

Chapter Unit 5.2 Assessment and Learning: Summative Approaches

Book: Learning to Teach in the Primary School

Year 2025

Paula Lehane  <https://orcid.org/0000-0003-0856-3505>

Alan Gorman  <https://orcid.org/0000-0002-5024-448X>

Abstract

This chapter provides opportunities to reflect on what summative assessment is, how it is used in primary schools, and the impact that it has on learners. The chapter considers a range of questions about summative assessment in order to fully understand its purpose and how it can be effectively used in classrooms. Alongside discussing purposes of summative assessment, the chapter also identifies sources of assessment evidence. Considerations are also given as to what counts as good evidence of learning. The chapter also presents an overview of current policies on summative assessment and reporting procedures. Comparisons are also made with formative assessment. The chapter concludes by inviting the views of readers on current assessment policy and practice.

Introduction

In this unit, you will have the chance to reflect on what summative assessment is, its uses and its potential impact on learners. You will also be able to consider some aspects of current policy on assessment. We start by considering some basic questions about summative assessment and by linking it with formative assessment. We will go on to identify purposes of summative assessment, as well as sources of assessment evidence and we will explain what counts as good evidence of learning. We also describe current policy on assessment and reporting. We finish by inviting your views on current assessment policy and practice.

Objectives

By the end of this unit, you should be able to:

- define and explain the potential purposes of summative assessment;
- understand the relationship between validity, reliability and fairness;
- know the kinds of assessment tasks that are effective in generating good evidence of learning;
- describe some aspects of policy on assessment and offer an informed opinion about the current emphasis on different assessment purposes and approaches.

What is assessment and why do we do it?

Assessment refers to the process of gathering, interpreting, and using information about a child's learning. It helps teachers to identify what their learners know, understand and can do. Without assessment, a teacher would be unable to identify what prior learning has occurred leading to a unit of work that could be full of repetition or inappropriate content. Assessment is therefore a crucial part of teaching and learning. It is not something that should ever be seen as 'separate' from teaching. However, assessing learning is not a neutral or value-free activity – it is always bound up with attitudes, values, beliefs and sometimes prejudices on the part of those carrying out the assessment and on the part of those being assessed. When we make assessments of children's learning, we are always influenced by what we bring with us in terms of our previous experiences, personal views and histories. Children's responses to assessment are influenced by what they bring with them – their previous experiences and their personal views.

Summative assessment 'sums up' learning

Discussions on assessment approaches tend to refer to two important types: formative and summative assessment. Sometimes, these are also called Assessment for Learning (AfL) and Assessment of Learning (AoL). The previous chapter explores formative assessment or AfL in more detail. We will focus on summative assessment or AoL approaches. Assessments that claim to serve a summative purpose generally focus on grading or ranking students. Summative assessment is retrospective: it evaluates the knowledge or skills that have been achieved by a certain point in time e.g. at the end of Key Stage 2. Gathering the evidence necessary for such evaluations and judgments can involve a range of assessment tools such as tests, portfolios or presentations.

Purposes of Summative Assessment

Summative assessment is carried out for several purposes. When designed by you, the teacher, it can provide you with a summary of learners' achievements that will inform your future teaching and, of course, your planning for future learning (This is close to the notion of formative assessment described in the previous chapter). Second, it provides information that can be shared with others about a learner's progress e.g. parents. Thirdly, summatively assessing learning can provide a numerical measurement that can be used in league tables. In England, these allow comparisons of achievement across schools within an education system. These usually involve formal exams or standardised tests that are not set by the classroom teacher. Finally, summative assessments are often used for certification or selection purposes, particularly in post-primary education. For example, A-Levels in the UK are used to certify a student's skills and understanding at the end of second-level education.

As you can see from these examples, the tools and methods used in summative assessments can sometimes be very far removed from the day-to-day work of teachers. How isolated the method for summative assessment is from the daily running of the classroom changes the 'stakes' of that assessment. Indeed, summative assessment is often discussed using the terms 'high-stakes' or 'low-stakes'. For example, A-Levels examinations are considered 'high-stakes' as they can be used to select students for further and higher education. The outcomes on these examinations can have a significant impact on a learner's future. End-of-term tests and examinations set by schools or teachers are relatively low-stakes in comparison as the consequences for the learner are relatively short-term. Murchan and Shiel (2018) note that international assessments like the Programme for International Student Assessments (PISA) can be considered both low-stakes and high-stakes. As these assessments do not report results for individual schools and learners, they are quite low-stakes for them. Yet, they are very high-stakes for each country's or government's department of education as they can be used to rank how 'successful' a country's education system is.

