

Characteristics of effective summative assessment

We know that summative assessment drives practice in schools. We also know that current forms of summative assessment inhibit both curriculum and pedagogical innovation because of their focus on 'knowledge' (as viewed in [a knowledge based curriculum](#)¹). The challenge is to find new forms of summative assessment which satisfy the criteria against which they will be judged. Those criteria should include:

Validity and reliability

At its most basic validity is the extent to which the assessment measures what it is intended to measure. There are many facets of this, including:

- does it measure what it appears to measure (e.g. if a science exam requires extensive reading and writing competence then it may be testing literacy rather than science);
- does it measure the things that are important (i.e. the things we want children to have learnt).

Reliability is about the extent to which the outcomes are both accurate and consistent (e.g. would two people who are equally competent in the thing that is being measured end up with the same result every time?).

The validity and reliability of an assessment capture the extent to which it is **relevant** and **credible**. Critically, unless an assessment is measuring the things that are important then it can't be effective, no matter how reliable and valid in other ways it might be. It has to be measuring something that is relevant as well as being credible.

Practical and scalable

Assessment has to be practical in the sense of it being feasible to implement. The more time consuming, expensive and operationally difficult it is to carry out an assessment the less likely it is to be used.

Within the formal education system large numbers of learners need to be assessed. The more practical the assessment is then the more scalable it is likely to be.

Assessment that is integral to the learning process is more practical than assessment that is an additional discrete activity.



Formative and positive

Even when the key objective of assessment is summative it should also inform future learning (i.e. be formative). Summative assessments that serve a formative purpose are better than those that do not.

Whilst criticism can enhance learning, it normally only does so when it is constructive. Effective assessment should also be constructive, for example by focussing on successes rather than failures.

¹ <https://halfbaked.education/knowledge-based-curriculum/>

Ethical

Clearly, it is important that every element of education, including assessment, is ethical.

It would be unethical for assessment to damage one person in order to benefit another. It should therefore be criterion rather than norm referenced (because norm referencing dooms some people to fail in order for others to succeed).

Ethical assessment should enable everyone to succeed - through being at an appropriate level for each learner and/or through the provision of multiple pathways to success.

It would also be unethical to assess people covertly, for example collecting data about them without their knowledge or consent. Indeed, the way in which assessment outcomes are determined should be transparent - assessment should not be a black box.

Single use

Finally, the way in which an assessment is used is critical. Any one assessment should be being used for one clearly identified purpose. If the aim is to assess student achievements then that should be the only purpose for which the assessment is used - such assessments should NOT also be being used to evaluate the quality of educational provision (teaching/the school).

Where an assessment is being used both to assess the students' learning and the quality of education provision this often leads to gaming of the system. For example:

- schools offering study options that maximise the benefits for the school rather than for the student;
- teachers increasing pressure on students to do well because the teacher is concerned about how the assessment outcomes will reflect on their performance;
- schools targeting support on students who are just below grade boundaries because that is more beneficial for the school than trying to raise all students' grades.

In some senses this overlaps with the issue of validity - the assessment measuring what it claims to measure (e.g. students performance OR teaching quality). However, it is such a critical issue that it needs to be pulled out as a separate criterion. Many of the worst consequences of current summative assessment practices are related to dual use of summative assessment outcomes data.

Easily reported

As the key element of summative assessment is to evaluate what someone has learnt up to that point in time, having a concise summary of the outcomes of the assessment is important. This facilitates comparison (with their previous performance and/or with external standards and/or with other learners).

Conclusion

Devising assessments that meet all of these criteria is incredibly difficult, not least because there may be conflicts between the criteria. However, we surely must be able to do better than continue to use norm-referenced, paper-based terminal exams that only result in a grade (e.g. A to F or 9 to 1).



Cite as:

Twining, P. (2020) *Characteristics of effective summative assessment*. Halfbaked.education blog, 4th November 2020.

<https://halfbaked.education/characteristics-of-effective-summative-assessment/>

© Peter Twining [CC-BY-NC 3.0](https://creativecommons.org/licenses/by-nc/3.0/)²

² <https://creativecommons.org/licenses/by-nc/3.0/>