & Schmid, 2020).

Alternative practice systems include mastery or standards-based grading, specifications grading, contract grading and ungrading.

Table 1

Characteristics of three AGP families

Mastery grading or standards-based grading	Approaches in which the degree to which objectives have been met, rather than performance on an evaluation, is graded, and in which each evaluation is clearly associated with one or more objectives.
Specifications grading or contract grading	Approaches in which bundles of assignments are predefined, with each bundle that students satisfactorily complete corresponding to a given final grade.
Ungrading	Approaches in which qualitative feedback, self-assessment and peer-evaluation are encouraged, and as little work as possible is graded.

The boundaries between these grading approaches are permeable, and the systems implemented by AGP proponents sometimes have features of several of these families of approaches simultaneously. Some teachers even use approaches inspired by both AGPs and conventional practices to gradually familiarize themselves with AGPs, to comply with certain administrative constraints, or both (see, for example, Mercier, 2023, for an account of the practice of such a hybrid grading approach).

This variety of approaches means that AGPs can be adapted to a wide range of educational contexts. What they have in common is:

- offering students several opportunities to demonstrate their mastery of learning objectives (Clark & Talbert, 2023; Feldman, 2024; Nilson, 2015);
- not penalizing their errors during the course of learning, taking into account only those traces that best represent the learning achieved (Feldman, 2024; O'Connor, 2010);
- providing continuous, individualized qualitative feedback about progress with respect to each objective (Hattie & Timperley, 2007; Koenka *et al.*, 2021);
- determining the final grade based on the degree of attainment of objectives associated with the competencies to be developed (Guskey, 2001; Townsley & Schmid, 2020);
- grading the degree of attainment of course objectives using a limited number of levels, typically two to five, which improves intra- and interjudge reliability (Brookhart, 2018; Clark & Talbert, 2023; Nilson, 2015);
- ensuring that these terminal objectives are clearly defined and transparently communicated to students (Feldman, 2024).

Table 2

Characteristics of AGPs enabling them to achieve the two broad goals of evaluation

To better support learning	**To better attest to achievement**
AGPs **do not penalize error during the course of learning**, which means offering **several opportunities** to demonstrate attainment of objectives.	Teachers using AGPs **determine the final grade based on achievement of course objectives**, rather than by adding up marks**.**
To better guide learners' efforts, teachers provide **precise information about the achievement of each learning objective** related to the competency to be acquired.	Course objectives are graded on the basis of a **limited number of levels**, usually between two and five, which are clearly and transparently defined.

At the end of this article, we present examples of evaluation systems that possess these characteristics, but first, let's look at what might justify the effort of a change in grading system.

Reconstructing the link between grades and learning targets

In their certification evaluation, a student may clearly demonstrate that they have not developed the competency to a satisfactory level. However, the marks they have accumulated may still result in a mark of 60%. **How can we avoid situations where there is a contradiction between our professional judgment and the final grade obtained by summation?** And above all, **how can we more clearly inform students** during the session of the gap between their learning and the learning targets? These were two of the questions we were asking ourselves, and to which we found answers with the help of AGPs.

This problem of information quality is well highlighted by the Conseil supérieur de l'éducation in its report *Evaluating So It Truly Counts*: "the informational value of grades remains scant. A criterion-referenced scale would allow to more accurately impart what has been successfully acquired and what remains to be achieved" (CSE, 2018, p. 3). Summative grading doesn't communicate this path forward with clarity and transparency, because a variety of knowledge, skills, behaviours, etc.—related or unrelated to learning objectives—are aggregated into a single grade, a process that makes valuable information disappear (Clark & Talbert, 2023; Feldman, 2024).

To maintain coherence between the grade and learning targets, we must also avoid contaminating it with elements that fall outside the scope of learning targets, such as participation, meeting deadlines, completion of learning activities like homework or formative evaluations, etc. (Feldman, 2024). In fact, these attempts to control student behaviour through grades mean that the latter no longer accurately reflect the degree to which learning targets have been met, as well as undermining intrinsic motivation (Niemiec & Ryan, 2009; Sarrazin, Tessier & Trouilloud, 2006).