Think about it... PISA Shock

The Programme for International Student Assessments (PISA) is an international large-scale assessment that is administered every three years to representative national samples of 15-year-olds. It aims to assess students' 'preparedness' for their future lives based on their performance

in tests of reading literacy, mathematics and science. In 2022, eighty-one countries took part in PISA.

Many countries (rightly or wrongly) use PISA to help evaluate their education system's effectiveness. Pizmony-Levy & Bjorklund (2018, p. 251), drew on survey data from 30 countries and found 'evidence that national performance on PISA has a significant positive relationship to public confidence in education'. Therefore, it's not surprising that a country's performance in PISA and what 'ranking' they receive can often cause significant debate and discussion in the media. The poor performance of a country in PISA can lead to something called 'PISA shock'. This refers to the *reaction* of policy-makers and the public to student performance in PISA (Niemann et al., 2017). For example, when the first PISA results were released in 2000 the perceived poor performance of German children led to significant changes in education policy. Ireland also experienced something akin to 'PISA Shock' when a decline in performance in literacy and mathematics in 2009 'shocked' policy makers into launching a new 'National Strategy for Literacy and Numeracy'.

Think... Do you think that using PISA performance to evaluate a country's education system is a good idea? What are the potential advantages or disadvantages of such an approach?

Producing Good Evidence of Achievement

Characteristics of Assessment

It is important to appreciate that summative assessment can take a variety of forms – it need not just be a written test. This is only one tool that we have at our disposal. Before we investigate these tools, three key characteristics of assessment should first be considered: validity, reliability and fairness. While these can be relatively abstract ideas, we will consider them in the context of Mr. Murphy's Year 1 class of 25 students.

Validity refers to whether the information being gathered can support the judgment a teacher needs to make about the learner's progress. When designing a summative assessment, teachers should ask themselves if they are collecting the 'right' information for what they want to know about their learners. Imagine if Mr. Murphy wants to understand how much his students know about subtraction so that he can plan his next week of work in Maths. He gives his students

a short ten-question test. However, unbeknownst to him, he picks up the wrong test sheet. All of the questions he gives his learners are about subtraction *but* they involve renaming – Mr. Murphy has not yet taught his class how to do subtraction with renaming. Mr. Murphy has asked the wrong questions for what he wants to know. He might be able to glean some information about his students’ understanding of subtraction generally but not much beyond that. If he finds that his students perform poorly on the test (because they have yet to be taught how to do sums that involve renaming) and groans to his friend “they know nothing about subtraction!” we might have to stop Mr. Murphy and remind him that he’s not drawing an appropriate inference about his learners from this assessment source.

We should try to remember three key things about validity from this example:

- Validity takes into account how appropriate the information being gathered is for a teacher’s particular *purpose*.
- Validity has to do with *inferences* i.e. the conclusions you draw from the information you gather.
- Validity is not an ‘all or nothing’ thing – it’s a matter of *degree*. The inferences we draw from an assessment can be more or less valid depending on the situation.

A second important characteristic of summative assessments is reliability. Reliability is broadly concerned with accuracy and consistency. When we think about this characteristic of assessment, we are trying to determine whether or not the same results would have been obtained if the assessment had been given at a different time or in a different context. Let us go back into Mr. Murphy’s classroom again. Mr. Murphy has a student called Fred. Based on the written work he has reviewed from Fred’s history, geography and English essays for that month, Mr. Murphy has noted Fred’s written work has a number of errors in it; mainly in relation to punctuation and sentence length. By gathering multiple examples of Fred’s written work, Mr. Murphy can have some confidence in the reliability of his observations from his summative assessment tasks. This is because no single assessment can provide a ‘perfect’ overview of a learner’s knowledge, skills or abilities.

