

A young woman with dark hair pulled back, wearing a white graduation gown with a blue stole and a blue mortarboard cap. She is looking off to the right with a thoughtful expression. The background is blurred, showing other graduates in a similar setting. The image is overlaid with several semi-transparent blue geometric shapes.

Understanding Balanced Assessment Systems

Integrating assessment in a way that works for students and their families, the school, the district and the state.

Introduction

Assessment is an established part of the educational landscape. It has a critical role to play in improving educational outcomes by measuring student learning. But this landscape is evolving as new legislative frameworks, such as the Every Student Succeeds Act (ESSA), introduce new accountability requirements, and wider developments such as rapidly advancing technologies emerge. These changes in turn have an impact on assessment systems. Educational stakeholders at all levels should consider whether their assessment systems continue to provide the right information in a timely manner and in the appropriate format to ensure their system is still fully fit for purpose.

Understanding the different types of available assessments and how these can be connected to enable best assessment practice is a key step

to assuring fit for purpose. Within K 12, there are three main assessment types: formative, interim and summative. Each has a critical role to play in delivering the right data to the right people to meet their particular needs, from student, parent, educator and principal to district- and state level stakeholders. Together, these assessments combine to create a balanced system that provides insights to accelerate educational progress.

This briefing looks at the characteristics of each form of assessment and how each could be used and applied to yield evidence that can inform various decisions, whether at a policy, district/school or classroom level to support the improvement of education. It also looks at the benefits of integrating assessment to create a balanced system whose whole is greater than the sum of its parts.



Fit for purpose

The reason three different types of assessment are utilized in the K-12 arena is because each serves a different purpose. Understanding the goal of each assessment can help to ensure each is used to appropriately add value to improving overall learning outcomes. So, whether it's formative, interim or summative, it is important to be familiar with the function, and the limitations, of each form of assessment.

Formative assessment

Fundamentally, the purpose of formative assessment is to inform both students and teachers about learning in the classroom. Formative assessment occurs within the classroom, planned and orchestrated by the teacher and provides information that helps them to make decisions about what are appropriate next learning steps for students to move learning forward, and to support students as they gain insights into their own learning. Formative assessment can take many different forms, from purposeful listening to student discussions as they collaborate together and providing feedback to help them deepen their understanding, to bringing important ideas forward to the whole class, or to extending work on a project with rounds of feedback from peers. Any information gained from formative assessment activities should be useful in the moment.

A good analogy for thinking about the role of formative assessment is Roger Bannister breaking the four-minute mile barrier. Finally running a sub four-minute mile was a summative performance with a specific target reached. The times for all Bannister's practice runs were not used to calculate his average for the year, but all the practices were essential in order for him to achieve his 'summative' performance. In the same way, formative assessment informs and guides ongoing learning during the year until a culminating summative assessment.

Interim assessment

Interim assessment provides an opportunity to "check-in" on student learning at several points during the year and to get an estimate of likely performance

on the summative assessment. It is intended to provide a shared point of reference across teachers and classes within a grade level on student learning during the year. Interim assessment data can be used to examine group performance to address questions such as, "how does the performance of English Learners in our school compare to other students?" Data could be disaggregated by gender, race/ethnicity, disability status, or socio-economic status if there are sufficient numbers of students in a subgroup.

Data may be used to inform some adjustments in resources or curriculum strategies during the academic year, but may not necessarily influence classroom instruction in the way that a more targeted formative assessment would. Essentially, the information interim assessment delivers enables administrators and educators to understand where students are with respect to grade-level standards at given points in time during the year

Summative assessment

The goal of state-wide summative assessment is to evaluate student learning usually near the end of the school year. It may also be referred to as the accountability assessment. State-wide summative assessment provides a broad view of student and school educational performance and allows districts and states to measure how well learning and teaching is meeting required state standards. As it measures specific outcomes at a single point in time, it is useful for examining individual student's overall mastery of state standards and also for comparing performances of groups of students across schools or districts. Given the survey nature of the assessment - covering a year's worth of standards in a relatively short period of time - it produces aggregated data that is useful for state education agencies and districts for accountability and resourcing purposes. For teachers, it can identify student strengths and weaknesses broadly but they will need additional more targeted information during the academic year to inform ongoing instruction.

Figure 1 below illustrates how formative, interim and state-wide summative assessments exist together, illustrating how the stakes vary, and the scope of each assessment varies in terms of the standards assessed.

