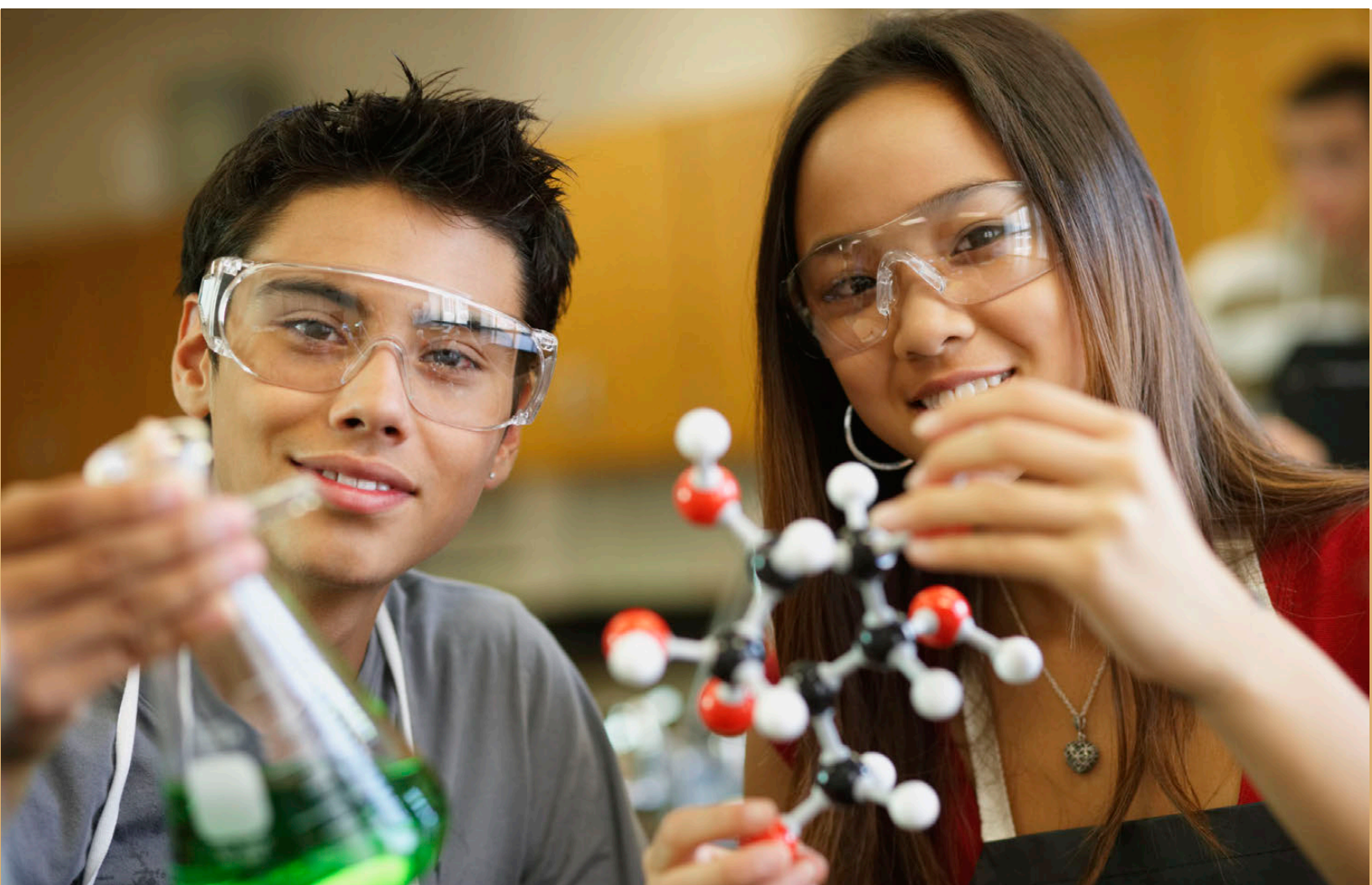


Assessment Design for Broader, Deeper Competencies

Report 12 of the MyWays Student Success Series



October 2017

Grace Belfiore and Dave Lash
for Next Generation Learning Challenges

The MyWays™ Student Success Series

All reports in the series are available for download at myways.nextgenlearning.org/report.

Visual Summary
Introduction and Overview

Part A: Adolescence in an Age of Accelerations

Summarizes specific real-world realities and conditions confronting today's young people.

- Report 1: **Opportunity, Work, and the Wayfinding Decade**
- Report 2: **5 Roadblocks to Bootstrapping a Career**
- Report 3: **5 Decisions in Navigating the Work/Learn Landscape**
- Report 4: **5 Essentials in Building Social Capital**
- Report 5: **Preparing Apprentice-Adults for Life after High School**


Part B: Broader, Deeper Competencies for Student Success

Provides a composite definition of student success in learning, work, and life.

- Report 6: **Welcome to the MyWays Student Success Framework**
- Report 7: **Habits of Success — for Learning, Work, and Well-being**
- Report 8: **Creative Know How — for a Novel, Complex World**
- Report 9: **Content Knowledge — for the Life Students Will Lead**
- Report 10: **Wayfinding Abilities — for Destinations Unknown**

Part C: Redesigning the Learning Experience for the MyWays Competencies

Brings the broader and deeper competencies of the MyWays Student Success Framework into educational practice.

- Report 11: **Learning Design for Broader, Deeper Competencies**
-  Report 12: **Assessment Design for Broader, Deeper Competencies**

About this report

Report 12, *Assessment Design for Broader, Deeper Competencies*, presents the evolution toward greater authenticity and multiple measures, and recommends the use of Five Assessment Strategies: *Formative assessment, Performance assessment, Multiple measures, Badges & micro-credentials*, and *Quality reviews*. The report also surveys the emerging state of assessment within each of the MyWays domains, as well as offering resources to support initial moves toward implementation.

Report 12 is the second of two reports in Part C of the *MyWays Student Success Series*. **Part C, “Redesigning the Learning Experience for the MyWays Competencies,”** explores how to bring the broader and deeper competencies of the MyWays Student Success Framework into educational practice, focusing on key constructs for learning design and assessment design.

The *MyWays Student Success Series* examines the through-line of four essential questions for next generation learning and provides research and practice-based support to help school designers and educators to answer these questions. The series consists of 12 reports organized into three parts, plus a Visual Summary and Introduction and Overview.

The **primary researchers and authors** of the *MyWays Student Success Series* are Dave Lash, Principal at Dave Lash & Company, and Grace Belfiore, D.Phil., Principal Consultant at Belfiore Education Consulting.

MyWays is a project of Next Generation Learning Challenges, an initiative of the non-profit EDUCAUSE. MyWays is supported through a grant from the William and Flora Hewlett Foundation with additional support from the Bill & Melinda Gates Foundation, the Barr Foundation, and the Oak Foundation.



nextgenlearning.org



© 2017 EDUCAUSE. This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Cover photo courtesy of [Project Lead The Way](https://www.projectleadtheway.org/).

REPORT 12

Assessment Design for Broader, Deeper Competencies

“ [T]he world has already shifted. A broader, richer range of capabilities is being demanded of our young people, and there are innovations in learning, teaching and organization that begin to support them. But **a focused, shared effort is needed to ensure that, collectively, we measure what we value across our broader expectations and make sure that our education systems reflect those measures to the greatest extent that learning science and evidence make possible.** [emphasis added]

—Global Education Leaders’ Partnership¹

Introduction

This ability to “measure what we value” in a world that has shifted is a mission-critical challenge to next generation educators. There is no finished blueprint for next generation assessment, no glossy catalog of proven assessment approaches and tools, and no formula for replacing low-cost, narrow accountability testing with thoughtful investments in gauging the progress of the whole learner. On the other hand, there are pockets of growing research and maturing practice in measuring learning and competency in some skill areas and in increasingly authentic settings. The challenge in designing assessment for the whole learner is creating new combinations comprised of established assessment tools, processes for feedback, and methods for reflection from within and outside the education sector, together with new tools and processes for collecting evidence where gaps exist. As Andrew Hargadon notes in *How Breakthroughs Happen*, this recombinant process is the tried and true nature of most innovation.²

Given the nascent nature of the field, measurement, or *assessment*, is the subject of the fourth big question at the heart of the MyWays Project. How do we, as next generation educators and students alike, gauge student progress in developing the broader and deeper competencies in the MyWays Student Success Framework? And how do we measure our performance in fostering this development as teachers, schools, and communities?

The MyWays Through-line

1	WHY the urgency to change? What are the real-world conditions that our students will need to address?
2	WHAT does success look like for students in a world of accelerating change? What competencies promote a broader, deeper definition of success?
3	HOW can our learning experience design help students develop the broader, deeper competencies?
4	HOW do we gauge student progress in developing these competencies? How can we measure our school’s success beyond proficiency in math and ELA to embrace whole child development?

When dealing with the *How* of next generation learning, Next Generation Learning Challenges and the MyWays team feel that it is important for educators and school designers to step back and re-visit the through-line; designing their *How*s not only from deep knowledge of their students, but also with intentional reference to *Why* the urgency to change and *What* kind of competencies students need in order to develop to those compelling whys. This helps to ensure that their work is not built on a collection of

the “latest trends” or even isolated “good ideas” — though there are certainly some excellent individual learning and assessment ideas that can be integrated into this intentional design. But one critical element of the MyWays Student Success Framework is that success in learning, work, and life in our changing world is what drives each stage of the process. We therefore urge you strongly to read the reports in Parts A and B of this series. (Or follow the tips for a quick drive through the reports, as noted in the box to the right.) In the meantime, here’s a quick summary of what’s behind our response to the last through-line question:

Tips for a quick drive through the series, for readers of Report 12

- Read the *Visual Summary* and the *Introduction and Overview*, and skim Reports 1, 6, and 11.
- Explore the Key Takeaways at the end of Reports 2–5 and the Competency Definitions and Key Principles in Reports 7–10.
- Visit the [MyWays website](#) for top-line concepts.

Using the MyWays Through-line to inform how we gauge student progress



In response to Question 1 (**WHY** the urgency to change?), Part A analyzes the nature of the **rapidly changing, more variable, more complex world**.

The nature and speed of this change shapes the response to Question 2 (**WHAT** success looks like), which is Part B’s exploration of **the four domains of broader, deeper competencies needed to address this change: Habits of Success, Creative Know How, Content Knowledge, and Wayfinding Abilities**.

In turn, the range and nature of these broader, deeper competencies set the requirements for the response to Question 3 (**HOW** can our design for learning support these competencies?), which is a **learning design that is more authentic, holistic, and focused on student agency**.

In response to Question 4 (**HOW** do we gauge student progress in developing these competencies?), this report suggests that the nature of the learning design requires **a system of assessments that are also authentic, student-driven, and holistic, and that these need to be integrated with and extend that learning**.

Given the integration of these last two steps of the through-line, we emphasize that, although we have presented learning design and assessment design in separate reports, they need to be planned and undertaken together. Accordingly, this report builds on the three learning design principles described in Report 11: *Whole Learning* (seven principles of holistic, authentic learning, and their use in junior versions of adult activities); a *Wider Learning Ecosystem* (spaces beyond the school walls that enable a coordinated portfolio of diverse, real-world learning experiences); and *Lever for Capability and Agency* (eight targeted practices based on learning and developmental science).

This report overviews assessment for broader, deeper competencies by presenting the following:

- A brief snapshot of the evolution taking place in assessment
- Two key paradigm shifts needed to assess broader and deeper competencies
- Five recommended assessment strategies we believe are essential to success, including for each strategy a one-page primer and a practice example from the Assessment for Learning Project
- A look at the trajectory of the field of assessment for broader, deeper competencies, including cautions, potential, and the ways those in the field are working together

- A scan of the state of assessment across the four MyWays domains
- A quick resource dive for next gen assessment, highlighting
 - Starter resources
 - Three simple MyWays evaluation tools and a case study for how to use them
 - A guide to sources for deeper assessment implementation tools

Building capacity within your faculty and staff to apply these strategies in integrated, mutually reinforcing ways will lead, over time, to systems of assessment that are more effective across the broader competency range. This report is intended to support you in evaluating your current assessment practices and identifying areas for development.

Assessment as a means of knowing the whole learner



What would it look like to, as the Global Education Leader’s Partnership puts it, “measure what we value,” in the most meaningful way possible for broader, deeper competencies? A growing number of educators have been thinking deeply about this end-goal, and indeed rethinking assessment for whole-person outcomes.³ For example, Michael Fullan highlights that next generation assessment would

- support rapid feedback cycles on learning progress;
- include a wider variety of participants in the assessment process, such as peers and outside experts;
- provide more complex assessment experiences, including collaborative problem solving, embedded performance assessment, game-based assessments, and online capture of process skills; and
- assess more complex learning products, such as student work artifacts.⁴



Other educators highlight a range of related priorities, such as integrating assessment with learning and teaching; using a wider variety of measures; emphasizing process as well as product; and paying attention to the context and transfer of learning. Youth development professionals emphasize the inclusion of more descriptive and longitudinal analysis and suggest that, “in recognition that the purpose of schooling is to prepare youth for tasks, problems, and opportunities outside of it, some assessment would focus on emergent expressions of lifelong capacities — cognitive, social, practical, civic, identity related.”⁵

We particularly like Andrew Miller’s vision of assessment as a “force for knowing our students”:

Truly, assessment can be a powerful force for knowing our students... We simply have to move past the baggage that comes with the term assessment, and understand that it can mean a lot of things. We can assess for content and skills, yes, but we can also assess for passions, interests, success skills, and the like for the purposes of the right instruction at the right time.⁶

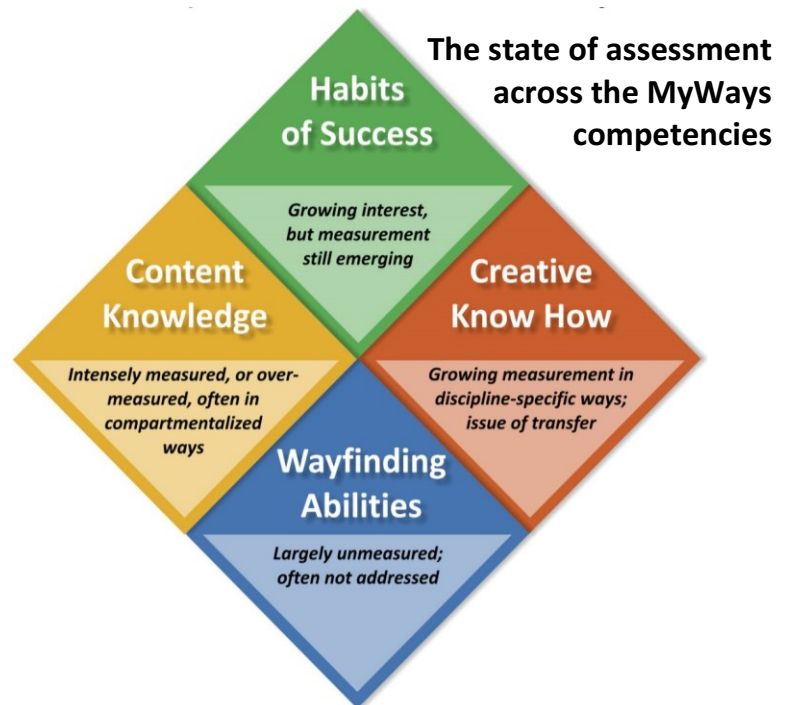
We will explore these and related aspirations in more detail in the remainder of this report. First, however, it is important to note that the state of assessment today across the broader competency range is fragmented, uneven, and fails to achieve many of the desirable attributes noted above. As we describe in more detail later, assessment varies dramatically across the MyWays domains (see the graphic below).