Recognizing the role of error in learning

In contrast to the case of the student who obtains 60% without attaining the objectives associated with the course competency, a student may encounter difficulties at the beginning of the session and obtain a final grade that does not adequately reflect their learning if they have overcome these difficulties later in the session. This case illustrates a paradox caused by certain grading practices commonly applied in college and in education in general. While most teachers believe that **error is an integral part of the learning process** (Astolfi, 2015), many give marks on learning activities (Bélanger & Tremblay, 2012), which has the effect of penalizing errors made during the course of learning.

If we wish to evaluate in accordance with the principle that error is inherent to learning, we must not penalize a student for an error in a summative evaluation in the fifth week of the session, especially if the incomplete learning has been achieved

Source: iStock/Anton Vierietin

subsequently. **To enable us to validate whether this incomplete learning has been achieved later in the session, it is necessary to plan repeat evaluations or remedial activities**. This evaluation practice is an integral part of all AGP systems. We're not talking here about deferred evaluations, but rather about additional opportunities for evaluation. Limits must, of course, be placed on the number of such opportunities, or on the last possible moment for demonstrating achievement of learning targets, both to avoid procrastination and to ensure that management of retakes does not become too burdensome for teachers (Clark & Talbert, 2023). To ensure that the grade accurately reflects the competencies acquired at the end of the course, these additional opportunities should be accessible to all, and the grade that can be achieved should not be capped either (Feldman, 2024): **it's the achievement of learning targets that should be valued, not the moment in the session when it occurs**. For the final grade to be a fair reflection of learning, two people who demonstrate the same competency at the end of the session should get the same grade, no matter how far they've come to get there.

Thus, in a course where AGPs have been implemented, summative evaluation can retroactively become formative: it no longer "counts" in the sense that it no longer enters into the composition of the final grade, but it counts a great deal in the sense that it has become valuable feedback for the student. In other words, all evaluation is formative until the learning target is reached. This optimizes the effectiveness of feedback by specifically informing students about the gap between their understanding or performance and the learning targets (Hattie & Timperley, 2007).

This approach to corrective feedback is consistent with research in the psychology of learning. Indeed, testing oneself or being tested is more effective in promoting learning than studying (Karpicke & Roediger, 2008; Roediger & Karpicke, 2006). This is known as the *testing effect*. What's more, the testing effect is even stronger in the presence of corrective feedback (Roediger & Butler, 2011). Repeat evaluations thus make it possible to use the learning levers of the testing effect and corrective feedback, specifically for learning targets that have not yet been reached.

Impact on motivation

Another reason for using AGPs is that evaluation during the course is not accompanied by a grade that can directly influence the final grade: the grade itself has a negative effect on motivation. Of course, to comply with the College Education Regulations (CER) and the colleges' Institutional Policies on the Evaluation of Student Achievement (IPESA), students must be awarded a final grade, and the passing grade must correspond to 60%.[1]

But is it useful or necessary to give a definitive, quantitative grade on each summative evaluation? Some people would say that a grade is a form of feedback, a way of situating the student in relation to what they have learned. However, qualitative written feedback alone is preferable to feedback accompanied by a grade: **the presence of a grade does not enhance learning, and is detrimental to intrinsic motivation** (Koenka *et al.*, 2021), especially for learners in difficulty. The case of students who consult the grade without bothering about the accompanying feedback is a good illustration of this phenomenon. It is therefore preferable not to give a grade when it is not necessary (for example, by evaluating as often as possible formatively rather than summatively) and, when it is, to give feedback first and only give a grade later. An approach where the only grade is a mention of whether or not the learning target has been reached, and where the feedback then clearly becomes an invitation to carry out learning with a view to the next opportunity for revision or retake, is compatible with these recommendations.

[1] One of the aims of our research project is to identify local constraints (institutional, departmental or cultural) that may limit the implementation of AGPs in certain colleges, and to propose adaptations to AGPs to take these constraints into account.