A fair assessment is one that allows all learners an opportunity to demonstrate their achievement. If an assessment is discriminatory or biased towards a learner, then an inappropriate inference may be drawn about their knowledge, skills or abilities. Common sources of assessment

bias include racial/ethnic bias, gender bias and socio-economic bias. Fairness is also seen in assessments that provide learners with an adequate opportunity to learn the content being assessed (unlike what happened with Mr. Murphy's earlier Maths test). Let us consider another example of this in relation to computer-based exams. If Mr. Murphy's students are required to complete a computer-based test for English writing, then all of his students should be familiar with the technology required. If they must type a response when they would normally handwrite their work in class, their lack of experience or proficiency in using a word processor could unfairly influence or bias their performance.

Think about it... Differentiating assessments

To support valid, fair and reliable inferences about the diverse range of learners you will be working with in your future classrooms, some adaptations may be necessary to your summative assessment approaches.

Accommodations are those adaptations to an assessment that do not fundamentally change what you are assessing. These changes focus mainly on the *how* of the assessment. This may involve using a larger text size in tests of reading comprehension to support those with a visual impairment. Accommodations often involve changes to the presentation of an assessment e.g. allowing extra time or how learners can respond. However, if there are changes to *what* is being assessed, then it is more appropriate to call it a modification. Modifications often change the content of the assessment e.g. asking a learner less questions on a maths test. It is important to ensure that *what* you are assessing does not fundamentally change because of your adaptations or modifications. For example, when assessing *reading* comprehension, teachers may decide to read out a text to a learner who cannot independently read long passages. This change by the teacher has altered the assessment – the teacher is no longer assessing reading comprehension. They are now assessing listening comprehension.

Think... A teacher is assessing her students' spelling skills in Year 4. For most of her students, she uses the same spelling list. However, for other students she uses a different spelling list. Is this an example of an accommodation or a modification?

Summative Assessment Approaches

Based on the above, we recommend that you think about three key questions that will help to guide your approach to any assessment, including summative assessment.

1. What do I want to know from the information I will gather?
2. How will I gather enough information to have an accurate and fair overview of the specific knowledge, skill or understanding of interest?
3. How well does the information gathered justify my inferences or future actions?

(adapted from Burke & Lehane, 2023, p. 76)

Summative assessments can include a broad range of approaches. We provide a brief overview of the following: Classroom Tests (Teacher-Designed), Performance-Based Assessments/Portfolios, Rubrics, Standardised Assessments and Examinations. It is important to note that there should be a close match between the learning being assessed and the methods used to assess it. This should guide which of the approaches you use. Furthermore, no one assessment approach is ‘perfect’. In a modern primary classroom, a mix of summative assessment approaches should be used to understand learners’ achievements and what ‘next steps’ are needed.

Classroom Tests (Teacher-Designed)

The word ‘test’ can be used to describe any written (or oral) assignment that consists of a series of questions/problems that must be answered by individual learners within a limited time frame in a classroom context (Brookhart, 2015). Teachers should always remember that tests can only assess a *sample* of what students have learned. It would be impossible to create an assessment that can address every learning target/objective/outcome. A range of item (question) types are available to teachers – constructed response (e.g. short answers, essays, completing maths problems), selected response (e.g. multiple choice, matching, binary choice (true/false, yes/no) and cloze procedures. See below for some examples of each.

Think about... Question Types
Imagine that you have just completed a unit of work about the Amazon rainforest. Here are some sample questions that you could include in your classroom test.
<i>Multiple Choice Questions (MCQs)</i>
What is the top layer of a tropical rainforest called?

- A. Canopy
- B. Understory
- C. Emergent
- D. Forest Floor

Fill in the Blanks (Cloze)

The _____ layer sits below the emergent layer. Trees have _____ to this environment by having glossy leaves and needing less sunlight.

Constructed Response

Rainforests are found in countries near the equator. What impact does this have on the climate there?

Think... Could images or illustrations be used in any of these questions?

When trying to decide what questions to include in your classroom test, it is good to consider how your test will be administered (e.g. paper, digital, oral), what can be done to minimise learner anxiety (e.g. teach test-taking skills) and how many questions you need to ask to ensure your inferences are appropriate.