Some of the shortcomings in current assessment practices relate to the narrow range of what is measured, but much of the challenge arises from the nature of the assessments commonly carried out, even for traditional competencies, due to a preoccupation with accountability.

Note: *In looking at assessment design for broader competencies, we focus in this report on assessment **for learning** and for progress to competency, rather than for accountability. While some of the forms of assessment for Content Knowledge and Creative Know How show potential for both accountability and learning (as in New Hampshire’s U.S. Department of Education waiver to include performance assessment in their accountability*

system), the Next Generation Learning Challenges MyWays team is convinced that the assessment of Habits of Success and Wayfinding Abilities is still too emergent to use in accountability measures today (and, indeed, some measures of these competencies may never be appropriate for that purpose). State-driven accountability is too blunt an instrument to lead — right now, anyway — to the thoughtful, nuanced practice that competency development within these domains requires.

We would also like to emphasize that this report is intended to introduce and summarize just a selection of the strategies and issues related to assessment — those particularly relevant to the expanded success competencies. It does not provide a comprehensive view of the exploding assessment landscape, nor do the examples cited represent the full range of what is out there or currently being piloted. (For more on this fast-moving, wide-ranging, and sometimes controversial field, see the [MyWays](#) website and the [Assessment for Learning Project](#) website.) That said, we were struck by how, out of all the changes in assessment, two major paradigm shifts are emerging as vital, particularly in relation to the development of broader and deeper competencies.



Two key assessment shifts: greater authenticity and an integration of multiple measures

In the *Introduction and Overview of the MyWays Student Success Series* and throughout Part A, “Adolescence in an Age of Accelerations,” we emphasize that operating effectively in the real world is essential to success in learning, work, and life. Accordingly, we define *competence* in any pursuit as the union of *capability* and *agency*, where capability is “knowledge and the understanding to use it in real-world situations” and agency is “a deep and durable self, acting to shape one’s development and environment.” In Report 11, we explore eight Levers for Capability and Agency.

Measuring competencies comprising a student’s capability *and* agency will require new approaches to assessment. Not only must we begin to assess hard-to-measure competencies such as creativity, social skills, and wayfinding abilities, we must also gauge how well students “own” these competencies and apply them in real-world settings. Comprehensive assessment systems that can address this challenge are some ways in the future. The crucial first step for next generation educators, we believe, is to adopt two paradigm shifts in assessment practice:

Two key paradigm shifts for next generation assessment



ASSESSMENT SHIFT 1: The shift to greater authenticity.

The first shift involves moving from poor proxies — for example, measuring a few narrow competencies (such as math and English) in tightly bounded ways (such as through multiple choice questions) — to assessments designed to include more rounded student performance in the outside world. Just as junior

versions of Whole Learning experiences (See more about these in Report 11.) create learning environments that are developmentally appropriate real-world approximations, assessments designed for those Whole Learning environments should involve developmentally appropriate tasks that approximate students' competence in the real world. For example, as we discuss later, well-designed performance assessments are an established tool for assessing how well students deepen and apply knowledge and skills within project-based learning experiences.

Moving to greater authenticity begins by recognizing just how far standardized student assessments — our most prominent assessment method — diverge from the tasks students will encounter in the outside world. Learning and assessment are inauthentic when students' knowledge of a subject is measured by how they address well-defined problems, working alone and without access to outside information or tools (see the chart below). Furthermore, many of the MyWays competencies defy simplification and compartmentalization into learning and assessment frameworks such as the chart's "Standardized Student Assessment" column. Instead, most competencies in Creative Know How, Habits of Success, and Wayfinding Abilities are closely aligned with the situated learning characteristics in the "Tasks in the Outside World" column (see Report 11 for more on situated learning).

Comparison of standardized assessments vs. real-world performance

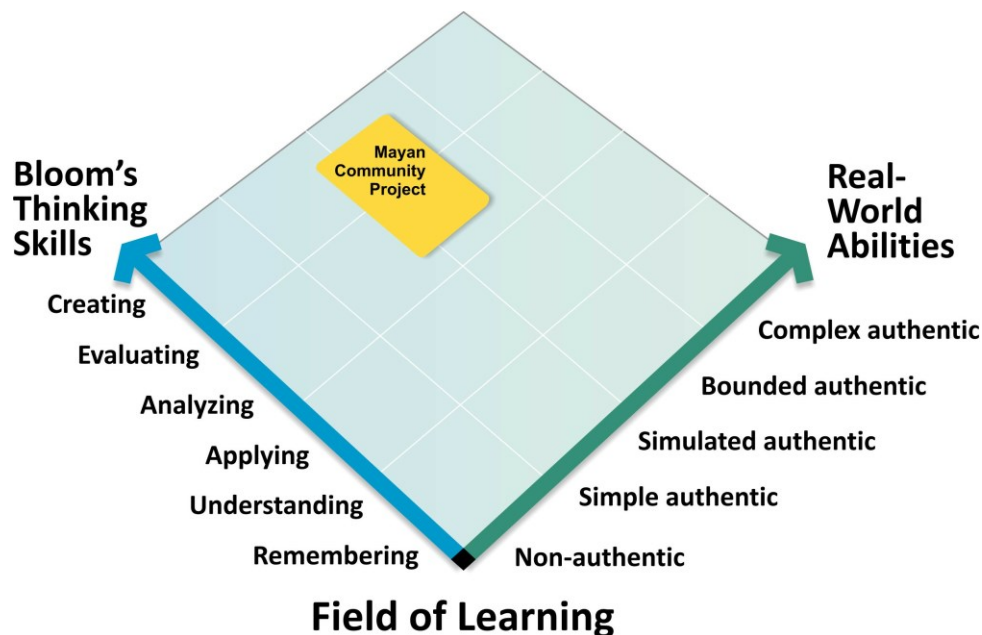
	Standardized Student Assessments	Tasks in the Outside World
Knowledge is:	Measured within a subject	Applied across disciplines , along with other skills, to solve real world problems, create products, and generate new knowledge.
Asked to address:	Facts and application of simple procedures to well-defined problems	Complex, disorderly problems in real-world contexts.
Work is done:	Individually	Individually and in groups of others with complementary skills to accomplish a shared goal.
Resources available:	Without access to outside information , and use only paper and pencil	The challenge is to evaluate information from a wide range of tools and resources to find what's relevant to analyze problems and create solutions

Adapted for the Stupski Foundation from *Transforming Education: Teaching and Assessing 21st Skills* (Cisco, Intel, Microsoft), 2010

In addition, learning science underscores dramatically how poorly learning in inauthentic settings, even for Content Knowledge and Creative Know How, transfers to real-world settings. For more on these findings, we point readers to the National Research Council's [Education for Life and Work](#) (2012) and [Make It Stick](#) (2014) by Peter Brown, Henry Roediger, and Mark McDaniel. If learning is more effective and durable in more authentic settings, then our assessment methods must become effective at providing feedback and guidance in those settings as well.

The MyWay Field of Learning graphic below offers a useful visual device for envisioning learning activities in terms of thinking skills and the real-world abilities they engender. The left axis of the

“ballfield” represents increasing levels of Bloom’s thinking skills. The right axis represents increasing levels of real-world abilities across a range of settings of increasing authenticity. This device was introduced in Report 11 to illustrate the difference in authenticity between traditional learning modes and learning design for broader, deeper competencies. For example, High Tech High’s Mayan Community Project (which we feature as an example of Whole Learning through project-based learning in Report 11 and an example of authentic assessment later in this report) is at the high end of the thinking skills axis while spanning both simulated and bounded authentic settings.



Bloom’s Thinking Skills

The left-field axis uses Bloom’s taxonomy* to key a familiar progression of thinking skills. While the taxonomy is not technically a hierarchy, in our usage here, the skills are cumulative as one moves out the axis: Applying, for example, includes Remembering and Understanding, while Creating includes all five of the “earlier” skills.

*Webb’s Depth of Knowledge would also work here, but we judged Bloom’s to be more broadly known and thus more suited to this purpose.

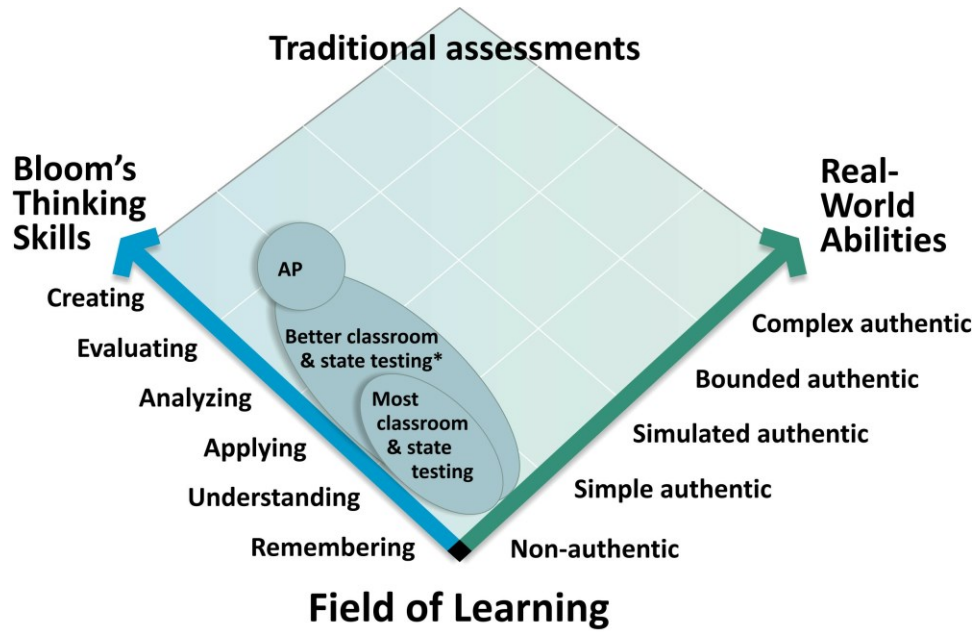
Real-World Abilities

The right-field axis indicates growing competence as the authenticity of the setting increases.

For example, a particular middle schooler might be competent in Self-Direction & Perseverance or Communication & Collaboration within “simulated authentic” settings (such as those within school), but not in “complex authentic” settings in the adult world.

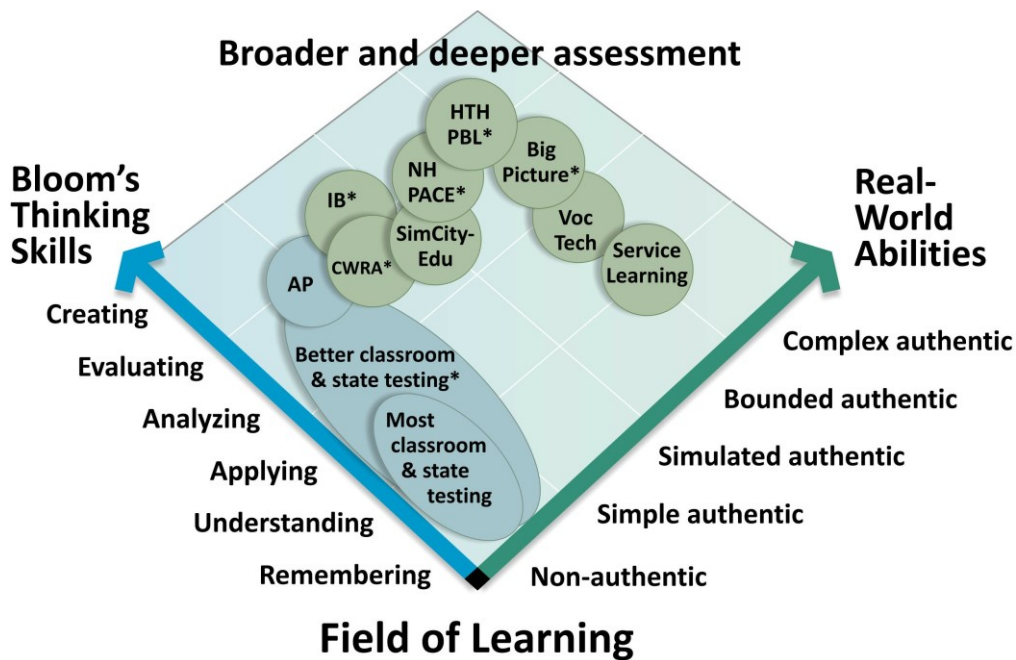
This progression also allows specific learning activities to be plotted by degree of authenticity.

Now that we have reviewed the basic diagram, let’s look at two versions illustrating traditional assessments versus examples of broader, deeper assessments that are higher on the right (authenticity) axis. The graphic below maps the common student experience in traditionally-designed public schools. Traditional assessments, including most classroom and state testing, focus on student performance on inauthentic measures, such as multiple choice questions. Better classroom and state tests, including PARCC, Smarter Balanced, and Advanced Placement (AP), include essays and other tasks that let students construct responses at a higher order of thinking, but they do not necessarily incorporate more complex, authentic contexts or settings.



*Asterisked area: Better classroom & state testing – includes PISA for Schools and Common Core assessments from PARCC and Smarter Balanced.

Conversely, the graphic below plots several examples of assessments that are aligned with Whole Learning principles and increasingly authentic learning contexts, including rich simulation, extended projects, and immersion in the real-world settings of the Wider Learning Ecosystem; these contexts increase the development of student agency, capability, and adaptability.



*Asterisked assessments are: Better classroom & state testing – includes PISA for Schools and Common Core assessments from PARCC and Smarter Balanced; IB – International Baccalaureate; CWRA – College Work & Readiness Assessment, a scenario-and document-based 90-minute assessment; NH PACE – New Hampshire Performance Assessment for Competency Education state testing (waiver); HTH PBL – High Tech High's Whole Learning-rich project-based learning approach; Big Picture – Big Picture Schools' experiential learning model, which integrates workplace-based learning. The latter two Deeper Learning models use extended, embedded performance-based assessment.