Implementing change

Implementing AGPs may seem like a considerable undertaking, but it can be done on at least two fronts: establishing clear standards (what does the demonstration of mastery of my course content look like?) and evaluation methods (what forms of evaluation and retakes can I envisage to observe this demonstration of mastery?). Without changing everything at once, it's possible to implement AGPs in small bites, for example:

- transforming a summative evaluation with low weighting into a genuine formative evaluation, intended to give qualitative feedback to students (a form of denotation [Blum, 2020]), a formative evaluation that can be done by peers;
- evaluating a written assignment on the basis of a list of characteristics that correspond to a satisfactory production, in which case the learning target is considered to have been reached (this can be described as *dichotomous evaluation*) rather than giving a percentage grade, and giving students the chance to rework this assignment without penalty if it does not reach the desired level;
- building our course's grading rubric around learning targets (objective 1, objective 2, etc.) rather than types of evaluations (exam 1, homework 1, etc.), and communicating the results accordingly, so as to highlight what has been learned and what remains to be learned by each student.

Source: iStock/Anton Vierietin

For teachers who, like us, dare to embark on the complete conversion of a course to incorporate AGPs, here is an approach inspired by those proposed by several authors (Cilli-Turner *et al.*, 2020; Clark & Talbert, 2023; Nilson, 2015; Townsley & Schmid, 2020).

1. What learning objectives do I want to measure?

First, we need to decide, in line with the program-based approach and the ministerial devis, and in agreement with our department, what the specific learning objectives are for our course. To be able to communicate them clearly to students, the objectives written in the general course outline, where it exists, may have to be reworded. In addition, depending on the degree of precision of the general course outline, the objectives often need to be clarified so that they can be used as a basis for constructing the evaluations administered throughout the session.

2. What does mastering these objectives look like?

This step of the process can be more or less complex, depending on the tools we already have: if the general course outline is well written, or if we have quality grading rubrics, we may already have the answer to this question. In other cases, such as when we start with a course in which success is determined by the sum of grades corresponding to at least 60%, rather than the achievement of learning objectives, this step may require more work. It is then relevant to specify the behaviours observable on the part of a student who has achieved the objectives at a level judged satisfactory for passing the course. A number of AGP

models suggest high expectations, especially when evaluation is dichotomous and a single demonstration of mastery is enough for an objective to be considered achieved.

It is crucial to ask ourselves what our students should be able to do by considering the next step in their academic or professional careers. This is the time to think about the types of evaluations, and the criteria for each one, that can be used to evaluate the achievement of mastery. The use of grading rubrics will help to improve the fairness of our evaluations and precisely define our expectations of students, which can improve their learning (Brookhart, 2018; Panadero *et al.*, 2012).

3. How will we determine the final grade in the course?

Different methods can be used to convert qualitative or dichotomous grades into a percentage grade at the end of a session. To do this, it is important to determine how to aggregate the information gathered throughout the session, in a way that makes sense in terms of the level of achievement of the course objectives. Ultimately, AGPs open the door to many possibilities for composing the final grade, as opposed to the usual grading practices that are limited to a single way of doing things, i.e. adding up the grades of all summative evaluations. In this article, to give an idea of this variety, we will illustrate two different methods in the case studies presented below.

4. When and how will students be able to demonstrate their mastery?

Since AGPs rely on students being able to demonstrate their mastery as their learning is completed, we need to design a system that gives them the chance to do so, at the right time, without overloading the process. How many opportunities do we want to offer? Will we offer repeat evaluations or review opportunities? Will we ask students to complete a revision activity before being entitled to a retake? Here again, the cases presented below present different avenues.

Two cases of AGP implementation at the college level

François adopts AGPs to better support his Nursing students: an example of specifications grading

The *Chemistry for Nursing I* and *II* courses are mandatory (and often considered unpleasant) for Nursing students who haven't taken the Secondary 5 chemistry course. In the fall of 2022, returning from paternity leave during which I had had the opportunity to reflect extensively on the subject, I was determined to implement specifications grading (Nilson, 2015) in this course to offer retakes of evaluations or assignments with transparent objectives.