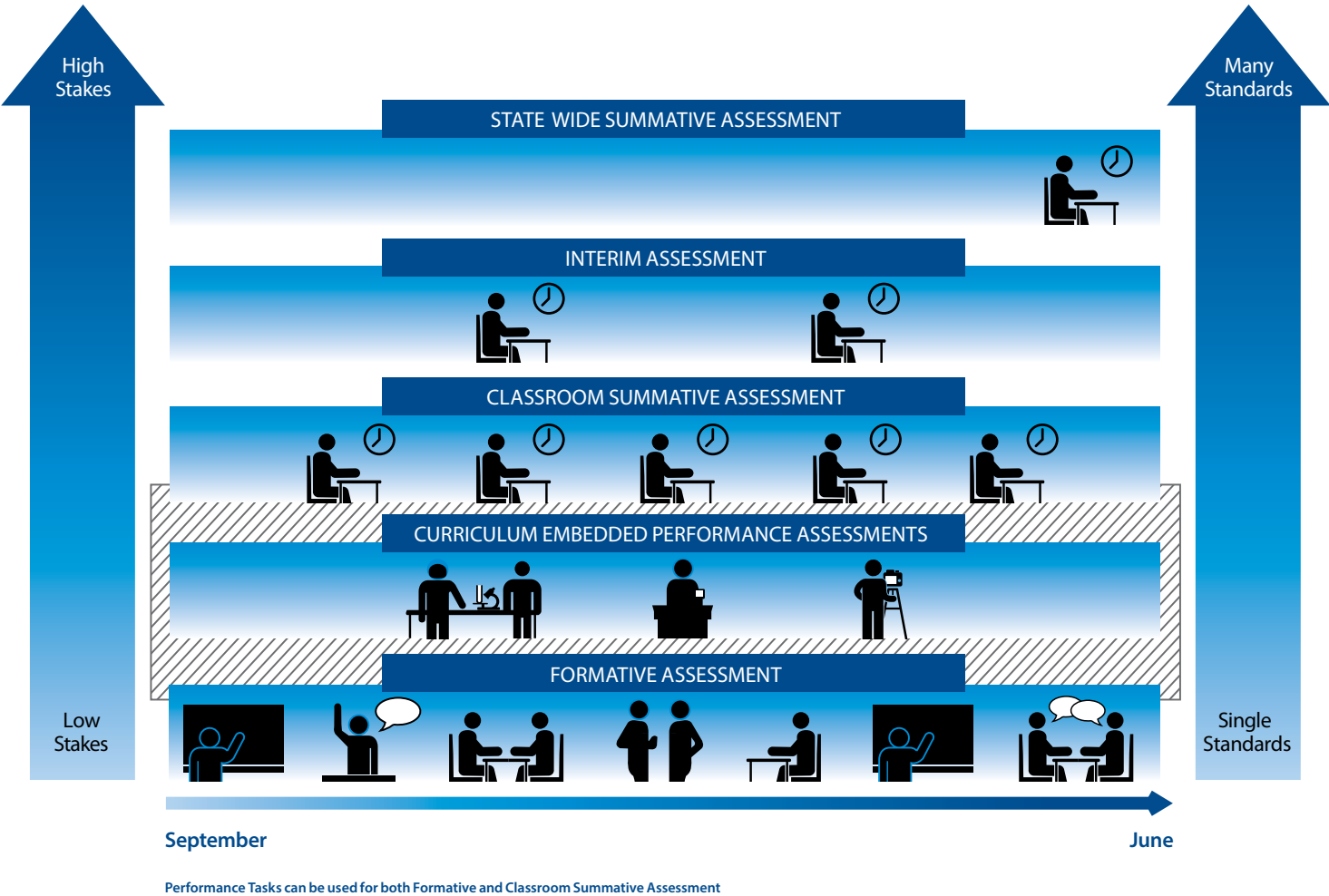


Fig. 1. Representation of a balanced assessment system

Note that while not the focus on this paper we recognize that teachers often use classroom summative assessments as part of the process of determining student grades and also there is an increasing interest in the role of performance assessments which can be used in either a formative or summative capacity (Wylie & Lyon, 2017).

Getting the most out of assessment data

Each type of assessment produces a different type and grain size of evidence, from the very individualized information of formative assessment to the broader year-long view of summative. Formative, interim and summative assessment work together to create a multi-faceted view of learning at an individual, class, district and state level and the differing size of data generated by each is key to their appropriate deployment and successful application.