For example, situated about one-third of the way out the authenticity axis is the Council for Aid in Education’s College and Work Readiness Assessment (CWRA), an “on demand” 90-minute task that offers a real-world scenario (such as how a city might deal with pollution from a now-defunct factory) and gives learners access to a library of online documents from which to create their responses. Compared with many traditional, on-demand assessments, the CWRA succeeds in simulating some elements of real-world tasks, such as evaluating and analyzing information, and problem solving.

Significantly further out the authenticity axis, the projects pursued by students at High Tech High involve work on complex challenges drawn from the real world, but structured and bounded both for practical reasons and to maximize learning conditions (to target the application of certain skills, for example, or offer time for feedback and self-reflection). Service learning and extended internships (like those offered by Big Picture Learning) move a step further toward complex authenticity because students are actually working in the messiness of the real world to deliver real products for companies or communities; furthermore, clients use and respond to their work, providing authentic feedback.

We offer further analysis of these models and many more in the section below on the Five Assessment Strategies. We also have a [set of Field of Learning slides](#) available, including an empty grid for your use.



ASSESSMENT SHIFT 2: The shift to multiple and varied measures.

The second assessment shift involves moving from single, narrow assessments to multiple forms of measures that are more varied, more developmentally nuanced, and better integrated. Although basic mastery of multiplication tables can be confirmed by a simple quiz or test, assessing competencies like creativity, social skills, and wayfinding abilities — the examples we used earlier — require a multifaceted approach.

Because multiple forms of measurement for any given competency are not the norm in traditional school models, we turn to state requirements for new drivers as a concrete and familiar example of such a system. At first blush, one might cite the road test as the qualifier for getting one’s license. However, over the past century, states have evolved systems of multiple, mandatory requirements to ensure the safety of drivers, passengers, and the public. In Massachusetts, for example, these requirements include the following:

- A written test of road rules (necessary for a learner’s permit)
- 30 hours of classroom instruction, with a test at the end
- Computer simulations (at some driving schools)
- 6 hours of official driving observation of another driver
- 12 hours of official driving practice
- 40 hours of additional practice, usually with parents
- A two-hour parent education class
- The road test

In addition, Massachusetts and many other states have instituted graduated driver licensing that begins with a learner’s permit, then graduates to various restricted, provisional, or probationary licenses in low-risk environments (such as driving with an adult, daytime driving, or driving without other adolescents).

In short, the Massachusetts “driving test” is actually a series of assessment-embedded learning experiences with multiple forms of measurement, addressing Content Knowledge and its application, Creative Know How, Habits of Success, and indeed Wayfinding Abilities, all through authentic performance.



Driver photo by [Lee Giles III, CC-BY-NC-SA 2.0](#)

The table below illustrates the multiple assessments involved with these Massachusetts requirements.

Assessment elements for MA driver’s license

MA requirements	Assessment involved
Written test of road rules — to get permit	Multiple choice, fact-based ; summative gateway to learner’s permit
30 hours of classroom instruction, with test at end	Formative feedback ; scenarios for understanding of skills, consequences; summative knowledge test
Computer simulations	Incorporating application of knowledge and skills
6 hours of official driving observation	Introduction to the authentic learning environment ; group/peer learning
12 hours of official driving practice	Practice loops in authentic environment with instant instructor feedback
40 hours additional practice, usually with parents	Practice loops in varying circumstances (different adult, different car), confirming transfer
2-hour parent education class	No assessment. Requirement is “programmatic”/seat time
The road test	Performance-based assessment in complex, authentic environment

(See also, in the *Introduction and Overview* of this series, our interpretation of learning to drive as an analogy for how the four MyWays domains all work together in the real world.)

Note the rich variety of assessments in the right column. The road test itself is actually just one component in a system of multiple, varied assessments embedded in a set of learning experiences. A system of this kind is an excellent model for assessing important life skills. The thoughtful creation of such assessment systems should be interwoven with the elements most vital to Whole Learning, student agency, and human development for life and work. The Latin root of “assessment” translates to “sitting next to” and “getting to know.” Accordingly, high-quality systems of multiple measures should allow students to observe mastery firsthand and receive feedback and coaching from those who know them, as well as empowering them to navigate their own learning.

The concept of multiple, varied measurement is essential in addressing broader and deeper competencies for several reasons. The driving test illustrates the highest level reason: to deal with broader, deeper competencies in a way that also furthers the development of meta-cognition and agency, learning and assessment must be multi-faceted. Summit Public Schools is one model that takes this richer view; for more on its model, see this case study on [Summit’s approach to varied assessments](#) and the *Multiple measures* box below. In addition, for some sets of competencies such as Habits of Success, individual measures have their own susceptibilities to bias. Thus, using two or more measurement methods and triangulating the results improves assessment. The [Mission Skills Assessment](#), highly reviewed by RAND researchers,⁷ measures six of the Habits of Success and Creative Know How competencies by triangulating student self-reports, teacher assessments, and situational judgment (in hypothetical scenarios) or other student-completed performance measures. For more on this, see the section on *Multiple measures* below.

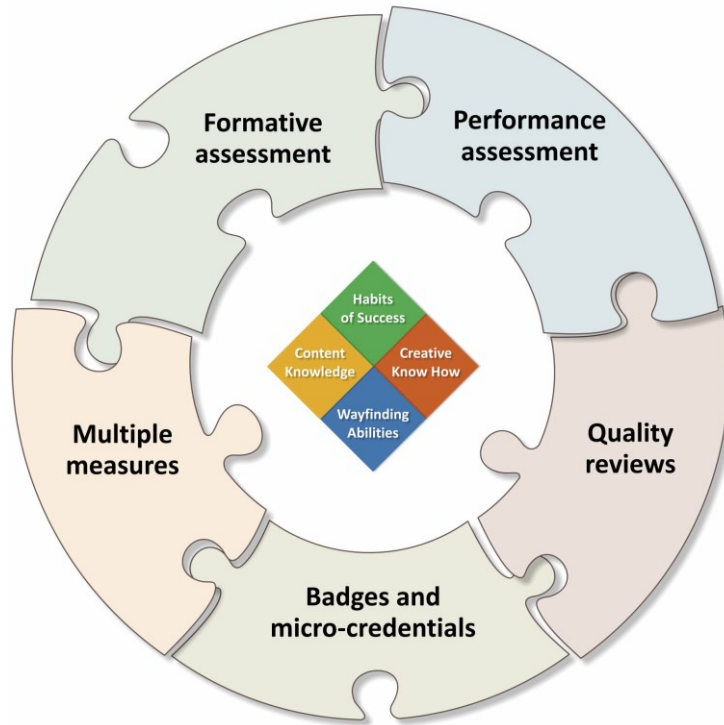
We turn our attention next to the set of Five Assessment Strategies that next generation educators can employ to address the shifts to both greater authenticity and multiple and varied measures.

Five Assessment Strategies that align with the two shifts

As with the assessment of driving and other important life skills, the next generation of student learning assessments will need to integrate learning and measurement, and use a variety of strategies in order to better develop and gauge progress on the different kinds of competencies involved. While the whole concept of assessment is complex, and assessment experts work in a range of specialized fields, educators thinking about next generation learning design can start by focusing on the following five key strategies:

The Latin root of “assessment” translates to “sitting next to” and “getting to know.” Accordingly, high-quality systems of multiple measures should allow students to observe mastery firsthand and receive feedback and coaching from those who know them, as well as empowering them to navigate their own learning.

Five strategies for next generation assessment



Integrating these five strategies will align assessment with holistic learning and help gauge and guide learner progress toward the new, broader goal-line:

Formative assessment

To provide the essential foundations for effective learning and personal development

Performance assessment

To provide the rich context for development and measurement of agency and capability

Multiple measures

To address the whole learner and the breadth of competencies within next generation learning environments

Badges and micro-credentials

To integrate “anywhere, anytime learning” within personalized learning approaches

Quality reviews

To ensure the quality of the learner experience when outcomes can’t be measured (and even when they can)

Representing these strategies as oddly shaped puzzle pieces is intentional — as is the fact that the pieces don’t always fit together seamlessly. Some of these strategies are much more developed than others, some include innovative measures around which there is little agreement, and others currently feature more strongly outside the formal education sector than within it. All of this makes sense given that Whole Learning designs are focusing more on complex learning in the Wider Learning Ecosystem, as well as incorporating a broader range of agency-related competencies. At the same time, the strategies can also overlap. Performance assessment can be used for formative purposes, badges can be a form of performance assessment, and multiple measures can incorporate any combination of the other strategies.

Assessment is a fast-moving field and the pace of change — as well as the introduction of new solutions — is likely to accelerate. As we see it, the organizational question today for next generation educators is: *What investments in assessment capacity-building will foster shifts to greater authenticity and multiple and varied measures, and facilitate the incorporation and integration of promising new assessment tools and practices from across the next generation learning space?* We address this capacity question later, because it is important first to understand the scope of the shifts and the strategies involved. The five strategies we are recommending combine two established imperatives for effective assessment — *Formative assessment* and *Performance assessment* — as well as three other, emerging strategies. We provide one-page primers and a practice example on each; but first, we offer additional comments on the role of each of the five strategies.

Formative assessment

Frequent iterations of measurement; verbal, written, and peer feedback; and, *perhaps most importantly, reflection* — aligned and integrated with learning design — have proven to be some of the most powerful strategies in enhancing achievement. *Formative assessment* is a vital tool for the learner as well as teacher; done right, it encourages the ownership of one’s learning. Accordingly, a crucial area of innovation for next generation educators is tying formative (as well as performance) assessment to learning progressions and rubrics for hard-to-recognize, hard-to-measure Creative Know How, Habits of Success, and Wayfinding Abilities competencies such as creativity, social skills, and finding needed resources.

Performance assessment

Although they are not commonly found in traditional school models, *Performance assessments* have been used for decades. Here, we focus on authentic performance assessments through which students demonstrate the broader range of knowledge and skills by performing real-world tasks that require those skills. Curriculum-embedded performance assessments, within quality project-based or Wider Learning Ecosystem experiences, provide the greatest opportunity for assessing broader and deeper competencies.

Multiple measures

Like our driver’s license example, any complex competency requires more than one assessment type. This is certainly true of hard-to-measure Creative Know How, Habits of Success, or Wayfinding Abilities competencies like entrepreneurship, self-direction, and navigating personal journeys. In addition, multiple measures will likely be needed to gauge both the capability and agency aspects of any competency. *Multiple measures* should include formative and performance assessments; they might also usefully include diagnostics, pre- and post-tests, adaptive testing, and summative tests for the purposes of end-of-topic, end-of-course, and progression-to-next-level evaluations. For specific measure types, see the *Multiple measures* row in the chart on p. 25, as well as the *Multiple measures* primer and practice box. In today’s world, most educators will also need to include tests for the purposes of accountability. Accordingly, next generation educators need the ability to identify, administer, evaluate, and integrate multiple and varied assessment elements, tailoring the mix with both the learner and the competency’s purpose and nature in mind.

Badges and micro-credentials

Any system of assessment should make learning visible, transparent, and portable. *Badges and micro-credentials* help accomplish these goals for the aspiring learner, for teachers and student advocates, and (as screening/signaling devices) for subsequent academic institutions and employers. Badge systems range from those requiring micro-performance assessments to those verifying participation in a quality experience likely to promote competencies. (For Wayfinding Abilities, for example, a badge related to exploring college options would require not just “visiting three colleges,” but “creating a list of four characteristics most important to my college experience, visiting three colleges, and providing reflections on how each matched up to [or altered] the four characteristics I chose to look for in a college.”) Two new job options for next generation educators include roles for those with the ability to design meaningful

internal (home-grown) badges, and those with the expertise and Wider Learning Ecosystem connections to vet and manage external (regional or standards-based) badges and stackable credentials, including those for out-of-school experiences.

Quality reviews

For some competencies, there is little consensus on valid, reliable, context-sensitive measures of student outcomes. For these hard-to-measure competencies, educators can still evaluate the learning experience itself to ensure quality and maximize the potential for student development. A quality review involves evaluating qualitative and quantitative data on the experience design and implementation against defined learning objectives and a logic model of how those objectives can best be met. In many cases, *Quality reviews* can be informed by excellent, established quality standards based on longstanding real-world practice such as those for internships (Big Picture Learning) and career exploration (Linked Learning) — see details and links in the *Quality reviews* primer. Reviews can focus on user experience and learning design, the school climate and culture likely to support competency development, and educator expertise in human development and brain science. While essential for hard-to-measure competencies, *Quality reviews* can also complement assessment of any student competency and foster program improvement.

Going deeper with the Five Assessment Strategies

These five strategies apply across all four MyWays domains and the 20 individual competencies, as next generation educators develop combinations and configurations that best support the development of broader and deeper competencies, individually and collectively.

While all five strategies are important to next generation assessment today, ***the first two strategies are at the core and are interconnected: expanded use of curriculum-embedded performance assessment that also includes strong formative assessment elements is central to the kind of learning and assessment needed to prepare students for the world after high school graduation.*** This foundation of performance and formative assessment makes possible more authentic assessment of deep and durable learning in Content Knowledge and Creative Know How, as well as offering opportunities for learners to develop and gauge progress in Habits of Success and Wayfinding Abilities in ways that might be lacking in more transmission-based learning and traditional assessment.

The following section provides a two-page spread for each of the five strategies. The one-page primers are not intended to be comprehensive nor do they provide in-depth analysis of each strategy's technical merits. Rather, their purpose is to get design teams thinking and discussing the level of variety, nuance, and integration needed to develop assessments for broader and deeper competencies. The facing practice boxes provide a glimpse into the efforts of just some of the Assessment for Learning Project's members to develop and reflect on new assessment approaches and tools. (See p. 27 for more about ALP.)

Note: after the Five Assessment Strategy primer and practice box pages, we provide a chart summarizing key points from this section on one page. This summary is intended to promote discussion rather than capture all possible characteristics and examples of the various strategies.

Formative Assessment

The foundational role of formative assessment in effective learning design

“Formative assessment is both good teaching and good learning, because it empowers students to self-assess and guide their own learning.”