I planned for four possible grades at the end of the course: 50, 60, 80 and 100%. The relatively small number of final grades made it possible to establish and communicate significant differences in the level of achievement of the course objectives. This gave each student the chance to plan the extent of their engagement in the course in full knowledge of the facts. Indeed, each grade was matched by an increasing number of assignments and tests (hereafter referred to as "productions"), each of which could be retaken or revised. Aiming for a higher grade therefore meant agreeing to work harder. Finally, each production was considered "accepted" or "not yet accepted" according to a list of specifications, i.e. quality indicators sufficiently detailed for their presence or absence to be unequivocally established.

As I was teaching *Chemistry for Nursing I* in the fall of 2022 and its sequel in the winter session, I had the opportunity to adjust. I realized that retakes had to be carefully planned and, above all, limited. Indeed, the absence of limits, or blurred limits, is an invitation not only to procrastination, but also to a lack of seriousness when it comes to evaluations. In the following session, in *Chemistry for Nursing II*, I reduced the number of required productions and limited the retakes to one per production. At the same time, I improved the frequency and consistency of feedback. Once the evaluation process had been clarified in this way, no one was surprised by their grade or even failure at the end of the session.

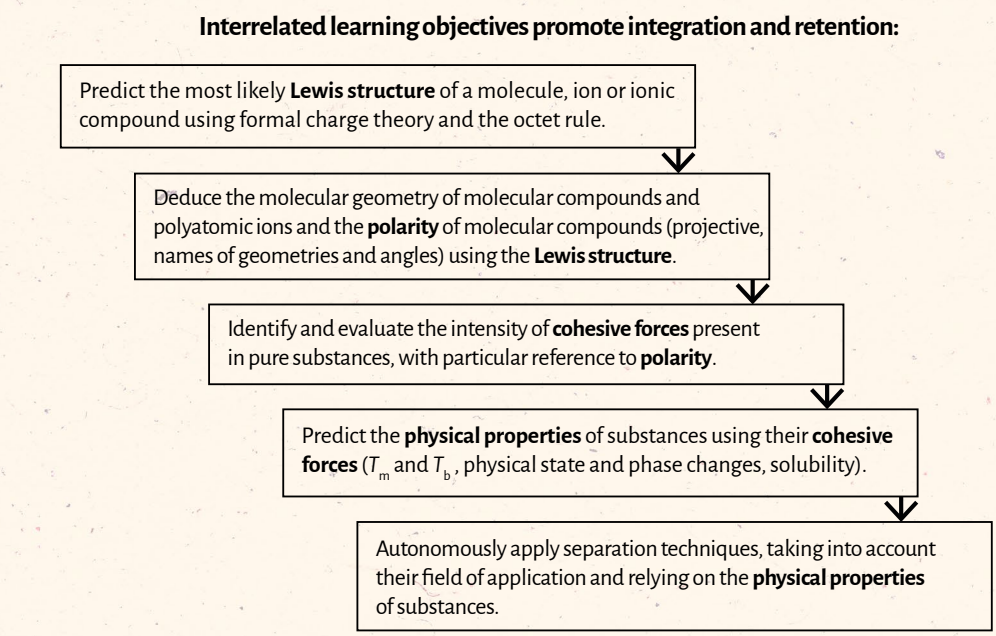
In cases where the planned retake led me to conclude that the competency was still not developed to a sufficient level, I would invite the students to review the content and come and see me to discuss what they had learned. Almost invariably, at the end of one of these informal visits, I could testify favourably to their acquired competency. For me, offering these multiple and varied opportunities to demonstrate the achievement of objectives, while limiting them, meant reconciling the two purposes of evaluation with each other, and with my work capacity.

