EVALUATING SO IT TRULY COUNTS

Report on the State and Needs of Education 2016-2018

SUMMARY

Conseil supérieur
de l'éducation

Québec

Evaluating So It Truly Counts

Report on the State and Needs of Education 2016-2018

SUMMARY

"We must learn to measure what we value rather than valuing what we can easily measure." — John MacBeath

In its 2016–2018 Report on the State and Needs of Education, *Evaluating So It Truly Counts*, the Conseil supérieur de l'éducation examines the goals of evaluating learning, from pre-school all the way to university, including vocational education and adult education. Evaluation takes up a lot of time, can cause a lot of stress and comes with its fair share of frustration. With this Report, the Conseil hopes to dispel some myths and contribute to the dialogue necessary to establish practices better aligned with and more supportive of evaluation's intended goals, both over the course of and upon completion of learning.

As outlined in current programs and policies, evaluation should be a constructive and insightful process that enables the individual being assessed to advance as well be part of it. Yet **reduced to grades and examinations** for all intents and purposes, evaluation continues to be a generally negative experience (one is subjected to it), and perceived as a verdict handed down from whoever is doing the teaching. Consequently, pupils and students—who are not always clear on what is expected of them—often try to guess the right answer rather than try to understand…

In *Evaluating So It Truly Counts*, the Conseil focuses in particular on the context producing the problems observed, and presents guidelines to help make evaluation a more useful tool for both teaching and learning.

Why and How to Evaluate

In education, the evaluation of learning serves two broad goals: to **support learning** (for the development of an individual's full potential) and to **certify achievement** (to publicly and transparently attest to and ensure the validity of diplomas).

A criterion-referenced evaluation makes it possible for these goals to be adequately achieved, as it places the pupil or student on a scale of what is expected:

✓ During the course of learning, where a criterion-referenced evaluation (combined with quality feedback) allows room for mistakes, shows

> **Criterion-referenced evaluation** is an assessment method where an individual's performance of a specific task is measured against a success criterion or threshold, formulated in one or more clearly targeted objectives, independently of the performance of any other individual (Legendre, 2005, pp. 577-578, free translation). Criterion-based evaluation does not rank individuals in relation to each other.

learners being assessed where they stand in relation to set targets, what they can do to improve, as well as the strengths they can build on.

✓ At the end of a given period (e.g. cycle, course or internship), where a criterion-referenced evaluation uses defined performance standards (rather than in cumulative points or comparing with other students) to assess whether learners have acquired the skills needed to attain certification or advance to the next level.

Additionally, as criterion-referenced evaluation is built on known and understood expectations, it can help pupils and students develop a sense of responsibility as well as self-assessment and self-correction skills, thus making sustainable and life-long learning possible.

Educational Goals Overshadowed by Administrative Functions

At an administrative level, the evaluation of learning also serves a number of functions: the management of the system, ranking and selection, for example. In the context of accountability, however, the Conseil notes that in fulfilling these administrative functions it can be all too easy to lose sight of educational goals, creating imbalances, confusing evaluation with grading, and ultimately diverting evaluation from its fundamental goals.

While traditional norm-referenced evaluation practices (assigning grades by cumulative points) might properly tick the boxes on a number of management functions, they do not play their intended role on an educational level. Indeed, the mechanics of the report card are feeding an unhealthy relationship with mistakes and making evaluation a tedious task for teachers. Grades, class averages and report cards are currently receiving the lion's share of attention yet do not clearly impart what individuals have achieved and what they still need to acquire, consequently making it difficult to support learning and not necessarily presenting an accurate picture of achievement. Moreover, grading practices prematurely rank individuals, which is causing major drifts in evaluation.

> **Norm-referenced evaluation** is an assessment method where an individual's performance is compared to that of others in a peer group using the same measurement. (Legendre, 2005, p. 584, free translation) Norm-referenced evaluation refers to grading on a curve and peer group averages.

Grades: A Flawed Symbol

Although they appear to be ingrained in the evaluation of learning and are currently an important benchmark for society and many parents, percentage grades derived from cumulative points are not guided by any pedagogical need. They are simply a symbol (and a far from perfect one at that) through which the outcomes of assessments are expressed in a summative way. In fact, percentage grades:

✓ Have scant informational value. Indeed, is there a quantifiable difference between 58% and 61%? What information do these figures provide about what one learner is missing that the other is not?

- ✓ Help convey a technicist, numeric view of evaluation, with cumulative point totals replacing professional judgement, and learners not likely to be motivated unless "it counts."
- ✓ Obtained from cumulative points indelibly penalize pupils or students for mistakes, something which is at odds with a competency-based approach and can have a significant counterproductive effect on motivation and perseverance among those who experience more challenging beginnings. Education should be visualized as a marathon to be finished, not as a sprinting race to be won.