—McREL, [Re-Balancing Assessment](#)

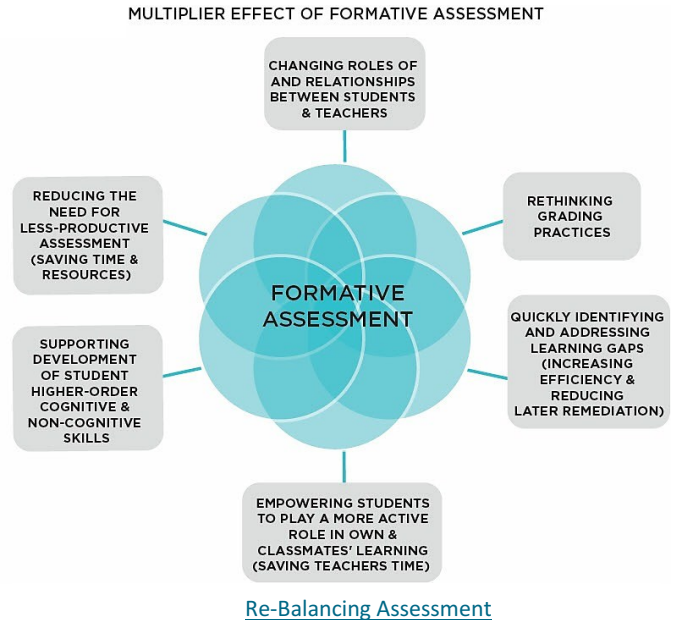
Why it’s important:

Formative assessment is a process involving actions undertaken by teachers and students within a learning activity to gauge progress, provide feedback, and adjust further learning. Frequent iterations of measurement, feedback, and reflection — aligned and integrated with learning design — have been shown to be one of the most powerful strategies in enhancing achievement, especially for lower-performing students (See [Hattie](#)). In addition to engendering deeper, more durable learning of knowledge and skills, formative assessment is a vital tool for the learner as well as the teacher; done right, it encourages the ownership of one’s learning. Accordingly, a crucial innovation area for next generation educators is tying formative assessment to learning progressions and rubrics for hard-to-recognize, hard-to-measure competencies such as creativity, social skills, and wayfinding.

The graphic to the right underscores the multiplier effects of formative assessment.

Examples of this type of measurement:

- Teacher-initiated formative feedback, including: gathering rich evidence of student progress toward transparent learning goals through a variety of means (observation, checks for understanding, questions, student response systems); providing feedback that is rapid, descriptive, and focused at the task, process and self-regulation levels; and using feedback to adjust learning and instructional activities.
- Student self-assessment and self-reflection, where students are deeply involved in gauging their own progress toward learning goals, and reflecting on their own learning processes. Includes student-run conferences.
- Peer-assessment, including gallery walks, feature critiques, pair-and-shares using rubrics, or even group discussions where students give each other feedback on ideas that are then further developed.
- Digital forms of formative feedback and adaptation of instruction through adaptive software and adaptive learning games or simulations that are set up to respond to performance as it happens.



Through the MyWays lens:

- The focus on process as well as product provides the opportunity for the development and assessment of *capability* and *agency*, which comprise competence. See more concrete examples in the Levers section of Report 11.
- Self-assessment and self-reflection are particularly important for the new goal-line of activating agency across all four competency domains. As [Rick Stiggins](#) says, “If you want to appear accountable, test your students. If you want to improve schools, teach teachers to assess their students. If you want to maximize learning, teach students to assess themselves.”
- For inspiration on the educator role, try turning to **teachers and coaches** of novice-to-expert, practice-based subjects like music, drama, and sports; **youth developers** who work with youth on self-reflection; and **special educators** who work with strategy-based (read: process-based) instruction.

A few resources as food for thought:

- Ron Berger et al., [Leaders of Their Own Learning: Transforming Schools Through Student-Engaged Assessment](#) and [website](#).
- McREL/Measured Progress. [Re-Balancing Assessment: Placing formative and Performance Assessment at the Heart of Learning and Accountability](#).

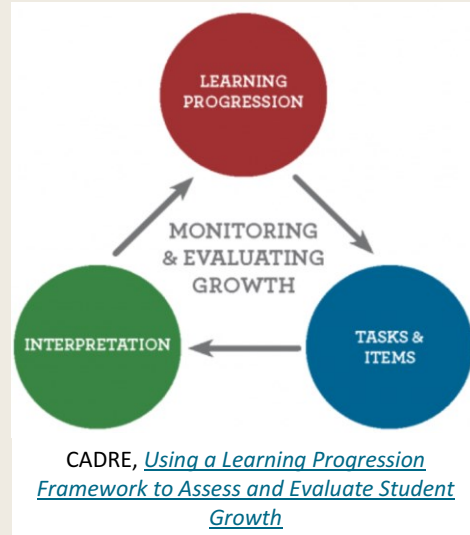
FOR MORE RESOURCES, see the [MyWays website](#).

FORMATIVE ASSESSMENT

Developing Deep Educator Expertise in Formative Assessment

Assessment is often thought of as an event or an instrument, rather than a process. As a result, many define formative assessment by when it occurs, instead of how it occurs. A small quiz given at the end of class — an exit ticket — can show whether students understood what was taught, but it does not shed much light on the more important question: how will students learn the concepts that come next? This is the job of formative assessment, and it requires teachers to engage deeply with the material to be learned and the students who are grappling with it. Quality formative assessment, then, is less about the tool and more about the educator expertise in understanding how students progress. So how do educators develop this expertise and incorporate formative assessment into their instructional practice? Here are a few examples:

Learning progressions. In order to give good feedback, teachers need to know how students develop increasingly complex understanding of the material they're learning, at a very incremental level. One very powerful way to accomplish this is for teachers to dig deeply into student work, deconstructing what it shows about how different students are making sense of the material. The [Colorado Education Initiative](#) is helping a multi-district collaborative of teachers map out these developmental progressions, working side by side with researchers from the University of Colorado Boulder [Center for Assessment Design, Research, and Evaluation](#) (CADRE)'s [Learning Progressions Project](#). Teachers then use the learning progressions in their classrooms, testing whether the documents actually capture the developmental progression for the students in their class, revising as needed, and discovering what is generalizable and what is unique to each student. This is a rigorous *process of shared inquiry* into student learning. It creates not only a tool, but a deep *expertise in how the understanding of complex concepts emerges* in developing minds. For an excellent overview to this project and other uses of learning progressions, see this [EdWeek](#) article.



Quality feedback (or, feed-forward). Formative assessment has the greatest impact when there is a culture of feedback among young people as well as adults in classrooms and schools. [Henry County Schools](#) in Georgia is building this culture by inviting students to give “feedback on feedback.” Henry County is training teachers in formative assessment using a [Feedback Loop protocol](#). As part of the process, after students receive feedback from teachers they are prompted to give reciprocal feedback on how helpful the teacher’s feedback was. The system is designed to provide real-time information to improve the quality of feedback, while at the same time deepening the culture of learning.

Self-assessment and peer feedback. Formative assessment works when it prompts students to better understand criteria for success, and to think deeply about their own learning as it relates to these criteria. Self-assessment and peer feedback are powerful tools to accomplish this. Through the [Student Agency in Assessment & Learning](#) project at [WestEd](#), teachers are engaging in these processes themselves and with students to identify ways they can improve opportunities for students to self-assess and give feedback to peers. Teachers take video of students learning in their classrooms, and then they reflect on the student experience using “continua” that describe criteria for high-quality self assessment and peer feedback. After they’ve reflected on their own classroom, they share the video with peers for feedback. Together, the teachers generate strategies for how to integrate these practices more deeply into their classrooms.

Performance Assessment

The centrality of performance assessment in measuring capability and agency

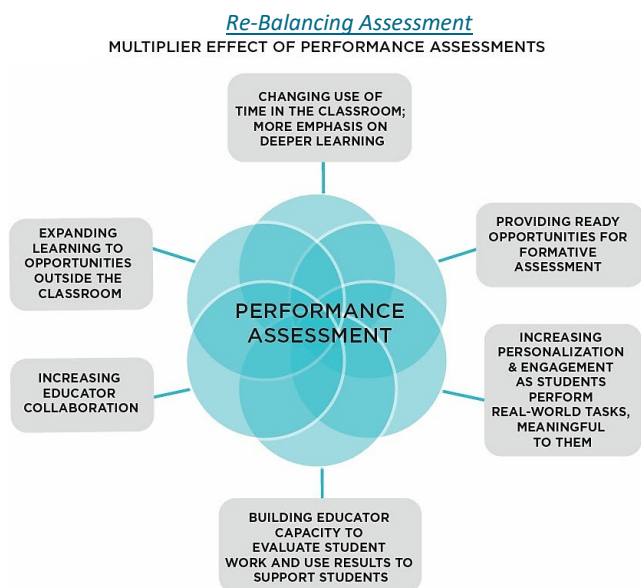
“Ultimately, the promise of CEPAs [Curriculum-Embedded Performance Assessments] is that they provide a more motivating, robust, and balanced way to measure student learning.”

—McREL, [Re-Balancing Assessment](#)

Why it’s important:

Performance assessment (PA) is the process by which the learner demonstrates knowledge and skills by doing real-world tasks that require those skills. It is associated with features such as real-world scenarios; authentic, complex process; higher-order thinking; authentic performance; and transparent evaluation criteria. Although they are not commonly used in traditional school models, PAs have been in use for decades. Curriculum-embedded performance assessments (CEPA), within project-based or other experiential learning, provide the greatest opportunity for assessing broader and deeper competencies.

The graphic below underscores the multiplier effects of PA.



Examples of this type of measurement:

- Assessment of real-world tasks or junior versions of them using assessment rubrics — both extended, curriculum-embedded models (CEPA) and bounded PAs from task banks. (See Report 11 on junior versions.)
- Public performances and public exhibitions of the outcomes of project work.
- Portfolios, student logs, and journals that document the learning process, reflections, and revision of work.
- CEPA is incorporated into some deeper learning, project-, work-, and service-based learning, and

competency-based learning models, such as [EL Education](#), [Envision](#), [High Tech High](#), [New Tech Network](#), [Sanborn](#) (NH), and [Summit Public Schools](#).

Through the MyWays lens:

- CEPA has much in common with the Whole Learning design presented in Report 11 (applied, holistic, authentic, complex, and connected to the adult world). The use of PA, like that of holistic learning, is central because it engages the full range of broader, deeper competencies and provides rich opportunities for student agency and capability. CEPA follows all seven Whole Learning principles, and is highly aligned with the Levers for Capability & Agency described in Report 11.
- Like Whole Learning, PA carries within it a paradox — that it should be experienced by learners as holistic and integrated, yet must be carefully designed to assess specific learning outcomes at the right “grain size” to effectively evaluate them. Therefore, educators need effective tools for designing high-quality PA (clear learning targets; standards for content, skills, and habits; process protocols; product rubrics; progressions; and conceptual maps) *and* coaching on how to use them without allowing the student learning and assessment experience to be disjointed and molecularized.

A few resources as food for thought:

- Bob Lenz, [Transforming Schools: Using Project-based Learning, Performance Assessment, and Common Core Standards](#) and [website](#); Ron Berger, [Leaders of Their Own Learning: Transforming Schools Through Student-Engaged Assessment](#) on EL schools [website](#).
- The [Center for Collaborative Education’s Microcredentials Stack](#) in Performance Assessment Design, Performance Assessment Validation, and Competency-based Rubric Design offers guidance, open source resources, and potential to earn micro-credentials.
- The Buck Institute for Education’s [Resource List: Assessment in PBL](#).
- Chris Sturgis, [Learning Progressions: Are Student-Centered State Standards Possible](#), offers a useful discussion of and links for progressions, in addition to examining the accountability issue.
- For further tools related to PA, such as learning progressions, rubrics, developmental frameworks, and rich task banks, see the “Implementation Tools” box at the end of this report. These tools are more plentiful for Content Knowledge and Creative Know How, but we also link to emerging work in the other two domains.

FOR MORE RESOURCES, see the [MyWays website](#).

PERFORMANCE ASSESSMENT

Measuring Creative Know How with Performance Assessment

Deeply committed to a broader definition of student success, the educators at [Two Rivers Public Charter School](#) in Washington, D.C., are finding ways to measure not only core content and basic skills, but also three additional dimensions of student success: critical thinking and problem solving, collaboration and communication, and character. Performance assessment is a natural fit with their project-based and problem-based curriculum. And it provides the opportunity for authentic assessment that engages a broad range of these success dimensions.

To assess Two Rivers' critical thinking and problem solving dimension — aligned with the MyWays domain of Creative Know How — the school designed a series of performance tasks, rubrics, and instructional supports. The performance assessments measure students' ability to transfer these skills to new situations through content-neutral tasks. For a great description of the work behind developing the performance tasks, in collaboration with the Stanford Center for Assessment, Learning, and Equity (SCALE), see this [EdSurge MyWays series article](#). Descriptions of the competencies within the critical thinking and problem solving dimension — schema development, decision making, effective reasoning, problem solving, and creativity and innovation — as well as the rubrics and tasks tailored for pre-kindergarten through eighth grade, are available in the organization's [Deeper Learning Assessment folder](#).



What's the impact?

According to the director of curriculum and instruction at Two Rivers, Jeff Heyck-Williams, “Implementing performance assessments of critical thinking and problem solving has galvanized teachers to be more intentional about their explicit teaching of these skills.” Specifically, he continues, “Teachers have begun to implement thinking routines that cross disciplines and are intended to become habits of mind that students can apply widely in various situations.” One example is how students are developing a thinking routine for problem solving in any subject and for any project, which allows them to transfer the skill in flexible ways: First, identify what they know about the problem, what they need to find out, and what ideas they have for approaching the problem. Next, monitor their problem solving process to determine if they are on a fruitful path. Last, evaluate not only their solution but the process that they took to arrive at a solution.

Understanding student performance over time

Assessing performance against pre-defined tasks isn't a panacea, however. When it comes to collaboration and communication, and character skills, Two Rivers is moving down a somewhat different path. They are currently, for instance, using student self-assessment and teacher assessment against [rubrics](#) to understand how student performance over time in a class reflects the school's *Scholarly Habits* (I work hard; I am responsible and independent; I care for my community; I am a team player). While not involving pre-defined, content-free tasks, this is similar to the more longitudinal and embedded performance assessment practiced by Summit Public Schools and Envision Education, and employed in the Sanborn CARES approach described in the box at the start of Report 7. The school team is still grappling with the best way to capture data related to these success dimensions, but Heyck-Williams sees the next steps for this work involving consulting social-emotional learning (SEL) frameworks to refine the Scholarly Habits definitions, and engaging their staff, students, and families in creating concrete “look-fors” in student performance.

Multiple Measures

The importance of multiple measures within next generation learning environments

“Whenever possible, we recommend using a plurality of measurement approaches. While time and money are never as ample as would be ideal, a multi-method approach to measurement can dramatically increase reliability and validity.”

—Duckworth and Yeager, *Measurement Matters*

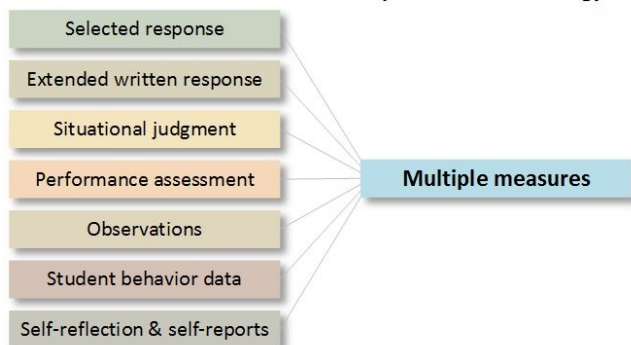
Why it’s important:

There are many important reasons to use multiple, diverse measures: to cover a broader range of very different competencies; to address the student agency element of competency, which requires more nuance and context; and, as noted above, to enable triangulation of results, which is good practice in any event, but particularly important while some types of measures are still in development. And, of course, different types may be suitable for diagnosis, for demonstrating mastery, and for other summative and accountability purposes. Accordingly, next generation educators need the ability to identify, administer, evaluate, and integrate multiple assessment elements.