Bruno's AGP model for Science: an example of mastery grading

In the winter of 2022, I converted a course to include AGPs for all evaluations for the first time: the *General Chemistry* course, the first chemistry course in the Science program. In this experiment with AGPs, I broke down the learning targets defined in the general course outline into 16 specific objectives, 9 of which were considered essential and the other 7 complementary. Most objectives were evaluated by means of dichotomous graded tests ("mastered" or "learning"). For each objective, a grading rubric detailed the characteristics of a satisfactory degree of mastery, which was quite demanding. These one-page tests were given at the start of class, during most weeks of the session. The three "exams" were an opportunity for students to repeat objectives they had not yet mastered: they did not therefore have to answer questions relating to objectives they had already mastered. To avoid fragmentation of knowledge and encourage retention, we took care in the department to ensure that the objectives were interrelated, especially since a single demonstration of mastery for a given objective was sufficient. For objectives relating to laboratories and scientific writing, students were given the opportunity to revise their work rather than repeat it.

Figure 1

Statement of five interrelated objectives of my General Chemistry course



The final grade was simply calculated based on the number of objectives mastered, provided that the objectives considered essential had been mastered. If this was not the case, the final grade was capped at 55%, in line with an approach proposed by Robert Howe (2006). These essential objectives had been chosen based on the skills that we in the department considered essential for further progress in the chemistry courses. In fact, this was one of my motivations for adopting this system: I was always amazed at the extent to which, in the past, students who found themselves in later courses were in difficulty because they had not mastered the basic skills taught in the *General Chemistry* course, even though they had passed this course. By defining essential objectives along with a dual pass threshold, I was sending students a clearer message: they had to focus on acquiring these skills first, until they had mastered them, even if it meant setting aside the complementary objectives.

Implementing this model within the administrative constraints imposed by my college required a great deal of concerted effort. First, the department agreed to derogate from the general course outline on the grounds of pedagogical experimentation. I was also careful to have my model approved at a departmental meeting before the start of the session. Since then, we've chosen to write our general course outlines clearly specifying the tasks that students should be able to do at the end of the course, rather than specifying the evaluation methods. This gives us greater latitude in the choice of grading practices, while ensuring equitable evaluation and program coherence.

Conclusion

Alternative grading practices have enabled us to align our practices with ideas and values we consider fundamental in education, such as justice, equity and caring. They also operationalize recommendations made by the Conseil supérieur de l'éducation in 2018 that we feel have found too little resonance in the college environment, by enabling us to reconcile grading with the two broad goals of evaluation, namely, to support learning and attest to achievement.

After a first AGP implementation in Winter 2022 with the support of our CEGEP, we have now started a research project on the subject thanks to PAREA funding (2023-2026). We aim to shed light on the effects of AGPs on learning and on various motivational variables. We are also seeking to describe ways of implementing AGPs in various college contexts.

The world of AGPs is little known in Quebec. Although we've only touched on it in the few pages of this article, we hope we've piqued your curiosity about the possibilities they offer! ▬

Bruno Voisard has been a teacher in the Chemistry Department at Cégep André-Laurendeau since 2001. He has implemented several innovative pedagogical practices, including the flipped classroom, alternative grading methods and inquiry-based laboratories. He is also interested in developing students' problem-solving skills and research autonomy. Moreover, he has been involved in the revision of the Science program, both locally and at the provincial level.

bruno.voisard@claurendeau.qc.ca

Caroline Cormier has been teaching chemistry at Cégep André-Laurendeau since 2008. She has led several educational research projects, notably on the flipped classroom, oral scientific communication, and autonomy in the laboratory. One of her priorities is to apply her research findings to professional practice. In addition to her teaching duties, she is also a columnist on the radio program *Moteur de recherche*, broadcast on ICI Première.

caroline.cormier@claurendeau.qc.ca

François Arseneault-Hubert has been teaching chemistry at Cégep André-Laurendeau since 2013. Confinement and distance teaching have given him the opportunity to overhaul his evaluation and grading methods, in line with his keen sense of justice. He owes this, among other things, to his two young children who, in addition to chronically depriving him of sleep, constantly remind him that there is no age limit to being a person worthy of respect.

françois.arseneault-hubert@claurendeau.qc.ca

References

Arnaud, C. H. (2021). "How an alternative grading system is improving student learning," *Chemical & Engineering News*, vol. 99, nº 15 [Online].

Astolfi, J.-P. (2015). *L'erreur, un outil pour enseigner* (12th edition), ESF éditeur.