The relevance of real-time data

Formative assessment provides very fine grain information, sometimes targeting only a single standard or aspect of a standard, which may be tailored to an individual student or a small group and of a particular moment. It is timely and informative, providing real-time feedback that teachers are then able to quickly apply to adjust their teaching plan to better reflect specific needs, or that students can apply to their own work to improve it. It is the immediacy and relevance of the insights gained that makes it highly effective. This allows teachers to incorporate the evidence of student learning into their planning and act on insights to augment their classroom-based instruction immediately, making a positive impact on deepening student learning. This process of timely adjustment to meet student learning needs as they are emerging has a positive impact on student learning^{1,2,3}. There is no delay between the capture and application of data and this real-time characteristic is crucial to effective formative assessment.

Formative assessment can be wide-ranging, from more practice-based activities such as quick, verbal

checks-for-understanding, to more formal types of assessment such as extensive tasks that support deeper learning and that are designed to provide more scaffolded or supported learning opportunities. Evidence may take the form of notes that the teacher makes about questions to ask students about their writing drafts during conference time the next day, patterns across a set of exit tickets that students complete at the end of a class that will then inform groupings for an opening activity the next day, or student self-reflections or feedback to peers. In some cases, a teacher may share evidence with another teacher to see if she has observed similar patterns in student work in order to strategize an effective alternative representation to help students better understand an important concept. However, it is less likely that evidence will be reported or shared beyond a very immediate, local context. Most critically, if the evidence is truly formative then it will provide information to be acted on immediately, either confirming for the teacher that the direction she is going in is appropriate or suggesting a different next step, but in either instance the information will become quickly out-of-date. In short, the teacher and students obtain information about learning, both student and teacher can respond to that and then the learning has moved forward.

Research suggests that teachers need ongoing professional support to develop and deepen their formative assessment practices^{4,5}. Collaborating with peers to plan and create shared tools and approaches to elicit evidence of student learning, to analyze student work together and to plan ways to deepen student learning based on evidence of current learning, are all important professional learning experiences for teachers. Teachers need time and opportunities to develop and practice these skills.

¹ Black, P., & William, D. (1998). *Assessment and classroom learning. Assessment in Education: Principles Policy and Practice*, 5, 7-73.

² Heritage, M., & Heritage, J. (2013). *Teacher questioning: the epicenter of instruction and assessment. Applied Measurement in Education*, 26, 176-190.

³ Andersson, C., & Palm, T. (2017). *The impact of formative assessment on student achievement: a study of the effects of changes to classroom practice after a comprehensive professional development programme. Learning and Instruction*, 49, 92-102.

⁴ Gotwals, A.W. & Birmingham, D. (2016). *Eliciting, identifying, interpreting, and responding to students' ideas: Teacher candidates' growth in formative assessment practices. Research in Science Education*, 46: 365. <https://doi.org/10.1007/s11165-015-9461-2>

⁵ Furtak, E.M., Kiemer, K., Circi, R.K. et al. (2016). *Teachers' formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. Instructional Science*, 44: 267. <https://doi.org/10.1007/s11251-016-9371-3>

The benefits of interim assessment

School and district leaders need to have a view of student performance as the school year unfolds so they can make informed, local decisions such as where to deploy coaches, or what type of professional development needs to plan for. Similarly, teachers want to be able to gauge student performance against summative expectations at the end of the year and adjust curriculum and learning strategies for both the student and cohort at regular points accordingly.

Interim assessments provide additional opportunities to monitor student progress using a set of content that is broader than formative, before reaching the summative end-of-year testing. Performance on the interim assessments will confirm a teacher's formative assessment judgments about student learning, or help teachers to identify students who may be performing more strongly or more poorly than the teacher had realized, and focus instruction in these areas. Students can sometimes have changes in learning that go undetected by the teacher and these interim check points can draw attention to these students. While interim data has greater longevity than its formative partner, it must still be provided in a timely way if it is to be acted on to drive educational improvement and to be effective.

Beyond accountability

State, district and school leaders may want to understand student performance in aggregate, both in terms of absolute attainment and progress over time, but also by sub-groups to identify disparities and monitor the effectiveness of approaches being used to reduce achievement gaps. This is the role of summative assessment – to provide data that can support meaningful comparisons across groups of students, classes, schools, districts and so on. This macro data provides districts and schools with an overall pulse on how

students are progressing by grade, by school and by content area. It also measures student achievement against required state standards to deliver the type of information that may then be used to develop educational policies at a state and federal level.

However, summative assessment has applications beyond accountability. The data can help district or school leaders to identify areas for professional learning, and it can support teacher reflection on teaching strategies or curriculum at the end of the year and inform adjustments ahead of the next year's instructional planning. In addition, summative data might be used for planning at the start of the year, with the receiving teacher using it to get a snapshot of the new students entering their classroom and to think about the appropriate level to begin instruction.