Examples of this type of measurement:

- For Content Knowledge and Creative Know How, while formative and performance assessment are central, models designed around these domains also use traditional selected response and essays, as well as adaptive software and simulations, to develop and gauge progress on knowledge and skills.
- For Habits of Success and Wayfinding — as well as agency in general — the field has yet to converge on common valid, appropriate measures for guiding learning and development in these areas. Educators might measure student perseverance, for instance, using 1) performance assessment employing a rubric in a PBL approach similar to [EL Education’s](#); 2) [Newton’s Playground](#), a computer simulation that stealthily measures learning of physics concepts, as well as student perseverance indicated by trying alternative strategies when obstacles occur; or 3) [The Partnership in Education and Resilience \(PEAR\) Institute’s Holistic](#)

Possible assessments in a multiple measures strategy



[Student Assessment](#), other self-reports, or teacher observations).

Through the MyWays lens (mostly focused on Habits of Success, Wayfinding Abilities, and agency):

- On the measurement of agency traits, as Angela Duckworth (known for her work on grit) recently warned, [we are not “there” yet](#). Such measures, usually based on self-reports, are susceptible to issues of reliability, validity, fakeability, and context-dependency. More embedded forms, such as performance assessment, also need careful consideration.
- At the same time, researchers and innovators are working on improving and targeting such measures, drawing on international practice to improve performance assessment, such as through social moderation, and developing forced choice pairwise preference, and other more sophisticated forms of self-reports. Major SEL/agency/character players are also developing what they hope will be new, improved measures. See this report’s “Moving forward” section.
- Other innovative approaches also show promise — including situational judgment assessments, computer simulations, learning games, and MUVES (multi-user virtual environments); the use of adaptive learning software for basic skills; evidence-centered concept maps for assessing higher-order thinking; and real-time mining of learning-related, online behavior data.
- There is also room for this field to borrow more from other fields, including **early education**, **youth development**, **developmental psych**, **workplace HR**, and **special education** (especially relating to executive functioning, ADHD, and spectrum disorders).

A few resources as food for thought:

- Valor Collegiate Academies, a leader in Habits development, has an excellent [Resource Guide](#) providing public access to a wealth of documents relating to its well-regarded Compass model for student development.
- Reviews of agency and Habits measures by educators and youth developers include [Evolving Assessments for a 21st Century Education](#), CCR; [Measuring 21st Century Competencies](#), Asia Society and RAND; and [From Soft Skills to Hard Data](#), Forum for Youth Investment.
- Coming in winter 2018: CASEL/[Measuring SEL’s Assessment Guide for Educators](#), an online tool for selecting and using currently available SEL measures.
- To keep on top of emerging agency-related measures, sign up for Transforming Education’s [MESH newsletter](#), or join the Measuring SEL’s [Collaborators’ Network](#).

FOR MORE RESOURCES, see the [MyWays website](#).

MULTIPLE MEASURES

Assessment is Wide and Deep at Summit Public Schools

[Summit Public Schools](#), a charter management organization with 13 schools in California and Washington, is a thought- and practice-leader in the next generation learning space, supporting over 330 schools across 40 states through the Summit Learning Program. Summit is taking a multiple measures approach to assessing their four elements of college readiness: Cognitive Skills, Content Knowledge, Habits of Success, and Sense of Purpose. Their use of multiple, varied measures spans not only across these four elements, but also within them.



Cognitive Skills are measured using a [Cognitive Skills Rubric](#), developed in partnership with the Stanford Center for Assessment, Learning, and Equity (SCALE). The rubric enables teachers to assess student proficiency of the 36 skills within every subject in every grade level. Cognitive Skills are learned primarily through projects and the rubric is most often used in connection with performance assessment of projects.

Content Knowledge is measured through a set of assessments embedded within the [Summit Learning Platform](#), an online tool to personalize learning through goal-tracking, playlists of self-paced content, and deeper learning projects. Both diagnostic and culminating assessments are available on-demand, and students must correctly answer 8 of 10 questions on the content assessment to pass.

The organization does not yet have a system-wide, defined approach for assessing **Habits of Success**. As Chief Academic Officer Adam Carter [wrote](#) in 2015, Habits of Success “are the invisible thread that ties together the fabric of relationships and organizations — they are bound intimately with motivation and achievement. They count, but we don't yet count them. We should.” Summit has been moving forward methodically in this area (in a design thinking, test-and-see way, not a research-extensively-before-doing-anything way) because the research community has not yet coalesced around how these habits are best developed and measured. Carter noted that he believes “**a portfolio approach to assessing these Habits of Success is possible**, and we have laid much of the groundwork to collect such assessment information.”

Summit continues to pursue this portfolio or *Multiple measures* approach because it respects the inherent complexity and nuance of Habits of Success. The organization has piloted various ways to measure habits including Emotional Intelligence, Self-Directed Learning Behaviors, Learning Strategies, Academic Mindsets, and School and Classroom Culture. Summit relies on both quantitative and qualitative information, piloting a mix of surveys (self-report and adult-observation), back-end data analysis of online behavior metrics, tracking of goal-setting and achievement measures, badges, and “pop-up” questions in the learning platform.

Through the pilots, the Summit team has also concluded that schools need more of an “actionable, in-the-moment” approach to assessment. The data are fed into 1:1 mentoring between students and educators and weekly dashboards within the Summit Learning Platform to enable “understanding through conversation” with students, as well as rapid-cycle interventions. A second approach involves culminating, high-stakes, summative, but also future-facing assessment: Summit is exploring the possibility of a capstone graduation requirement like a portfolio oral defense. Portfolios are, by their nature, composed of multiple measures of proficiency, and they would add a cumulative element to Summit’s multiple measures of Habits of Success.

Badges and Micro-credentials

The benefits of integrating badges and micro-credentials for personalized learning

“We think Badges Will Be Big. They provide an efficient way for employer groups to signal requirements for job clusters and a way for learners to accumulate and share a growing portfolio of skills.”

—Tom Vander Ark, [Vander Ark on Innovation](#)

Why it’s important:

Students pursuing Whole Learning for the full range of success competencies spend more of their time learning outside of school in the Wider Learning Ecosystem (WLE). Badges are an assessment and credentialing mechanism that can help make the results of this learning visible, transparent, and portable. (For example, clicking on a digital badge can link not only to criteria for earning it, but also to evidence from a completed project and comments from collaborators or evaluators.) Some educators are looking to create badging systems appropriate for assessing and encouraging Habits of Success and agency within school settings. External badging and other more structured forms of work-related micro-credentials are also increasing as ways to capture new aspects of personalized learning pathways and informal learning.

Examples of this type of measurement:

- The open badging movement was pioneered by the [MacArthur Foundation](#), [Mozilla Open Badges](#), the [Badge Alliance](#) (now IMS Global), and [HASTAC](#).
- Out-of-school-time programs have been early adopters of badging. [Chicago City of LRNG](#) awards badges for quests and experiments that roll up into city-level STEAM badges. [Providence After School Alliance’s](#) badged learning experiences, from science internships to business pitches to Social Venture Partners of RI, are accepted for credit by the school system, and recognized by several state higher education institutions. For more on a variety of community learning networks, including [Hive](#), [Education Innovation Clusters](#), and [Remake Learning](#), see the WLE section of Report 11.
- Micro-credentials, also known as stackable credentials, include educational credit, non-credit, and industry certifications. The [Work Readiness Credential](#), for example, is available in 23 states. The National Academy Foundation’s [NAFTrack](#) certification includes performance assessment resulting in artifacts, which can be added to portfolios, and workplace observations, leading to NAFTrack Certified Hiring benefits. [Jobs for America’s Graduates](#) supports 88 competency-based modules ranging from workplace competencies to character development.
- [The Credential Engine](#), a non-profit alliance led by Lumina, George Washington University, and the Business Roundtable, is now creating a credential

registry that will enable job seekers, students, workers, and employers to search for and compare thousands of credential programs through the [Workit](#) search app, to be launched in winter 2017. Other efforts to strengthen credentialing systems and provide individuals with ways to navigate competency-based micro learning include offerings like [Workcred](#), [Degreed](#) and [Pluralsite](#). See Michael Horn’s take in this [Forbes article](#).



[Using Digital Badges as Stackable Career Credentials](#),
Association for Talent Development

Through the MyWays lens:

- The badging ecosystem is still evolving; some groups are working to align performance learning standards to those used in schools, while others argue that informal learning should remain interest-driven and student-directed. “Too quick a move towards badges runs the risk of destroying the complex but fragile ecosystem within which participatory learning thrives,” claims USC’s Henry Jenkins in [this EdWeek article](#).
- Several next generation schools have explored badging to recognize work in Habits of Success or Wayfinding. [Summit](#) has piloted badges for note-taking and meta-cognition skills, as well as expeditions that reflect students’ passions, like coding or gardening. [Valor](#) awards badges for its Compass character activities within an ongoing, mentored program of progressively advancing practice and reflection. Also see Del Lago’s Competency X on the next page, and more examples in the [EdWeek](#) article just below.

A few resources as food for thought:

- [Students Earn Digital Credentials for Adding New Skills](#), *Edweek* article by Michelle Davis
- [Expanding Education and Workforce Opportunities through Digital Badges](#), Alliance for Excellent Education.
- “[Assessing Learning in Digital Badge Systems](#),” a chapter in Design Principles Documentation Project’s *Four Functions of Digital Badges Report*.
- [Understanding CTE and Stackable Credentials](#), Transitions Academy.

FOR MORE RESOURCES, see the [MyWays website](#).

BADGES & MICRO-CREDENTIALS

Badging Scholars in Science and Engineering Practice

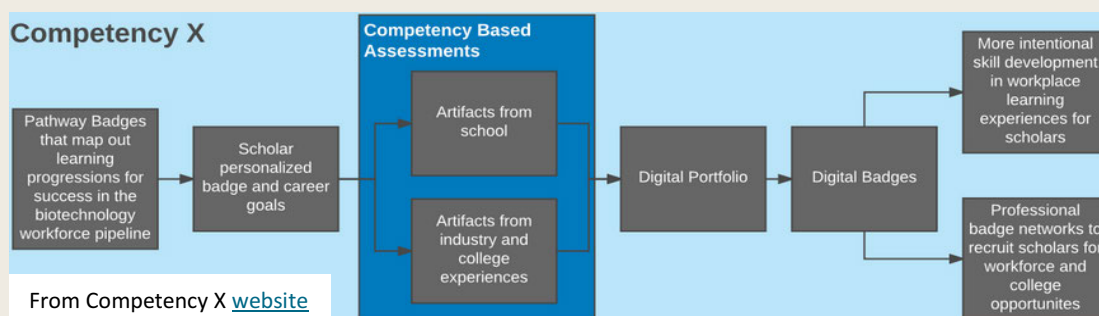
“Digital badges fill in gaps for how we describe what scholars know and can do in the real world. Traditionally, most scholars only have a transcript of coursework to represent what they can do. Digital badges unbundle the competencies within both courses and workforce experiences to help fill in the gaps of larger credentials (e.g., degrees and certifications). This allows them to be more precise about what a learner is capable of accomplishing.”

— Alec Barron, Del Lago Academy⁸

[Del Lago Academy - Campus of Applied Science](#) (DLA) is a district high school in Escondido, CA, created to “engage scholars in real-world learning in order to prepare them better for success in college and career.” DLA was designed around the [Next Generation Science Standards \(NGSS\) for Science & Engineering](#), with interdisciplinary projects and biotech industry internships at the center of the academic program. As the DLA team reflected on how to support students to transfer knowledge and engage in scientific inquiry, they realized that there was a bigger question to answer: How do people actually learn to become practicing scientists?

To answer this question, DLA held a summer institute in 2016 with students, teachers, and professional scientists. The institute began by posing this question to partners: When students come into your professional context, what skills, knowledge, and dispositions do you want them to have in order to be successful with the tasks you’d assign a new employee or intern? The result was DLA’s [Badge Progression](#), which identifies applied skills in three categories: Experimenting, Analyzing, and Communicating.

DLA’s [Competency X](#) website now provides a hub for this system, which also houses digital portfolios through which students track and reflect on evidence of their competency. Some foundational badges are required for all students; others can be pursued based on a student’s individual interests and goals. Badge requirements are co-created and validated by industry and college partners, and badges are earned when scholars meet specific criteria for these practices in their digital portfolio. Badges cover a range of attributes, from technical skills (e.g., measuring small volumes) to Habits of Success (e.g., reflective thinking), and everything in between. For examples, see [Skeptic Level 1](#) and [Elevator Pitch](#) badges.



Students collect evidence during school labs and internships. Using their 1:1 iPads and a free blogging platform, students capture photo and video evidence of their lab work and write short blog posts describing this evidence. When a student feels she has collected the evidence of learning to earn the badge, she submits it for review by a teacher, who either grants the badge or provides feedback on how the evidence of learning needs to be improved. The [digital badges](#) are used by scholars as evidence of their competency with the knowledge and skills required to be a scientist for further internships and college applications. DLA highlights the [benefits of a digital platform](#) in linking to evidence of performance, allowing learners to reflect on their progress and solicit feedback from a community of practice, and enabling prospective employers to assess the credibility of the credential and see individual skill development. Find out more in this [Competency X overview video](#) (9m) and at the [Competency X blog](#).

Quality Reviews

The added value of evaluating programs as well as students

“Shifting the focus of measurement away from only examining outcomes, towards the broader environmental aspects of learning conditions, could prove to be a powerful outcome.”

—Michael Fullan, *Towards a New End: New Pedagogies for Deep Learning*

Why it’s important:

A quality review of learning experiences is a systematic process of gathering and analyzing data (qualitative and quantitative) to inform learning and action. Importantly, it helps uncover insights related to context, the “how” and “why,” what works when, and what factors help or hinder. When unable to assess hard-to-measure competency outcomes directly or when lacking confidence in the reliability of assessments, it is particularly important to conduct a rigorous evaluation of program design and implementation, holding it up against learning objectives and a logic model of how those objectives would be met. Assessment can take many forms, and at times, a quality review of learning experiences is the only available way to foster student growth and program improvement.

