

How can we tell that students have learned? Why traditional exams and grades do more harm than good

Joana Stella Kompa, May 2017



Picture: College students taking an exam, Credit: Chris Ryan via Getty Images

The ideas behind traditional exams and grading

In traditional teaching, a teacher presents a learning unit by introducing new knowledge to the class via a series of lectures and presentations. Typically, brief question and answer sessions allow students to probe, at least to some extent, what they have not fully understood. In addition, pupils are given homework to apply the new knowledge in given exercises. At the end of the learning unit, an exam or test verifies the learner's competence to replicate and apply the new knowledge. The resulting individual grade is regarded as a reliable and truthful standardised assessment of a learner's competence of achieving the stipulated learning outcomes.

This is, in essence, the brief idea behind grading which was developed since the 16th century and gained its momentum in the 19th century with the introduction of compulsory education. The nagging question that is debated among educators is if standardised grading is actually measuring students' learning and if it is not, what else it is measuring. Another question is how useful or even harmful standardised assessment plays out in real life.

What is assessed and how?

Constructivist learning pedagogy promotes active learning and assesses not only **cognitive abilities** (such as the application of specified mental operations to a set of varied problems), but includes furthermore the measure of students' **study skills, their individual ways of learning, verbal and nonverbal communication skills, social competencies, intrinsic motivation and metacognitive reasoning** (the ability to reflect critically on mental content, methodology, knowledge creation and its meaning for the learner). By comparison, traditional exams only

assess the minuscule part of cognitive abilities, or the ability to temporarily store subject knowledge in short-term memory, while ignoring all other relevant competencies.

In order to conduct sensible educational measurement, lecturers need to establish baselines of such competencies in order to understand students' varying prior conditions. A series of semi-structured interviews combined with preliminary tests or exercises can reliably provide information on students' prior conditions. As sociological research has demonstrated, students from academic families fare generally better as compared to students from working class families or students brought up by a single-parent. Socio-economic settings as well as age- and family-related factors set students at distinctively different starting levels for their studies. The responsibility of a fair education is to mediate and diminish such differences.

Let's assume we have two struggling students with similar aptitude trying to pass the next exam. Student A comes back home from school to a household with two younger hyperactive siblings that deprive her of any opportunity to study and focus, whereby student B is luckier and has parents that can afford to send her to tutoring classes. In the final exam, student A fails while student B has achieved a decent grade. But what was measured by the exam was not a superior cognitive ability of student B, but a given social advantage. Or we may imagine a student with brilliant ideas, but incapable of time management to organise them, or a younger student who has exactly the same mental potential as everyone else, but is distracted by joining his friends on the football field while neglecting his studies. Or we can imagine a highly intelligent student who fails by not being able to handle exam anxiety, and so on and so forth. In most cases, we evaluate the influence of confounding factors on learning, but not learning itself.

Grades are not only unsuitable as an impartial evaluation tool, they also interfere with the very motivation to learn. Many students, sadly enough, learn in order to achieve good grades, not because they enjoy learning something new, exciting or useful. The main learning outcome is often not the acquisition of competencies, personal growth or new knowledge and skills, but a good grade. This is the point where the traditional learning system has turned ad absurdum: when the reward for learning is represented by the affirmation of prior social status via grades, true learning has lost its relevance as a driver for the development of young people. In many schools, colleges and universities, grades have become the symbolic trophies for representing achievement. What was originally intended only as a tool to evaluate learning outcomes by a simple scale has turned into a central outcome by itself. If education was strict science, grades would be removed as a confounding variable from the setup in an instant.

We could compare grades to political polls. Originally designed to objectively measure political trends within a population, polls have become a strategic tool for political parties to influence their voters. By employing polls, perceived instrumental threats and opportunities start to govern decision-making rather than good policies, arguments and well thought-out concepts. The same is true for students who calculate their minimum attendance requirements and bare pass investment in studies.

Tracking progress: What sound assessment entails

Institutions that blindly grade students based on standardised tests measure de facto a plethora of prior conditions, rather than learning. Measuring actual learning progress would require the implementation of two more conditions. Firstly, that students need to be provided **(a) adequate formative feedback for self-improvement** on above-mentioned competencies and **(b) follow-up investigations by teachers to verify if and how formative feedback was successfully internalised**. As weekly assignments and projects continue, actual learning progress can be measured and facilitated efficiently in this manner.