Performance Based Assessments/Portfolios

Performance-based assessments involve students applying their newly acquired knowledge and skills to a set task leading to a ‘performance’ or product that represents their learning (Brookhart, 2015). Many people argue that performance-based assessments are more ‘authentic’ than classroom tests but that really depends how much of the task is bounded by real-world contexts and constraints. Some examples of a performance-based assessment used in primary classrooms could include:

- Designing a house for a family of field mice
- Hopping, skipping, running or jumping
- Blending sounds to read a word
- Planning a trip to another country with a sample budget and itinerary

- Representing a character from a nursery rhyme in a drama
- Constructing a bar graph from the data provided
- Working in a group to create a multi-media presentation on an historical event

As you can see from this short list, performance-based tasks can be quite restricted (e.g. have only one correct response) or relatively open-ended (e.g. have multiple correct responses). Which one a teacher chooses is very dependent on the learning they wish to assess. Sometimes, performance-based assessments can be organised into a portfolio – a collection of ‘artefacts’ that can showcase the learner’s work and development over time. It is very common in subjects like visual arts, physical education, science and English. Portfolios often include reflective entries or annotations from the learner that offer insight into to their thinking.

Rubrics

Rubrics have been called many things depending on how they look – performance checklists, scoring criteria, rating scales. Regardless of the term used, they usually state the factors from which a learner’s work or performance can be evaluated. These criteria can be informed by the curriculum, teachers’ own knowledge/experience or learning progressions. Rubrics are not used to generate assessment information or evidence – they are used to evaluate them to support teacher inferences and judgements. Sometimes (particularly when grading essays), rubrics can be used to decide on a grade. These rubrics include descriptions indicating standards of attainment for different levels of performance, success, or competency (Brookhart, 2015). Two broad categories of rubrics exist – holistic and analytic. A holistic rubric is one in which each category of performance (e.g. excellent, very good, good, satisfactory) contains several criteria to help give an overall rating. An analytic rubric is one in which each ‘feature’ of the performance is identified and scored separately. The purpose of the assessment should help determine which form of rubric could be used. Both have advantages and disadvantages. The images below show how the same performance-based assessment could be evaluated using a holistic or analytic rubric.

Think about... Rubrics

Imagine that you are evaluating learners’ performance in a debate on the motion ‘Animals should not be in zoos’. Consider each of the following rubrics.

Rubric 1 (Analytical)

	Excellent	Good	Satisfactory	Needs Improvement
<i>Delivery</i>	Speaks consistently and clearly with excellent volume, variation, and inflection. Uses several persuasive techniques to excellent effect (e.g. humour, rhetorical questions). <i>10 points</i>	Speaks clearly with good volume, variation, and inflection. Uses some persuasive techniques effectively (e.g. humour, rhetorical questions). <i>7 points</i>	Speaks relatively clearly with appropriate but inconsistent volume, variation, and inflection. Shows knowledge of persuasive techniques but with little effect or impact (e.g. humour, rhetorical questions). <i>4 points</i>	Speaks in a low volume with little variation and inflection. Inadequate use or absence of persuasive techniques (e.g. humour, rhetorical questions). <i>2 points</i>
<i>Organisation</i>	Begins with a clear and purposeful statement of position. Provides multiple, highly relevant arguments. Contains an effective logical flow and structure. <i>10 points</i>	Begins with a statement of position. Provides relevant arguments. Contains a reasonably logical flow and structure. <i>7 points</i>	Statement of position is somewhat unclear. Small number of arguments provided; not all are relevant. Flow and structure are poorly organised. <i>4 points</i>	Does not begin with a clear position. Provides few or no real and relevant arguments. Difficult to understand the flow of the presentation due to inadequate structure and flow. <i>2 points</i>
<i>Content</i>	All arguments had specific evidence to support them (e.g. examples, statistics). All information was highly accurate and clear. <i>10 points</i>	Most arguments had some evidence to support them (e.g. examples, statistics). Information was accurate and clear but further elaboration was needed. <i>7 points</i>	Few arguments had evidence to support them (e.g. examples, statistics). Some inaccuracies in the information provided. <i>4 points</i>	Inadequate evidence provided to support arguments or statements (e.g. examples, statistics). Significant number of inaccuracies in the information provided. <i>2 points</i>
<i>Preparation</i>	Includes a wide range of sources. Provides well considered, elaborate and in-depth counterarguments. <i>10 points</i>	Includes a range of sources. Provides appropriate counterarguments with varying levels of depth. <i>7 points</i>	Includes a limited number of sources. Provides rudimentary counterarguments provided with little depth. <i>4 points</i>	Includes very few, if any, sources. Provides no counterarguments. <i>2 points</i>