Examples of this type of measurement:

- Evaluation of learner experience, looking at how and why, and what seems to work when and in what context. What opportunities and depth does the experience offer from the learner perspective?
- Evaluation of inputs and activities using quality standards such as those created for internships (see Big Picture Learning’s [ImBlaze](#) and [Learning Through Internships Toolkit](#)); career exploration (see [Linked Learning’s report](#) on the criteria for designing and assessing the quality of a career exploration experience); or service learning (see [National Youth Leadership Council’s standards](#)).
- Evaluation of school climate and culture, especially in relation to support for agency, Habits of Success, and Wayfinding Abilities, using established and newly-developed school surveys, reflections, and so on.
- Evaluation of educator understanding and use of holistic learning, next generation assessment, and the fundamentals that stand behind them — that is, brain and learning science, human development, growth mindset, and new educator roles in student-directed learning.
- The [Scottish system](#) of “experiences and outcomes” incorporates quality review of learning experiences at a national level, recognizing “the importance of the quality and nature of the learning experience in developing

attributes and capabilities and in achieving active engagement, motivation and depth of learning.”



Through the MyWays lens:

- Work-based learning for all students is a critical element for broader student competency. Educators looking to evaluate internship designs will find excellent recent “gold” [standards for high-quality internships](#) from the National Academy Foundation, and tools for design, compliance, and quality assurance embedded in [ImBlaze](#), Big Picture’s new internship management platform. ConnectEd’s [Guide for the Linked Learning Pathway Quality Review Process](#) also provides useful insight into the importance of quality processes.
- As new approaches to learning continue to develop, we will continue to look for ways to tell if and how they are working. For questions and suggestions — though no hard answers — see, for example, [How Should We Measure the Impact of Makerspaces?](#)
- Even for competencies with reliable outcome measures, as Srikanth Gopal of FSG noted, [quoting a colleague](#), “You can’t fatten a pig by weighing it.” Nuanced quality review can get at the situational and contextual factors that underlie the broader competencies and student progress on developing them.
- For inspiration, Fullan also [suggests](#) looking at **working practices in knowledge-based organizations**. Finland’s view of “**teachers as action researchers**,” [as noted by NGLC’s Andy Calkins](#), is also relevant here.

A few resources as food for thought:

- Srikanth Gopal, “[Stop \(Just\) Measuring Impact, Start Evaluating](#),” FSG Consulting blog.
- The Learning Accelerator, [District Guide to Blended Learning Measurement](#), offers a step-by-step guide to when, what, whom, and how to measure.

FOR MORE RESOURCES, see the [MyWays website](#).

QUALITY REVIEWS

Quality Review at the Learning Environment and Community Level

What are the characteristics of a successful student of the Hawai'i public schools? Like many districts and states, Hawai'i has adopted a [framework](#) to answer the question, called BREATH, or in Hawai'ian, HĀ. BREATH is an acronym for: Belonging, Responsibility, Excellence, Aloha, Total well-being, and Hawai'i.

"When I walk into a Hawai'i public school, I want to close my eyes and know that I am in a school in Hawai'i...and not somewhere else."

– **Suzanne Mulcahy, Complex Area Superintendent (2015)**

This framework differs from many we have seen not only in its language and explicit responsiveness to Hawaiian history and culture, but also in how it conceives of these measures. HĀ is a vision not only for



Reproduced from the HIDOE website

individual student readiness, but also for community readiness. As the [Board-adopted policy](#) states, the framework “reflects the Department of Education’s core values and beliefs in action throughout the public educational system of Hawai’i.”

Values like belonging, responsibility and aloha (welcoming) are not only individual, but reciprocal — they are expressed at the community level as well. For this reason, assessing these outcomes only at the student level would be incomplete. As an educator explains in this helpful [video](#) (6m), “People ask, what is HĀ. The first things that I think of are systems outcomes.” In addition, as the HĀ policy states, “All six outcomes are interdependent and should not be used separately.”

The framework was developed through inclusive community dialogue, and as the Hawai'i public schools develop their approach to assessing it, they are beginning with assessments of the learning environment rather than students.

The process invites members of the community — both within and beyond the walls of the school — to reflect on the ways that the HĀ values are expressed in the learning environment. Using this process across several communities within Hawai'i, HIDOE is developing a common understanding of what are the elements of a HĀ learning environment. As more communities begin to adopt HĀ (as of this writing it is a voluntary pilot), they will draw on this common understanding to assess their own learning environments — a powerful form of quality review.



[Student sharing Hawaiian greeting](#)

Five Assessment Strategies for broader and deeper competencies

	Points and principles	Illustrative examples
<p>Formative assessments</p> <p>To provide the essential foundations for effective learning and personal development</p>	<ul style="list-style-type: none"> • Focus on process over product • Rapid, ongoing feedback loops • Adapting subsequent instruction or experience • Requires clear, transparent learning goals, progressions • Action and improvement cycle motivates growth mindset 	<ul style="list-style-type: none"> • Teacher checks for understanding: traditional quizzes, exit tickets, and student response systems, with follow-up. • Self-assessment/reflection: students gauge their own progress and reflect on their learning process; includes student-led conferences. • Peer assessment: gallery walks, feature critiques, pair-and-shares, and group discussions. • Revision, iteration, re-dos to mastery, improved quality. • Digital adaptive software or simulation with feedback loops; digital collection of behavioral data.
<p>Performance assessments</p> <p>To provide the rich context for development and measurement of agency and capability</p>	<ul style="list-style-type: none"> • Engages student agency • Less “proxy”; closer to complex adult world • Longer-term performance assessments embedded in PBL and competency-based curriculum offers the most opportunities to address broader competencies • Still establishing best “grain size” for learning and assessment 	<ul style="list-style-type: none"> • Assessment of real-world tasks or junior versions thereof, using assessment rubrics and learning progressions. • Public performances, especially if assessed in part by members of professional communities of practice. • Public exhibitions of project work, including learner explanation, demonstrations, and response to queries. • Student logs and journals that document process, reflection, and revision.
<p>Multiple measures</p> <p>To address the whole learner and the breadth of competencies within next generation learning environments</p>	<ul style="list-style-type: none"> • Multiple measures create rounded learner profile • Coverage across diagnostic, formative, and summative • Different types of measures are needed for different types of competencies and aspects of competency such as agency • On Habits of Success and Wayfinding, as well as agency generally, tools are not yet “there”; multiple measures create checks 	<ul style="list-style-type: none"> • Performance and formative assessments (above) are vital for authenticity, complexity, and integration, and need to be augmented with badging and quality reviews (below). • Also useful as multiple measures: extended written response, situational judgment (hypothetical scenarios), direct observation, behavioral data (or “bio-data”), improved self-reports, forced choice, self-reflections, teacher reports, and so on. • Includes required accountability testing, pre- and post-tests, and non-performance summative assessment. • “Assessment” may not even look like traditional assessment, from stealth assessment in computer gaming to informal teacher assessment of student interests and real-world capacities.
<p>Badges and micro-credentials</p> <p>To integrate “anywhere, anytime learning” within personalized learning approaches</p>	<ul style="list-style-type: none"> • Badges can recognize participation, skill, and achievement • Flexible assessment for learning outside school, especially in the community (PASA, Cities of LRNG), at work (internships), and in agency and Habits of Success areas • Requirements: modular, visible, transparent, and portable 	<ul style="list-style-type: none"> • All badge types are useful, but badges based on demonstrated and verified skills (micro performance assessment) hold the most promise for broader competencies. • Examples of badging systems: Microsoft, Chicago City of LRNG, Summit Public Schools, and Del Lago Academy’s Competency X. • Other micro-credentials: Career readiness certificates (JAG’s competency modules to ACT WorkKeys); also industry, educational, and non-credit certificates that can be “stacked” to lead to degrees (see NAFTrack).
<p>Quality reviews</p> <p>To ensure the quality of the learner experience when outcomes can’t be measured (and even when they can)</p>	<ul style="list-style-type: none"> • In Habits of Success and Wayfinding as well as agency overall, student outcome tools are still emerging. Indeed, schools are only now adding coverage; evaluation of inputs, experiences, and outputs is likely to increase quality and effectiveness even if student outcomes cannot be measured precisely 	<ul style="list-style-type: none"> • Evaluating learner experience, looking at how and why, what seems to work when, and in what context. • Evaluating inputs and activities by using quality standards documents for internships (BP), career exploration (LL), service learning (NYLC), and so on. • Evaluating the school climate and culture to support agency and Habits of Success (established and new survey tools). • Evaluating educator understanding and use of holistic learning and next gen assessment.

Moving forward: cautions, potential, and ways to work together

The primers just presented on the Five Assessment Strategies provide samples of what’s going on in the field, the practice boxes provide a window into some prototyping of new approaches, and the resource boxes at the end of this report offer links to resources that review specific types of assessment for strengths, limitations, and suitability for particular purposes. But, as even the primers make clear, the field of next generation assessment is still emerging. Careful thought is warranted about the use of various assessment measures, especially for agency and Habits of Success; nevertheless, given the demonstrated importance of these broader competencies, it seems equally important to avoid “analysis paralysis” and to forge ahead in testing out new tools and measures in next generation learning and assessment (though not accountability) models. As Angela Duckworth and David Yeager urge at the end of [Measurement Matters](#), their paper on assessing “personal qualities”:

Given the advantages, limitations, and medium-term potential of such measures, our hope is that the broader educational community proceeds forward with both alacrity and caution, and with equal parts optimism and humility.⁹

Cautions

Among other warnings, Duckworth and Yeager’s influential article provides a useful set of cautions from people involved in researching and developing measures for the trait of *grit*. They point out that many of the existing Habits of Success and agency measures were created for research purposes, not as evaluation tools or tools to measure student outcomes. They review three of the most common types of measures: self-reports, teacher-reports, and performance assessments, describing their strengths (all can be “remarkably predictive” under the right conditions) and their limitations. With regard to limitations, for example, self- and teacher-reports are subject to misinterpretation by the participants, reference bias (with different groups of people having different standards or reference points when they rate themselves), faking, and social desirability bias (answering in a way that others will view favorably). While performance assessments avoid some of these concerns, they can be misinterpreted by researchers and be vulnerable to task impurity and extraneous situational influences. The authors suggest that educators “seek out the most valid measure for their intended purpose(s),” adding that “[w]henver possible, we recommend using a plurality of measurement approaches. While time and money are never as ample as would be ideal, a multi-method approach to measurement can dramatically increase reliability and validity.”¹⁰

Promising developments

Duckworth and Yeager, as well as numerous other practice researchers, also underline the potential for improving existing assessments, as well as for innovations that produce new types of assessment. Potential improvements to existing assessment types include: enhancing self-reports (such as by using anchoring vignettes to decrease reference bias by helping users calibrate responses; sampling random subsets to decrease assessment impact on individual students, and using quasi-ipsative pairwise-preference options to improve forced choice measures); mining behavioral or bio data (analyzing online learning behaviors and

communications in real time, and applying factor analysis to data such as attendance, GPA, and suspensions); and improving performance assessment (using more sophisticated tools and processes such as learning progressions, social moderation of assessment, and the embedding of performance tasks online as a way of facilitating information collection and performance reporting). For more on addressing the limitations for performance assessment, see the ETS paper, [Psychometric Consideration for Performance Assessment with Implications for Policy and Practice](#).

Ways the field is working together

Educators, researchers, psychometricians, and others are collaborating in a variety of ways to develop new measures and assessment approaches to improve educator practice and help students progress over time in developing agency and capability.

One notable approach here is the creation of communities of learning and practice. The [Student Agency Improvement Community](#), (SAIC), for example, led by Anthony Bryk and David Yeagar for the Carnegie Foundation, is a networked improvement community that brings together academic research on psycho-social learning with improvement science practices to help teachers advance the development and measurement of student agency.

The SAIC is, for example, creating a *practical* measure to help improve classroom practices that build student agency. The practicality is achieved because the measure is short, and it is easy for practitioners to embed it within their daily work routines; however, it also employs items that are “demonstrated to be powerfully predictive of important educational outcomes. Its development has been guided by theory and linked to specific work processes and change ideas” introduced in the improvement community. One SAIC member, Summit Public Schools, has provided [thoughtful insight on its activities](#) related to student agency and its multiple approaches to measurement.

Similarly, the Center for Innovation in Education and Next Generation Learning Challenges launched the [Assessment for Learning Project](#) (ALP) to support and connect educators who are “fundamentally rethinking the core role that strategies and systems of assessment should play to advance student learning.” ALP has made grants to 17 teams of educators nationally who are fundamentally rethinking the role(s) of assessment in learning and system design. These projects are piloting assessment innovations designed to cultivate student agency, activate broader definitions of student success, and improve the cultural responsiveness of assessment strategies. ALP includes projects led by classroom teachers, school districts, professional learning non-profits, and state policy leaders. By bringing these perspectives together into an active community of practice, ALP is learning how assessment FOR learning can be supported from the classroom to state education agency.

Several other collaborative developments are also underway. The [Assessment Work Group](#) of the Collaborative for Academic, Social, and Emotional Learning (CASEL), in collaboration with California’s CORE Districts, Transforming Education, RAND, and Harvard University, has launched the [Measuring SEL website](#) and a growing [Collaborators’ Network](#). The group is developing [a practical guide/online tool](#)

for selecting and using currently available and practical social-emotional learning (SEL) competency measures that will be released in January 2018.

The [Center for Curriculum Redesign’s Assessment Research Consortium](#) (ARC) bills itself as “a collaborative entity — a ‘pre-competitive Research & Development consortium’ modeled after industry’s similar endeavors in semiconductors, biotechnology, etc.” The ARC’s goal is to “redesign systems of measuring Learners’ progress, aligned to 21st century competencies and desired education outcomes.” The organization published two landmark reports in 2016 — [Evolving Assessments for the 21st Century](#), and [Workforce Assessments: What Do We Actually Measure](#) — both of which do an excellent job of curating research, publications, and assessment tools. The [California Performance Assessment Collaborative](#) (CPAC), supported by the Learning Policy Institute, collaborates on performance assessment development with California districts and schools, several of the Deeper Learning school networks, the Stanford Center for Assessment, Learning, and Equality (SCALE), and others. Its [newsletter](#) makes some of its resources available to all.

In addition to these efforts and high-profile work from the likes of Angela Duckworth’s [Character Lab](#) (grit) and Carole Dweck’s [Mindset Works](#) (growth mindset), a range of assessments are being developed by educational research organizations, collaborations between next generation educational models and psychometric organizations, and other partnerships. Researchers at the University of Chicago Consortium on School Research, whose *Teaching Adolescents to Become Learners* and *Foundations for Young Adult Success* inform much of the work in this area, have completed two pilots of their [Becoming Effective Learners Survey](#), which aims to consolidate existing survey scales to create a comprehensive measurement instrument and simultaneously measure student agency factors and classroom context/instructional factors. [Transforming Education](#) is working with the CA CORE on measures for their MESH (Mindsets, Essential Skills, & Habits) framework. [MHA \(Means & Measures of Human Achievement\) Labs](#), an R&D nonprofit with a network engagement approach, is prototyping, testing, and validating 21st century skills and workforce readiness performance assessments based on performance reviews completed by employers and mentors/instructors for summer jobs programs over the past five years. Next Generation Learning Challenges grantees [Valor Collegiate Academies](#) and [Generation Schools Network](#) are partnering, respectively, with [6seconds](#), an emotional intelligence (EQ) training and assessment company, and [Pairin](#), an HR solutions company, to develop SEL measures. Johns Hopkins’ [Talent Development Secondary](#) is developing diagnostics based on its highly regarded agency program elements that link agency skill levels, as determined by self- and teacher-reports, to the impacts of poverty.