Despite these limitations, grades have garnered societal importance—becoming an end in and of themselves—that tend to make it easy to lose sight of the true meaning of **evaluation**: not one synonymous with graded school work (assignments, exams, etc.), but one that **must contribute to learning by being an integral part of it.**

Moreover, grading practices could be much more efficient and not handed down to learners like a verdict. For example, a qualitative criterion-referenced scale provides information on what is expected in terms of levels of progress. Learners would thus know where they stand in relation to these expectations and also be able to choose which direction they would like to go or what programs would best match their interests, potential, and effort they are willing to invest. Efficient grading practices also do not pit learners against one another, nor are they used to punish behaviour.

It is worth remembering that without any supporting feedback, 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.

Supporting a Paradigm Shift

To support learning, evaluation must therefore be placed at its very core. This may amount to a major paradigm shift for those doing the teaching: **supporting pupils or students in grasping content is indeed much more complex than simply transferring this content.** In other words, the skill to evaluate is exercised daily by observing the progress and difficulties of **pupils or students** (reinforced with the appropriate feedback), rather than being reduced to preparing assessment tools. This competency needs to become a cornerstone in teacher training.

The Conseil has noted, however, that teacher training programs tend to include very few courses specifically devoted to the evaluation of learning. The program-based approach can still be useful in ensuring that this competency can be acquired in other components of the programs, notably courses in pedagogy and internships. For their part, colleges and universities do not require their faculty members to have teacher training. Post-secondary institutions typically provide these members with pedagogical tools and resources to support their professional development, albeit on a voluntary basis. In fact, ongoing training is not mandatory in either compulsory or post-secondary education, and when it is provided, it is often offered piecemeal, without any real consideration for continuity.

Challenges for Every Level and Sector of Education

The context in which the responsibility for evaluating learning is exercised does not make it easy to keep a steady focus on intended goals, for a number of reasons that vary depending on the level or sector of instruction. Here are a few examples for each:

Current guidance documents applicable to General Education in the Youth Sector have been planned in silos: frameworks for the evaluation of learning and the provincial report card were designed around cumulative points, while the Québec Education Program and the Policy on the Evaluation of Learning place greater emphasis on the goal of evaluation as a support for learning. Moreover, the Conseil has noted that many teachers felt there was a discrepancy between the Program's objectives, the Policy's underlying principles, report card rules and certification requirements.

In Vocational Training, guidance documents all deliver the same message, namely that evaluation throughout the course of training needs to contribute to learning and not be a tabulation used solely for the purpose of certification. It nevertheless appears crucial that teachers be made more aware of and be equipped for evaluation that supports learning throughout training, to better position them to support learners until graduation.

In Adult General Education, new programs are currently being phased in, requiring the adoption of different evaluation practices, notably the use of feedback during learning. This new reality highlights the need for teacher training and support. There is major change taking place here—an opportunity not to be missed.

In College Education, cumulative points are still very much being used, and the need to feed the R score poses a significant hurdle for evaluation practices to be consistent with a competency-based approach. The dominance of the R score as a university selection criterion right off the bat eliminates excellent candidates for post-secondary studies without necessarily ensuring that those who are selected would fit the target profile. The R score is the culmination of a system where evaluation is dependent on a selection that favours candidates who benefit from more advantageous family and socio-economic conditions from the outset.

In University Education, the large number of students enrolled in a course is a major challenge. Undergraduate courses are increasingly attended by well over one hundred students, making it difficult to adopt evaluation practices that support learning and assess complex tasks and high-level competencies.

Key Findings

The work the Conseil carried out to produce its 2016-2018 Report on the State and Needs of Education revealed that in general, teachers do not question the goal of evaluation to support learning. Yet this goal appears to be difficult to put into practice, namely because not all current guidance documents deliver the same message or the favourable conditions to do so are lacking.

Criterion-referenced evaluation integrated into learning is the primary evaluation tool teachers have that can support learning and thus lead each individual to develop his or her abilities to the fullest (i.e. raising everyone a notch higher). Criterion-referenced evaluation can therefore facilitate **setting the bar higher for each individual**, without necessarily having the same expectations for everyone.

Yet in response to some societal pressures and given the current context, **when it comes to evaluation the education community is in a defensive stance.** Teachers gather evidence and might be tempted to stick to what is easier to measure, to the detriment of what is important. As for pupils and students, they tend not to take any risks (to avoid losing marks) and might limit themselves to superficial and short-term learning (e.g. memorizing answers for an exam).

Favourable conditions for **moving from norm-referenced evaluation to criterion-referenced evaluation** (namely a fair societal view of evaluation, cohesive guidance documents, teacher support) are currently lacking. Taking ownership of evaluation criteria is a shared responsibility of teachers; it is therefore vital to free up time and earmark budgets for this collaborative task, as well as have management exercise educational leadership in supporting it.