Rubric 2 (Holistic)

Performance Levels	Success Criteria
Excellent <i>30-40 points</i>	Speaks clearly and consistently using a wide range of well-supported, evidence informed arguments. Presentation is very well organised and effectively includes a range of persuasive techniques. Offers excellent counterarguments and defends position very well when asked questions.
Good <i>20-29 points</i>	Speaks clearly using a range of well-supported arguments. Presentation is organised and includes some persuasive techniques. Offers appropriate counterarguments and can defend position when asked questions.
Satisfactory <i>10-19 points</i>	Speaks relatively clearly using a small number of arguments that had limited evidence to support them. Presentation is somewhat disorganised and includes very few persuasive techniques. Offers some counterarguments but can only offer a limited defence of their position when asked questions.
Needs Improvement <i><10 points</i>	Does not speak clearly or consistently with little reference to any arguments or evidence. Presentation is poorly organised and contains few or no persuasive techniques. Offers no counterarguments and is unable to defend position when asked questions.

Think...

- 1) Do you think rubrics can eliminate teacher bias?
- 2) What are you going to learn about each learner's performance using the analytic or the holistic rubrics? Do you favour one more than the other and why?

Standardised Assessments and Examinations

When people think of summative assessments, they usually think of public examinations or standardised assessments. While they can be useful to classroom teachers' practice (e.g. identifying learners who have specific needs), certain authors (e.g. McMillan, Shepard) do not consider these tools to be relevant to the field of classroom assessment as teachers often have very little to do with their creation or selection i.e. they are obliged to do it for accountability reasons

rather than for learning. However, it is still important for teachers to have a clear understanding of them.

Standardised tests assess the abilities of learners under very controlled conditions e.g. strict administration, scoring and interpretation guidelines. These tests also have to follow careful rules in terms of their construction. How these tests are constructed can help teachers to understand the *meaning* of learner outcomes in accordance with the interpretative frameworks used. Two main interpretative frameworks are associated with standardised tests: norm-referenced and criterion-referenced. Some tests can use both frameworks. Norm-referenced tests allow teachers to compare their students' performance to a 'norming group' using test norms. The following steps are involved in creating norms:

1. Select a relatively large and representative sample of students from the country i.e. students who are similar in age and other characteristics to those who will be using the test in the future. This is the norming group.
2. Administer the test to the norming group.
3. Score the test and examine the appropriateness of the content.
4. Analyse the scores and use them to represent the performance of all learners who will be taking the test.

For example, if learners in Mr. Murphy's class take a norm-referenced standardised maths test at the end of Key Stage 1, their performance will be compared to the performance of students who helped 'norm' this test. Performance in a norm-referenced test can be summarised using percentile ranks, scaled scores or standard scores (see below for a brief explanation of two of these measures). Criterion-referenced standardised tests use 'proficiency levels' or 'performance bands' to score students e.g. 'meets expected standards'. The Phonics Check performed at the end of Key Stage 1 in England uses a criterion-referenced score to identify what children have or have not met the expected standard for phonics decoding.

Think about... Standardised Tests
Performance on a standardised test can be summarised in different ways. We will just explain two in this box.

A student's **percentile rank** summarises the percent of students in the norming group that the learner outperformed. It does *not* represent the amount they got correct in the test. For example, if Jack received a percentile score of 80, then that means he performed better than 80 percent of the students in the norming group.

A student's **standard score** summarises the distance of a student's performance from the mean. Lots of tests use standard scores that range from 55 to 145 and have a mean of 100 (with a standard deviation of 100). Scores between 85 and 115 are within the average range. If Lottie received a standard score of 109, then we can say that Lottie's performance is within the average range when compared to peers of the same age.

Think...

What if the norming group of a test is very different from the learners who take that test? Can a valid interpretation be made about the learners from their scores? Why or why not?