It is also worth reiterating the potential for adapting and learning from good practice outside of academic assessment. Throughout the primers, there are references to the potential of learning from practitioners in a range of sectors that have long focused on whole person development, including: *coaches* of novice-to-expert practice-based pursuits like music, drama, and sports; *youth developers* and out-of-school-time providers, who have long focused on agency and Habits of Success; *special educators*, who work with strategy-based instruction and assessment and interventions for executive functioning, ADHD, and autism spectrum disorders; *early educators and developmental psychologists*, who have established instruments

and approaches for child development, agency, and habit-related traits; and *workplace human resource staff*, who have long used psychometric testing in employee hiring and development. Educators have as yet only scratched the surface of collaboration with these experts on both learning and assessment, and the MyWays team believes there is a lot of potential still to be explored.

The state of assessment across the MyWays domains

As we illustrate in the previous sections, the effect of the two key paradigm shifts in assessment is a re-balancing — away from a heavy focus on traditional, single-subject-based, inauthentic, summative testing and toward greater emphasis on the continuous, embedded, authentic assessment that is naturally integrated within formative and performance aspects of Whole Learning. This directional shift is bolstered by the use of multiple measures, badges and micro-credentials (especially for learning increasingly happening in the Wider Learning Ecosystem), and the evaluation of new kinds of learning experiences (especially for parts of the domains where measures for student outcomes are still evolving).

It is important to stress that the use of all Five Assessment Strategies extends across all MyWays domains and competencies. What will not work, we are certain, is to try to isolate each of the 20 competencies and create a (likely inauthentic) way to assess each one. Indeed, educators would not even want to take each of the MyWays domains and choose a different assessment strategy for each.

One of the critical design features of next generation learning and assessment is that it maximizes the holistic, authentic nature of the learner’s experience, while being built on detailed, behind the scenes design work that ensures coverage, depth, rigor, and developmentally appropriate opportunities for advancement. From the learner’s viewpoint, next generation assessment shares the intentional, meaningful, and holistic nature of its partner, Whole Learning: each learning experience or junior version is designed to address a manageable number of competencies, including the associated aspects of agency and capability, while assessment will focus on that particular mix of knowledge, skills, habits, and personal wayfinding in as holistic and authentic a way as possible (suggesting, in many cases, some type of performance assessment). That performance assessment will provide information on the learner’s progress in most of the chosen competencies, while other assessment components will be added as needed for competencies (or, for instance, elements of agency) that may require a different approach.

All that being said, from the viewpoint of those charged with designing learning and the educational environment it occurs in, the suitability of assessment approaches, relevant cautions, and promising developments vary somewhat across the four competency domains. So, for next generation educators who are looking to understand the overall assessment landscape, as well as those who have already identified assessment gaps for particular competency domains, the following quick review focuses on considerations particularly relevant to each domain. Please note that these are just some summary points to initiate discussion. The relevant sections below are also included in the four domain reports (Reports 7–10), where you can see them as part of the full domain discussion, which covers why the domain is important, competency definitions, and key principles for implementation.

Considerations for assessing Habits of Success



- Academic Behaviors
- Self-Direction & Perseverance
- Positive Mindsets
- Learning Strategies
- Social Skills & Responsibility

Summary: The growing interest in measuring Habits of Success is partly due to the growing belief in the importance of a broader skillset for all students and partly due to the addition of a non-academic element to ESSA accountability. Effective measurement is

still emerging, however, and we need to be cautious about what kinds of metrics to use and how to share them with learners. In particular, opinion about the use of this domain in accountability systems is sharply divided, and caution is urged.

The range of assessment approaches for Habits of Success includes the following:

- Three main approaches:
 - **Early Warning Indicators (EWI)** to monitor critical academic behaviors (Johns Hopkins’ [EWI](#), Stanford and Chicago’s [CRIS](#) – College Readiness Indicator System)
 - **Rubrics, reflections, and peer assessments within performance assessment** from [New Tech Network](#), [Summit Public Schools](#), The Center for Innovation in Education (CIE) and Educational Policy Improvement Center (EPIC)’s [Essential Skills and Dispositions Developmental Frameworks](#). For more on these and other tools on self-direction and social skills, see the “Practice Resources for the 4Cs and More” box near the end of the Creative Know How domain report (Report 8).
 - **Various agency and SEL assessments** from different fields, including Likert and other self-reports (such as the [Holistic Student Assessment](#) for resiliency available from the Harvard University/McLean Hospital PEAR Institute), behavioral observations, ratings by others, situational judgment, climate surveys, and other measures. Ideally, such measures would be undertaken within an enhanced guidance/youth development function by teams that include people with psychology or social work backgrounds.
- Other assessments:
 - **Simulation and game-embedded assessments** ([Newton’s Playground](#), which measures conscientiousness and persistence as well as the learning of physics concepts, [ZooU](#))
 - **Clickstream analysis of learning behaviors** (See, for example, the mention of clickstream analysis in this Summit [blog](#).)
 - **New instruments in development** (Subscribe to the [Transforming Education](#) MESH e-newsletter to track the latest developments.)

Ongoing challenges in assessing Habits of Success include inconsistent attribute definitions and “fakeability” (socially desirable responding); the need for care and nuance in sharing assessment results with learners; and potential for misuse in accountability (on the last, see Key Principle 3 in Report 7). For more on Habits of Success assessments, see the Habits of Success one-page competency primers at the end of Report 7 and two recent external reports: the Center for Curriculum Redesign’s [Evolving Assessments for a 21st Century Education](#) and the National Academies Division on Behavioral and Social Sciences and Education’s [Supporting Students’ College Success: The Role of Assessing Intrapersonal and Interpersonal Competencies](#).

Considerations for assessing Creative Know How



- Critical Thinking & Problem Solving
- Creativity & Entrepreneurship
- Communication & Collaboration
- Information, Media, & Technology Skills
- Practical Life Skills

Summary: Evidence collection through performance assessment (curriculum-embedded and bounded) and other approaches are maturing; however, issues relating to transfer and reliability are complex and still being addressed. The growing use of project-based/performance

assessment for Content Knowledge provides growing opportunity to collect evidence on the development of Creative Know How skills as part of the same processes, performances, and outputs.

The current range of assessment approaches includes the following:

- **Predominant reliance on rich, curriculum-embedded performance assessments (PA)**, including the use of the following tools:
 - Validated performance frameworks and/or tasks created by educators and networks or pulled from rich task databases (SCALE/SCOPE/CCSSO’s [Performance Assessment Resource Bank](#) has a few Creative Know How PA items; EdLeader21 has [announced](#) a forthcoming 4Cs Performance Assessment Bank)
 - Rubrics, learning progressions, completed exemplars, portfolio and exhibition protocols, and other tools (see more information and links in the Practice Resources box in Report 8 to New Tech Network’s student learning outcomes and rubrics; Summit’s discipline-specific skills rubrics and look-fors; P21’s rubrics; EdLeader21’s rubrics for the 4Cs; and the Buck Institute’s rubrics for assessment of 4Cs).
- **Innovative approaches, such as embedded assessments**, which can capture student processes in competencies ranging from problem solving to creativity (see, for example, the mention of clickstream analysis of ways of working in this Summit [blog](#) on assessment for the four components of their model, and this article on [Newton’s Playground](#), which measures creativity and conscientiousness as well as the learning of physics concepts through game-playing).
- **The use of multiple, varied measures**, such as the combination of self-report, situational judgment, and forced-choice methods offered within ProExam’s [Tessera Noncognitive Assessment System](#).
- **Bounded, on-demand performance tasks**, such as those including, most recently, collaborative problem solving, in the Program for International Student Assessment ([PISA](#)), [NextGen Science](#) assessments, the Common Core [PARCC](#) & [Smarter Balanced](#) assessments, and the Council for Aid to Education’s [CWRA+](#), a middle and high school level assessment that uses document-based real-world problem solving tasks to measure critical thinking skills.

Ongoing challenges in the assessment of Creative Know How include issues of **transfer** (such as which learning approaches improve the likelihood of transfer, especially across subject areas and widely varying circumstances), **educator capacity-building** in the assessment of Creative Know How skills, and the **development of performance- and portfolio-based digital platforms** that are flexible enough to house and track the kind of evidence required for Creative Know How skills, including evidence of processes and student reflection, as well as output and performance. (See examples in Key Principle 3 in Report 8.)

For more on Creative Know How assessments, see the Creative Know How one-page competency primers at the end of Report 8 and two recent external publications: the Center for Curriculum Redesign’s [Evolving Assessments for a 21st Century Education](#) and the National Academies Division on Behavioral and Social Sciences and Education’s [Supporting Students’ College Success: The Role of Assessing Intrapersonal and Interpersonal Competencies](#)

Considerations for assessing Content Knowledge



- English Core
- Math Core
- Science, Social Studies, Arts, Languages
- Interdisciplinary & Global Knowledge
- Career-Related Technical Skills

Summary: Content Knowledge is intensely measured or over-measured, often in compartmentalized ways; however, there are encouraging moves away from memory-based testing and toward higher level thinking, application of concepts, and more authentic

performance assessments.

The range of assessment approaches includes the following:

- **Traditional assessment**, including teacher-designed tests, quizzes, essays, research papers, and labs.
- **Increasing attention to formative assessment**, in the form of more frequent, granular checks for understanding and the provision of real-time, high-quality feedback (or, a definition we like, “using [insights into] student thinking as a basis for teaching and learning”¹¹).
- **Performance assessments for formative and summative purposes**, particularly in the form of more authentic curriculum-embedded performance assessments, including in project-based or other experiential learning.
- **Increasing use of tech-enabled assessment to contribute to diagnostic, formative, and individualized mastery purposes, including:**
 - Diagnostic and formative assessments: [MAP](#), or computer adaptive assessments, in math and reading; [OECD test for schools](#); tech-enabled, quick feedback assessment, such as [Poll Everywhere](#), [Google Forms](#), [Gooru](#), [FlexiQuiz](#); and other ways to help check for understanding.
 - Approaches amenable to student-managed and adaptive courseware (such as Summit Public Schools’ playlists).
- **Moves toward knowledge application and performance assessment within accountability measures include the following:**
 - Mixing application of knowledge with on-demand tasks or bounded performance assessment (Common Core [PARCC/SBAC](#) assessments).
 - Pilots with curriculum-embedded performance assessment for accountability ([New Hampshire’s PACE](#) state waiver, the [New York Performance Assessment Consortium](#), [California Performance Assessment Collaborative](#)).

Ongoing challenges in Content Knowledge assessment include ensuring that assessment for learning is prioritized over assessment for accountability, and that assessments are focused on key organizing ideas and higher-level thinking. For performance assessment, the challenge is to ensure that attention is given to

building educator capacity (which includes having educators organize thoughtful calibration and social mediation), and enabling students to separately collect evidence on progress related to Content Knowledge, Creative Know How, and Habits of Success.

For more on Content Knowledge assessment, see the Content Knowledge competency primers at the end of Report 9 and two recent external publications — the Center for Curriculum Redesign’s [Evolving Assessments for a 21st Century Education](#) and the National Academies Division on Behavioral and Social Sciences and Education’s [Supporting Students’ College Success: The Role of Assessing Intrapersonal and Interpersonal Competencies](#).

Considerations for assessing Wayfinding Abilities



- Survey the Learn, Work, & Life Landscapes
- Identify Opportunities & Set Goals
- Design & Iterate Prototype Experiences
- Find Needed Help & Resources
- Navigate Each Stage of the Journey

Summary: Wayfinding competencies are often unaddressed; even when they are addressed, they go largely unmeasured. To deal with this, Wayfinding innovators are turning to badging (often internally) and/or tracking and evaluating student learning experiences to help ensure that

students have opportunities that are likely to help develop agency and self-direction.

Next generation suggestions for Wayfinding assessment include:

- Start by tracking the extent to which you are providing learners with the needed experiences (for more, see the “Visions of the Possible” box in Report 10).
- Where outcomes cannot be measured, step back in the evaluation logic model from looking at outcomes to thinking deeply about inputs. Implement quality reviews of educator capacity and Wayfinding curriculum, as well as quality standards, criteria, or guidelines for service learning ([National Youth Leadership Council](#)), career awareness experiences ([Linked Learning](#)), or internships ([National Academy Foundation](#)).
- Use standards-driven badging (such as [Boston Afterschool and Beyond](#) digital badges or [Open Badges](#)) where applicable, and use participation-driven badging for more informal learning.
- Create school-based badging systems and/or incorporate external micro-credentials that include not just metrics for participation but also student preparation for and reflection on their experiences. For example, you might award students a badge not for simply visiting three colleges, but for reflecting on how each of the colleges matched up to (or altered) the four factors they had chosen ahead of time to look for in a college.

Ongoing challenges in the assessment of Wayfinding Abilities include: carving out time in schedules to include and assess Wayfinding competencies; providing the types of student experience around Wayfinding that truly enable student agency; and creating the necessary relationships to fully involve external partners, such as families, community organizations, employers, and postsecondary institutions.

For more on Wayfinding Abilities assessments, see the Wayfinding Abilities one-page competency primers at the end of Report 10 and two recent reports: the Center for Curriculum Redesign’s [Evolving Assessments for a 21st Century Education](#) and the National Academies Division on Behavioral and Social

Sciences and Education's [*Supporting Students' College Success: The Role of Assessing Intrapersonal and Interpersonal Competencies*](#).

A quick dive into broader, deeper competency assessment resources

Because the purpose of the MyWays Student Success Framework is to provide a rosetta stone for thinking about the richer, future-ready success definition for today's learners, we have focused on describing that definition in conceptual terms. We also believe deeply that school designers, educators, and individual learners need to invest in constructing and evolving their own goal-lines within the broader framework.

In doing this work, educators may find the resources on the following pages helpful:

Starter Resources for Assessment Design for Broader, Deeper Competencies

This list offers some resources to get you started. A separate box below offers the best links for databases or compendia of assessment tools. The resources here may also include some tool information, but the focus is on introducing overall assessment approaches and reviews of the field.

General resources:

Starter blogs:

- Andrew Miller, "[Using Assessment to Create Student-Centered Learning](#)," Edutopia blog, Sept. 2, 2015.
- David Wees, "[Formative Assessment: More than just an exit ticket](#)," The Reflective Educator blog (undated).

Reviews of established and new measures by educators and youth developers:

- CCR Assessment Research Consortium, [Evolving Assessments for a 21st Century Education](#).
- Asia Society and RAND, [Measuring 21st Century Competencies: Guidance for Educators](#).