With the introduction of the Every Student Succeeds Act (ESSA), state educational stakeholders now have the opportunity to revisit their assessment systems and to explore options that do not focus solely on a single end-of-year assessment. During this time of transition, some states are beginning to explore options around using multiple interim assessments for the purpose of accountability or greater use of performance assessments.



A balanced assessment

A strategically balanced assessment system is one that incorporates summative, interim and formative components in order to provide meaningful and interpretable information for stakeholders at all levels in the educational system. By working together, these individual components provide greater insights into where students are and where they need to be throughout their K-12 journey, supporting learning opportunities for all students that are addressing their individual learning needs to help improve educational success for all.

The need for integration

Formative, interim and summative – each type of assessment has a role to play in enhancing learning outcomes and driving forward standards in education. Understanding when to attend to each source of information is important. Usable, meaningful data is also timely data. Timely summative results allow teachers to use them to reflect on the year just ending and to use them as part of their planning for the following year. Interim results can be reviewed by grade level teams after each administration to identify any adjustments needed to instructional plans for the rest of the year. Formative assessment evidence – based on the same set of standards – will be ongoing throughout the year supporting teachers and students to identify current understandings from which to build next instructional or learning steps.

A coherent system built around common standards helps to create consistency across the different assessment components and means that the information generated can be interpreted more easily and productively. For example, if interim assessments are built to the same blueprint as the summative assessment, only shorter, as within a balanced system, then results can be linked directly to progress towards summative requirements. An effective assessment system is one that reconnects assessment to learning.

A balanced assessment system does this by allowing the various testing components within the system to interact with each other. Interim and summative assessments can use the same reporting scales and share item types, for example, making it easier to integrate and compare analyses. Formative assessment directly provides support for teachers to closely attend to student understanding to develop instruction that best meets their immediate learning needs, and supports students reflect on their own learning and that of their peers which also has a positive impact on their learning.

The role of assessment design

With advances in technology, summative assessments are able to more strongly signal what is important for deeper student learning by greater use of assessment items that model good instructional practices, requiring students to demonstrate understanding through writing. In the past state-wide assessments tended to use primarily multiple choice questions due to the cost of human scoring and lack of technology support for more enhanced item types. With artificial intelligence (natural language processing) technologies, student open-ended responses can now be scored in a reliable way without the cost of human scoring. This allows the summative assessment to more fully assess the breadth of standards, which is likely to have a positive effect by encouraging broader curriculum teaching rather than focusing instruction only on parts of the standard that were known to be assessed.

For assessments that are closer to instruction, value is added by providing information to help teachers plan next instructional steps for a student or cohort. Where assessments are able to make use of learning progressions that target key ideas in the standards and describe how student understanding develops from naïve to expert levels, the reports can support teacher planning by signalling what is likely to be the next developmental milestone for students.

Making better sense of data

A system that uses a common language where appropriate across all components, and where reporting focuses on meaningful, actionable next steps appropriate for that component enables stakeholders to more easily understand, communicate about, and take action in the light of students' learning. Connecting the various types of assessment can have a positive impact on analysis and reporting too.

A single dashboard that contains all the assessment data can be accessed and shared by all educational stakeholders at the appropriate grain size. For example, if an online reporting system contained both state-wide summative results and interim assessment results it could facilitate the use by state, district and school administrators to examine levels of student progress and attainment of state milestones, making

the sharing of data more straightforward, increasing opportunities for gaining insights about student progress using multiple source of data, and creating a more efficient approach.

How data is presented can also improve the effectiveness of assessment and reduce time spent analyzing data to pinpoint the key trends. An intuitive system that allows stakeholders to easily identify relevant information without extensive training will increase the likelihood that the reports are accessed and analyzed, and the information used.

Furthermore, advances in technology and the increasing availability of curated online teaching resources support the development of score reports that can link to additional materials that might be useful for next teaching or learning steps.



In Summary

K 12 assessment can seem complicated. Different types of assessment using different standards, reporting and delivery systems can produce a feeling that there is too much assessment producing too much data and not enough useful information.

Designed, developed and implemented effectively, assessment can play a valuable role in supporting learning outcomes and improving education. However, understanding the characteristics of formative, interim

and summative assessment is key to also understanding how together these assessment types can add value beyond the sum of their parts. A balanced approach to assessment connects all three components to create a more efficient pathway to improving educational outcomes for all students.



Copyright © 2018 by Educational Testing Service. All rights reserved. ETS, the ETS logo are registered trademarks of Educational Testing Services (ETS). MEASURING THE POWER OF LEARNING is a trademark of ETS.
All other trademarks are the property of their respective owners.



Measuring the Power of Learning.®

www.ets.org