Isaacs et al. (2013) defines examinations as 'an attempt to measure a learner's knowledge, understanding or skills within a certain subject or sector domain in a limited amount of time'. They can be used to certify students (e.g. getting a 'license' to practice medicine), select candidates for third-level or further education (e.g. A-levels being used by universities to select students for courses) or for accountability reasons. While exams allow learners the opportunity to demonstrate their knowledge under the same conditions as their peers, the fairness of this assessment approach can be influenced by several inter-related factors that should be borne in mind when considering the appropriateness of the interpretations that can arise from them.

Summative Assessment and Reporting: Current Policy

The most visible form of summative assessment used in England are the Standard Assessment Tests (SAT). These tests are based on the national curriculum and are only administered in England. They are currently completed at the end of Key Stage 2 (KS2) (end of

primary phase) and take place annually. Schools are statutorily obliged to administer them in state-funded schools. There are three specific SATs: (i) *Reading*, (ii) *Grammar, Punctuation and Spelling*, and (iii) *Mathematics*. In addition, a selected sample of pupils countrywide also participate in biennial assessments in science. The purpose of these forms of summative assessments is to monitor student progress and attainment. They are also used by the Department for Education to hold schools accountable for their performance.

The SATs are considered ‘high stakes’ assessments as they are reported to the local authorities and can be published at local authority, regional and national levels (as they were in 2022). This allows for the development of league tables where a school’s results can be measured against the results of all schools nationally as well as other schools in its similar catchment area. Schools are ranked in these league tables in order of their success in the assessments. These results are reported on school websites, within school prospectuses, and can also be reported in the media generally. The use of statutory assessments at the end of KS2 to create league tables are subject to critique, given they only provide a narrow picture of what is occurring in the school (Gardner et al., 2010). However, it can also be argued that these league tables equip parents with important information on how to select a school that caters for their children (Roberts, 2022).

Alongside the national curriculum assessment at the end of KS2, children also undertake a range of other summative tests and assessments during their primary schooling in England. In the first year of primary education, known as Reception, children undertake the Reception Baseline Assessment (RBA) within their first six weeks. This baseline assessment looks at (i) *maths* and (ii) *language, communication, and literacy*. The scores are not publicly available but are instead used within schools to monitor the progress pupils make from Reception until the end of KS2. Schools are also required to administer the criterion-referenced *Phonics Screening Check* in Year 1. The purpose of this assessment is to check if all children have learned decoding skills to an age-appropriate level. They complete the screen again in Year 2 if they do not reach the expected standard. While the results of individual schools are not published in the performance tables, the results will be made available to Ofsted (inspectorate) and to local authorities so that schools will be able to allow schools to benchmark the performance of their children.

In Year 2, schools can administer the Key Stage 1 (KS1) SAT in *English reading, Grammar, Punctuation and Spelling* and *Mathematics*. Since 2023, these assessments are optional for schools. They do not need to be published but can be used by teachers as a way of reporting progress to parents. In Year 4, English schools are also statutorily obliged to administer the *Multiplication Tables Check (MTC)*. The purpose of this assessment is to determine whether pupils can fluently recall their times tables. Like the *Year 1 Phonics Screening Check*, the results of individual schools are not published in performance tables but are made available to national and local authorities. The table below summarises some of the tests that are undertaken in primary schools in England as of 2023.

Table X Assessments in English primary schools (see Roberts, 2022)

Name	When	Content	Reporting
Reception Baseline Assessment (RBA)*	Within first 6 weeks of beginning Reception	Maths Language, Communication and Literacy	Results used by teachers within schools Not publicly available
Phonics Screening Check*	Year 1 (and Year 2 if necessary)	Word reading skills (decoding)	Reported at local authority level Reported to Ofsted
KS1 Standard Assessment Tests (SATs)	Year 2	English reading English grammar, punctuation and spelling Mathematics Science (for a sample of schools)	Results used by teachers within schools Can be reported at local authority level
Multiplication Tables Check	Year 4	Times Tables up to 12	Reported at local authority level

KS2 Standard Assessment Tests (SATs)*	End of Key Stage 2 (Year 6)	Reading Grammar, Punctuation and Spelling Mathematics Science (for a sample of schools)	Can be published at local authority, regional and national levels.
---------------------------------------	-----------------------------	--	--

* All English primary schools are obliged to administer these assessments

Think about... Publishing test results

England is unique in that it is the only jurisdiction in the UK that produces league tables based on the performance of individual schools at primary level. For example, while it is a statutory requirement for schools in Wales to annually administer online assessments in reading and numeracy for learners in Years 2 to 9, test results are not publicly available. Instead, the central purpose of these assessments is to help teachers plan next steps and to support learner progress.