Students are coached continuously and they understand how they have been assessed on multiple levels. A formative and reflected approach is obviously more helpful to fine tune ongoing learning than receiving a single summative grade at the very end. New technologies such as online feedback can speed up and further specify ongoing improvement.

Another harmful effect of grades is that they condition motivation. A weaker student may find in low grades the confirmation to be a 'born loser' or a 'failure', while good students can find reassurance to belong to a class of eternal winners. Grades are often perceived as a judgment of the Global Self, especially among adolescents, which explains how low grades tend to diminish self-esteem. In this light, grades represent a cruel tool to retroactively punish weaker students for their compromised position in life. On the brighter side, many teachers find that they do not need traditional tests. When conducting, e.g., more complex interdisciplinary projects, tests become redundant. It would be fairly ridiculous, for example, to conduct a traditional test at the end of a group's research project, which would not only trade an information-rich assessment for an information-poor one but would be superfluous as learning outcomes and assessment have already been achieved and documented.

Alternatives to traditional assessment:

Multi-perspective evaluation and scoring rubrics

This leaves progressive educators with at least two proven options for a more learning-centered evaluation. One is to assess projects in a PBL-like manner, which includes self-evaluation as well as the evaluation of others: How have I and others performed as team members, as problem-solvers and as researchers? How do we assess the shared learning process and outcome of a project? Here is also the opportunity for reflective journals that worked well with my students at Temasek Polytechnic in Singapore. The PBL assessment emphasises students' social roles in interaction with cognitive and metacognitive reasoning. This is why this approach is well suited to postsecondary education in situations where discussions with students and their facilitators/coaches play a central role.

The other option is to work with scoring rubrics that many educators are familiar with. The advantage of rubrics is that students know in greater detail how and why they are being assessed. Expectations are made clear from the beginning which makes this a more structured approach. Students also learn that explaining phenomena, relating facts to ideas and integrating knowledge is of greater value than to cut, paste and simply summarise information, a common habit among digital natives. In many institutions, rubrics are still translated into grades, but they are still a far cry from simple point and error accumulations followed by a final grade and perhaps a brief commentary by the teacher. Both a PBL-like approach as well as scoring rubrics assist

students' *learning to learn*, which is why they are preferable to traditional grading. The desired procedural learning outcome is that students achieve *mastery in learning*, one of the likely reasons why some of my best students became teachers themselves.

The Bell Curve paradigm

A more progressive assessment deviates from the assumption of a static Gaussian standard distribution also known as the 'Bell Curve'. It represents the idea that typically some students are naturally at the top, some at the bottom and most positioned in the middle of the spectrum of abilities. At the beginning of a course, student levels might indeed be represented by a standard distribution. However, if performance levels remain unchanged for an entire term then a Bell Curve signals that learning among students has not taken place. Students that were good at the beginning are still good at the end of the semester while students who failed in the beginning still fail towards the end. A static Bell Curve is a reliable indicator of the fact that the education system or the teacher has failed. **We can only tell how students have learned once we can demonstrate how constructive and motivating feedback has fostered their autonomy, contributed to their personal and social development and has built their competencies.** Without such provisions in place, blindly conducting tests and grading papers tells us nothing about how students have learned. In the traditional assessment, the learning process remains unexamined simply because the dependent variable of learning progress is not related to the independent variables of the learning environment and other controlled factors facilitating learning.

Development as a whole person

As an empirically-based domain, education should not fall into the traps of confirmation bias by repeatedly verifying prior conditions or conditioned responses. Education should elevate weaker students while providing new opportunities to stronger students. Every student is deserving of empowerment to develop into a strong, self-directed and lifelong learner.

But how about students that appear obviously lazy, who display negative attitudes towards their teachers or the institution and simply act out anti-socially in every perceivable way? One of the advantages of a multi-level assessment is that it includes personal factors. Difficult students can be counselled, which is better than failing them and pushing them over the edges of society. This entails of course that institutions of Higher Learning employ professional counsellors and that personal development is taken seriously as a hallmark of quality education - to connect to the ideal of a holistic humanistic education as first envisioned by Wilhelm von Humboldt (1767-1835) or the idea of developing one's personality as first outlined by Immanuel Kant (1724-1804). From a constructivist perspective, exams and grades represent a poor excuse for not understanding and not facilitating students' learning.