Bélanger, D.-C. and K. Tremblay (2012). *Portrait actualisé des croyances et des pratiques en évaluation des apprentissages au collégial*, Collège de Maisonneuve.

Blum, S. D. (ed.). (2020). *Ungrading: Why Rating Students Undermines Learning* (First edition), West Virginia University Press.

Brookhart, S. M. (2018). "Appropriate Criteria: Key to Effective Rubrics," *Frontiers in Education*, vol. 3, nº 22 [Online].

Cilli-Turner, E. *et al*. (2020). "Mastery Grading: Build-A-Syllabus Workshop," *PRIMUS*, vol. 30, nºs 8-10, p. 952-978.

Clark, D. and R. Talbert (2023). *Grading for growth: A guide to alternative grading practices that promote authentic learning and student engagement in higher education*, Routledge.

Crooks, T. J. (1988). "The Impact of Classroom Evaluation Practices on Students," *Review of Educational Research*, vol. 58, nº 4, p. 438-481.

CSE (2018). *Report on the State and Needs of Education 2016-2018 - Evaluating So It Truly Counts*.

Feldman, J. (2024). *Grading for equity: What it is, why it matters, and how it can transform schools and classrooms* (Second edition), Corwin Press inc.

Guskey, T. R. (2001). "Helping Standards Make the GRADE," *Educational Leadership*, p. 20-27.

Hattie, J. and H. Timperley (2007). "The power of feedback," *Review of Educational Research*, vol. 77, nº 1, article 1.

Howe, R. (2006). "La note de l'évaluation finale d'un cours dans l'approche par compétences : Quelques enjeux pédagogiques," *Pédagogie collégiale*, vol. 20, nº 1, p. 10-15.

Howe, R. and L. Ménard (1993). *Croyances et pratiques en évaluation des apprentissages : Étude des croyances et des pratiques des enseignants des cégeps à l'égard de l'évaluation des apprentissages*, Collège Montmorency; Programme d'aide à la recherche sur l'enseignement et l'apprentissage.

Karpicke, J. D. and J. R. Blunt (2011). "Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping," *Science*, vol. 331, nº 6018, p. 772-775.

Karpicke, J. D. and H. L. Roediger (2008). "The Critical Importance of Retrieval for Learning," *Science*, vol. 319, nº 5865, p. 966-968.

Koenka, A. C. *et al*. (2021). "A meta-analysis on the impact of grades and comments on academic motivation and achievement: A case for written feedback," *Educational Psychology*, vol. 41, nº 7, p. 922-947.

Leroux, J. L. (2009). *Analyse des pratiques évaluatives d'enseignantes et d'enseignants dans une approche par compétences au collégial*, Université de Sherbrooke.

Mercier, C. (2023). "Ma 1re expérience de notation par spécifications," *Éductive*, November 30 [Online].

Niemiec, C. P. and R. M. Ryan (2009). "Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice," *Theory and Research in Education*, vol. 7, nº 2, p. 133-144.

Nilson, L. B. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*, Stylus Publishing.

O'Connor, K. (2010). "Grades - When, Why, What Impact, and How?" *Education Canada*, vol. 50, nº 2, p. 38-41.

Panadero, E., J. A. Tapia, and J. A. Huertas (2012). "Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education," *Learning and Individual Differences*, vol. 22, nº 6, p. 806-813.

Roediger, H. L. and A. C. Butler (2011). "The critical role of retrieval practice in long-term retention," *Trends in Cognitive Sciences*, vol. 15, nº 1, p. 20-27.

Roediger, H. L., and J. D. Karpicke (2006). "The Power of Testing Memory: Basic Research and Implications for Educational Practice," *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, vol. 1, nº 3, p. 181-210.

Sarrazin, P., D. Tessier and D. Trouilloud (2006). "Climat motivationnel instauré par l'enseignant et implication des élèves en classe : L'état des recherches," *Revue française de pédagogie*, vol. 157, p. 147-177.

Townsley, M. and D. Schmid (2020). "Alternative grading practices: An entry point for faculty in competency-based education," *The Journal of Competency-Based Education*, vol. 5, nº 3 [Online].