Think... Do you think League Tables for primary schools are a good idea? What are some advantages and disadvantages for teachers, children and parents?

Summative and Formative Assessment: Key Comment

While summative and formative assessment have been considered separately in this book, harshly categorising an assessment tool as ‘formative’ or ‘summative’ may be somewhat counterproductive. For example, an end-of-week assessment designed by a teacher to measure

addition skills in mathematics may have been originally conceived as a summative measure of learning but it can *also* be used to identify areas for future learning and instruction e.g. renaming within 100. Therefore, it may be better to acknowledge that ‘supportive links can be forged’ between summative and formative assessment as both support the production of judgments about student learning (Black & Wiliam, 2018, p. 569). While some assessment tools are better suited for summative purposes (e.g. tests) and others are better for formative purposes (e.g. self-assessment), it is important to remember that it is the judgements or inferences arising and *not* the assessments themselves that are formative or summative. Therefore, any assessment tool that is controlled by the teacher can potentially be used for formative or summative purposes, or a combination of both.

Summary

In this unit, we have sought to define and describe summative assessment and its role in classrooms. We spoke about three key characteristics of assessment – validity, reliability and fairness. We discussed some common approaches to summative assessment as well – classroom tests, performance-based assessments, rubrics and standardised tests and exams. To recap the major points of the unit, we suggest that you revisit the learning objectives we noted on the first page. As you do this, you might consider the different ways in which you could demonstrate your understanding and knowledge of the topic.

Further Reading

Murchan, D., & Shiel, G. (2017). *Understanding and Applying Assessment in Education*. SAGE Publications.

This book provides an excellent overview of assessment across all levels of the education system. It is a ‘one-stop-shop’ for all issues in summative and formative assessment using examples of practice from the UK, Ireland and other international contexts. It goes into more detail on many of the topics outlined here in an easy-to-read, highly accessible manner.

Elwood, J. and Murphy, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182–192.

This editorial to a special issue on sociocultural perspectives on assessment, along with its accompanying articles, offers a challenging and theorised account of assessment research, policy and practice that is worth studying to sharpen your thinking and invite you to challenge some current practices.

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.

<https://doi.org/10.1080/0969594X.2018.1441807>

This piece from Black and Wiliam outlines how assessment, pedagogy, instruction and learning all relate to each other. Of particular interest is how the authors speak about the relationship between formative and summative assessment.

Reference List

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6), 551–575.
<https://doi.org/10.1080/0969594X.2018.1441807>

Brookhart, S. M. (2015). *Performance assessment: Showing what students know and can do*. Learning Sciences International.

Burke, P. & Lehane, P. (2023). *Weaving the literature on integration, pedagogy and assessment: Insights for curriculum and classroom*. National Council for Curriculum and Assessment.
<https://ncca.ie/media/6371/weaving-the-literature-on-integration-pedagogy-and-assessment.pdf>

Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment*. McGraw-Hill Education.

Isaacs, T., Zara, C., Herbert, G., with Coombes, S.J., & Smith, C. (2013). *Key concepts in educational assessment*. SAGE Publications.

Murchan, D., & Shiel, G. (2017). *Understanding and applying assessment in education*. SAGE Publications.

Niemann, D., Martens, K. & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education*, 52, 175–183. <https://doi.org/10.1111/ejed.12220>

Pizmony-Levy, O., & Bjorklund, P., Jr. (2018). International assessments of student achievement and public confidence in education: Evidence from a cross-national study. *Oxford Review of Education*, 44(2), 239–257. <https://doi.org/10.1080/03054985.2017.1389714>

Popham, J. (2011). *Classroom assessment: What teachers need to know* (6th Edition). Pearson.

Roberts, N. (2022). *Assessment and testing in primary education (England)*. House of Commons Library. <https://researchbriefings.files.parliament.uk/documents/CBP-7980/CBP-7980.pdf>.