In addition, because criterion-referenced evaluation is not currently an integral part of learning, it is seen as time taken away from the latter. **Time is a double challenge**, however:

- ✓ Respecting individual paces is easier when a school organization shows flexibility (the current leeway permitted by guidance documents is not always used);
- ✓ Evaluation integrated into classroom activities (ongoing observation) should make it possible to lessen the need for formal assessments and examinations.

As traditional evaluation pits individuals against one another and encourages superficial learning strategies, it does not foster the **development of 21st century skills** (the ability to take risks, learn by trial and error, teamwork). Admittedly, the ranking that results facilitates **selection** when needed, however this is a **specific need that makes it easy to lose sight of the real goals of evaluating learning**.

Proposed Guidelines

The Conseil wishes to reaffirm that **support for learning**, which enables learners notably to develop a sense of responsibility and acquire a realistic view of their potential, **is a fundamental goal of evaluation**, as important as an evaluation that attests to achievement (certification). There is a discrepancy, however, between these goals and the practices observed (norm-referenced evaluation in a defensive mode), which lead to ranking learners and seem to be dependant on the selection process.

For the development of evaluation practices adapted to intended goals, the Conseil supérieur de l'éducation has formulated three key guidelines:

Realign Evaluation of Learning With its Goals

- ✓ Clarify the goals of evaluation and challenge myths that lead to overvaluing the importance of grades (or confusing evaluation with grading).
- ✓ Strive for a position of support between extremes that tend to be opposing: meritocracy (elitism) and complacency (a race to the bottom).

Create Favourable Conditions for Moving Towards Criterion-Referenced Evaluation

- ✓ Develop collective expertise and foster collaboration between teachers, guidance counsellors, and school and university management (freeing up time needed to adopt criteria and allocate appropriate budgets).
- ✓ Rethink the rationale of professional development for teaching staff: the required change rests on structured support over a sufficiently long period of time.
- ✓ Reassess the way grades are assigned and for which report cards and transcripts are created, use other means of communication.
- ✓ Reconsider the importance of grades in the selection process.

Build a Positive Relationship with Evaluation

- ✓ Break away from a vision of evaluation reduced to one of assessments and examinations.
- ✓ Communicate the outcomes of evaluation without automatically ranking learners.
- ✓ Empower pupils or students so they can align their strategies with their goals and fields of interest (involving them in the process and moving past superficial learning).
- ✓ Use ministerial examinations solely for the purpose of steering the system.

Conclusion

The Conseil has noted that in compulsory education, the major drifts in the evaluation of learning (overinflating the value of grades and rankings) appears to be linked to selection requirements from higher levels of education. The Policy on Educational Success announced an initiative "toward the modernization of pedagogical frameworks and procedures for the evaluation of learning" (MEES, 2017). The Conseil sees in this an opportunity to update, streamline or drop some documents, question the current percentage grade report card and ensure that all guidance documents in the education system are consistent with criterion-referenced evaluation practices.

In addition, to realign the evaluation of learning with its goals, the Conseil calls on the entire education system to adopt a paradigm shift, in other words, apply a rationale of observation and criterion-referenced evaluation based on feedback to efficiently support learning, help learners develop a sense of responsibility and better attest to achievements—all as early as elementary school.

In doing so, pupils and students can build a positive relationship with learning and evaluation much earlier on their education path, one free from the competitiveness of norm-referenced evaluation and looking toward the goals of achievement. In this process they will also have learned how to work in teams, take risks, learn from their mistakes, ask for help and express their needs.

These competencies, 21st century skills all, would trade the practice of having to memorize content the day before an exam simply to get a better mark than someone else for a more personal, sustainable success. At the same time, learners would also be able to make better targeted vocational choices. Those wishing to pursue studies further would know where their weaknesses lie and what strengths they can build on.

So it truly counts, this paradigm shift—and not the traditional notion of ranking—must become the guiding principle of evaluating learning along the education continuum. Teachers and faculty at every level and sector of the system should no longer continue to evaluate as they themselves had once been… Undoubtedly, support for this change of course will greatly depend on a better understanding of the goals of evaluation on the part of society and a long-term vision supported by policymakers.

Bibliography

Legendre, Rénald (2005) *Dictionnaire actuel de l'éducation*, 3rd Ed. Montréal: Guérin.

Ministère de l'Éducation et de l'Enseignement supérieur (2017) *Politique de la réussite éducative: le plaisir d'apprendre, la chance de réussir.* [Policy on Educational Success: A Love of Learning, a Chance to Succeed] Québec: Le Ministère. Retrieved April 9, 2018, from http://www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/politiques_orientations/politique_reussite_educative_10juillet_F_1.pdf.