Resources on performance assessment (including embedded formative assessment):

- McREL/Measured Progress, [Re-Balancing Assessment: Placing formative and PA at the heart of learning and accountability](#).
- Center for Collaborative Education, [Quality Performance Assessment: A Guide for Schools and Districts](#); also see CCE's tool-packed, open-source [Performance Assessment for Learning Micro-credentials](#).
- Ron Berger et al., [Leaders of their Own Learning: Transforming Schools Through Student-Engaged Assessment](#) book and [website](#).
- Bob Lenz, [Transforming Schools: Using Project-based Learning, Performance Assessment, and Common Core Standards](#) book and [website](#).
- The Buck Institute for Education's "[Resource List: Assessment in PBL](#)."

Resources on multiple measures for Habits of Success and Creative Know How:

- [From Soft Skills to Hard Data: Measuring Youth Program Outcomes](#), Forum for Youth Investment.
- For a shorter survey, see Patrick Kyllonen, ETS, [Soft Skills for the Workplace](#).
- For cautions on using existing measures, see Angela Duckworth and David Yeager, "[Measurement Matters: Assessing Personal Qualities Other than Cognitive Ability for Educational Purposes](#)."
- To keep up to date with new developments, subscribe to the [Transforming Education](#) MESH newsletter.

Resources on badges and micro-credentials and quality reviews

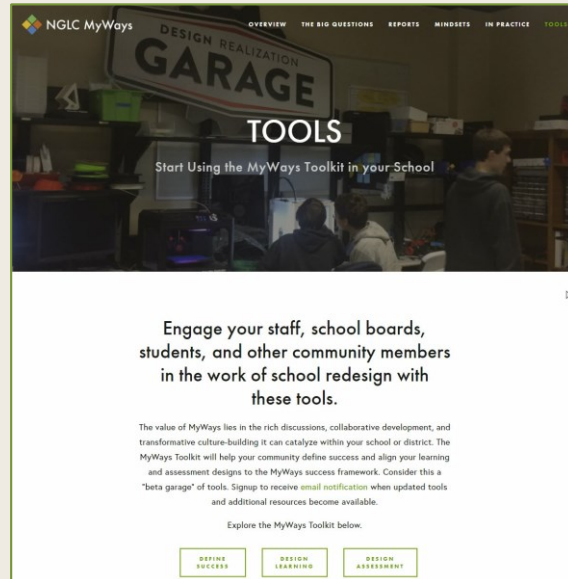
- [Expanding Education and Workforce Opportunities through Digital Badges](#), Alliance for Excellent Education.
- "[Assessing Learning in Digital Badge Systems](#)," in *Ten Principles for Assessing Learning that Apply to Badging and More*, Design Principles Documentation Project.
- [Understanding CTE and Stackable Credentials](#), Transitions Academy.
- Srikanth Gopal, "[Stop \(Just\) Measuring Impact, Start Evaluating](#)," FSG Consulting blog.
- [District Guide to Blended Learning Measurement](#). A step-by-step guide to when, what, whom, and how to measure from The Learning Accelerator.

MyWays Tools and the Mayan Worked Case Study

The MyWays Toolkit, available on the [Tools page](#) of the MyWays website, includes simple matrix tools to help you reflect and promote discussion on the MyWays concepts, and evaluate your existing assessment practices or design new ones.

The Toolkit also includes a worked case study that uses a selection of the MyWays evaluation tools to analyze a High Tech High middle school project on Mayan culture. The *Mayan worked case study* serves as:

- one example of the kind of Whole Learning and authentic assessment design required to address the competencies needed for a world of acceleration, and
- a way of illustrating how you might use a few of the MyWays diagnostic/design tools with your educator and designer teams, and to build the will to transform among your fellow stakeholders.

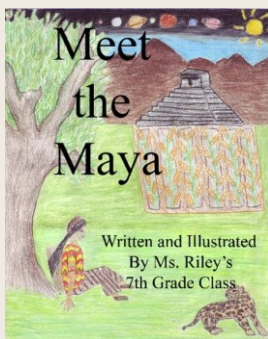


Two assessment tools from the MyWays Toolkit

We highlight here two of the MyWays assessment-related tools. Both are designed to help you use MyWays and the Five Assessment Strategies to address the fourth big question in the MyWays Through-line: *HOW do we gauge students' progress in developing richer competencies?* **The first is the Assessment-Competency Correlation Tool.** This tool helps you ask: How well is my school employing the five strategies that support assessment for broader and deeper competencies? **The second is the Assessment Strategy Analysis Tool.** This tool helps you ask: How well is my school using elements of the five strategies to assess student progress within our learning projects? These two tools are matrices to help you evaluate and improve assessment experiences. Simple, easy-to-follow instructions for using the tools are provided in the toolkit.

The goal is to equip your assessment design team with a reliable process for critiquing emerging assessment approaches — strengthening the extent to which you integrate your assessment with your learning, moving the team toward more authentic tasks, and increasing the range of assessment approaches so that you can address broader and deeper competencies. Even at a quick, conceptual level, these tools can flag key issues and “help change the conversation” within your team with respect to transforming assessment as a force for teachers to better know and guide their students, as well as for learners to understand themselves more fully.

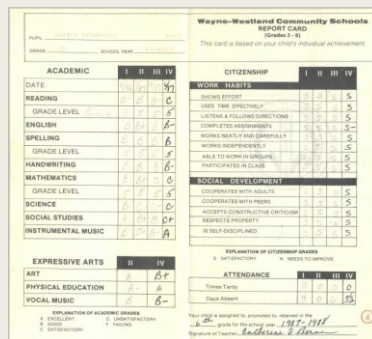
Using the tools: High Tech High's Mayan Community Project



In Report 11 we introduced the Mayan Community Project, an experiential learning project developed by Heather (Riley) Lovell, a seventh grade teacher at a High Tech High middle school. We chose the project because Lovell and High Tech High offer an abundance of materials on the [High Tech High Project website](#) that let us share with you many different aspects of the experience that are relevant to the MyWays competencies and authentic assessment. And of course because the project provides an inspiring example of Whole Learning and authentic assessment! In Report 11, the Mayan project was analyzed using three of the learning-related tools. Here we extend that case to demonstrate how you can use the concepts we present in this report to align your assessment design with the two key assessment shifts and Five Assessment Strategies.

Potential tool for assessment reporting

As assessment becomes more authentic and multi-measured, we must also recognize the student-facing side of assessment reporting. For centuries, the simple report card has been the dashboard of student achievement and progress. As we move to broader, deeper competencies and more complex assessment systems, we must address a key question: *How will our reporting to students, parents, colleges, and employers add needed depth and meaning in readily understandable ways?*



We have received positive, early responses to the MyWays Whole-Student Competency Plot — an idea we borrowed from the Lumina Foundation’s Degree Qualification Profile. To enable educators to use the plot, we created two simple, publicly available tools in Excel that are easily customized and can be used to compare any two states, such as *student A vs. B*, *school A vs. B*, or, as in the example that follows, *today vs. future*. Two versions of **the MyWays Whole-Student Competency Plot Tool** are available on the [Tools page](#) of the MyWays website, along with a more detailed explanation. Currently, the tool is conceptual, but some members of the Next Generation Learning Challenges network are interested in using learning progressions, rubrics, and scoring protocols to make it empirically driven. (See, for example, this [EdSurge article](#) on Two Rivers Public Charter School’s assessment work.) For guidance on emerging tools of this sort, see the next resource box.

Tia’s competency profile today (black) and goals at graduation (purple)

This plot illustrates a shortcoming that plays out almost every day in schools using the traditional, narrow academic metrics for student success. Tia (a fictionalized composite of two real students in the Boston area) is a complex learner whose natural gifts and competencies in Habits of Success and Creative Know How are neither recognized nor developed. Using a visual tool with a broader and deeper set of competencies enables Tia and her advocates to visualize her strengths and set goals for her future success.



Guidance on Finding Tools for Performance Assessment, and Multiple Measures for Habits of Success, Creative Know How

This box provides guidance on sets of assessment tools used in the assessment strategies that have been of greatest interest to the MyWays Community of Practice. These resources particularly focus on links, databases, or compendia for 1) the tools related to performance assessment (such as learning progressions, rubrics, skills maps, and performance tasks. Note that most of these can also be used for formative assessment); and 2) multiple, varied, and emerging assessment measures used to assess competencies in Habits of Success, Creative Know How, and a few of the Wayfinding Abilities.

We know from our beta piloting work with next generation educators that those interested in and inspired by the MyWays Student Success Framework are also thirsty for practitioner tools, exemplars, and documentation. In some cases, practitioners may be tempted to latch onto tools and use them without the internal mindset-changing and learning-model-revising work required for successful implementation; we caution against this! We also realize, however, that many thoughtful developers and practitioners want and need to see more concrete exemplars and tools to better understand the broader, deeper goal-line; to help work through their own approach; and to help plan and implement their assessment activities — which is why we provide these links. Note that MyWays and Next Generation Learning Challenges do not endorse any specific tools for assessment or curriculum planning — and particularly urge practitioners to ensure that tools they use are authentic, holistic learning.

Performance assessment tools

- **The Institute for the Future of Learning’s open source tool repository**

As part of the Institute for the Future of Learning project (which produced the excellent report, [Assessing the Learning that Matters Most](#)), Julie Wilson created a database of learning progressions, rubrics, and tasks on the 4Cs and on self-assessment and social-emotional learning. The tools were provided by EL Education, New Tech Network, High Tech High, Mount Vernon, Two Rivers Public Charter School, Sanborn Regional School District in New Hampshire, Catalina Foothills, Science Research Academy, and KIPP Socratic Seminar — more than 75 documents in all. The tools are searchable by topic, school model, and grade level, and can be found on this [beta website](#).

- **EdLeader21’s 4C’s rubrics**

This is a nationally vetted set of [rubrics](#) for the 4Cs from [EdLeader21](#). The master set of 4Cs rubrics covers grades 3–4, 7–8, and 11–12 can be purchased from EdLeader21, but you can see adapted versions in links from this [blog](#) by Ken Kay, EdLeader21’s CEO, who noted that, “The rubrics are a great resource on their own, but you and your teachers can also adapt them to your needs. For example, some of our districts have [modified the rubrics](#) and [associated learning targets](#) to make them student-friendly.”

- **The Buck Institute rubrics for assessing the 4Cs in a PBL context**

These [rubrics](#) describe what 4Cs good practice looks like, specifically in the project-based learning (PBL) context, with different sequenced rubrics for K–2, 3–5, and 6–12. Critical Thinking and the “Process” section of Creativity & Innovation are organized by the four phases of a typical project. The Presentation Rubric is used only in a project’s last phase, when students share their work with a public audience. Collaboration is relevant to all phases. See this [blog](#) for more on how to use these rubrics.

- **P21 21st Century Skills Maps**

These [21st Century Skills Maps](#) address how to implement learning models that integrate the 4Cs into core academic content mastery. 4Cs skills maps are available for math, science, social studies, geography, English, languages, and arts; ICT skills maps are available for social studies, English, and math. Each skills map provides examples of the types of skills that are appropriate for 4th, 8th, and 12th grade levels.

- EdLeader21’s 4Cs Performance Assessment Bank (forthcoming)**
 According to the [EdLeader website](#), “For EdLeader21 members, the lack of easily accessible, easily deployable assessment tools focused on the 4Cs is an ongoing issue. Most of our members have expressed interest in the development of assessment instruments that will help measure 4Cs student performance. The 4Cs Performance Assessment Bank project aims to establish a 4Cs-aligned bank of performance tasks that can be customized locally; EdLeader21 members may develop their own tasks for inclusion in this bank (using common design guidelines). The focus will be on... formative instruction and assessment.”
- The Center for Innovation in Education (CIE) and Educational Policy Improvement Center (EPIC)’s Essential Skills and Dispositions Developmental Frameworks**
 This [set of developmental frameworks](#) covers collaboration, communication, creativity, and self-direction in learning. The frameworks define five components inherent to each skill and describe performance for each component across a beginner to emerging expert progression, informed by research on the development of expertise. Unlike discipline-specific learning progressions and rubrics, the developmental progressions reflect components essential to the skill itself and describe growth dependent on many years of active exploration, experimentation, setbacks, and reflection.
- New Tech Network’s learning outcomes, rubrics, and college-ready assessments**
 New Tech Network (NTN), working with Envision and the Stanford Center for Assessment, Learning, and Equity (SCALE), created [open-source learning outcomes and rubrics](#) related to: knowledge and thinking in different core subject areas; agency; collaboration; and oral and written communication. These tools are used in NTN’s curriculum-embedded performance assessments called [College Readiness Assessments](#). The network also offers a three-part Student Literacy [video series](#) (10–15m each) that guides users through the delivery of workshops focused on the creation of high-quality tasks, looking at student work, and the use of the knowledge and thinking rubrics (including the difference between grading and scoring).
- Two Rivers Public Charter School’s resources**
 Two Rivers Public Charter School, a high-performing [EL Education School](#), hosts its own excellent professional sharing site, [Learn with Two Rivers](#). Its tasks and rubrics that address critical thinking, problem solving/“expert thinking,” collaboration, and communication are currently being curated in this separate [Deeper Learning Assessment folder](#). For a public share of excellent resources on working with Habits of Success (Valor’s Compass program) by this thoughtful MyWays Community of Practice member, see links to over a dozen resources in their [“Working the Compass” Resource Guide, Summer 2017](#).

Multiple measures assessment tools for Habits of Success, Creative Know How

- CASEL/Measuring SEL’s [Assessment Guide for Educators](#), an online tool for selecting and using currently available SEL measures is forthcoming in winter 2017.
- Center for Curriculum Redesign’s Assessment Research Consortium (ARC), [Evolving Assessments for a 21st Century Education](#). Recent report that does an excellent job of curating assessment tool information, as well as related research and publications.
- Asia Society and RAND, [Measuring 21st Century Competencies: Guidance for Educators](#).
- Forum for Youth Investment, [From Soft Skills to Hard Data: Measuring Youth Program Outcomes](#).
- ETS, [Soft Skills for the Workplace](#). A shorter, older survey with some useful explanations.
- For nine more compendia, see the selected assessment list on page 9 of CCR ARC’s [Evolving Assessments for a 21st Century Education](#).

Series End

This is the final report in the *MyWays Student Success Series: What Students Need to Thrive in a World of Change*.

For further resources please see the Next Generation Learning Challenges [MyWays website](#). To meet the Community of Practice members and share your ideas see our [Community of Practice page](#), and to receive updates on MyWays join our [mailing list](#).

Report 12 Endnotes

¹ From a Statement by the Global Education Leaders' Program (GELP), "Transforming Global Education with New Metrics," issued in June 2014. This statement is no longer available on the GELP site, but see the [New Metrics](#) page for links to subsequent reports on assessment from GELP.

² Andrew Hargadon, *How Breakthroughs Happen: The Surprising Truth about How Companies Innovate*, Harvard Business School Press, 2003 (especially Chapter 2).

³ Asia Society and RAND, the Gordon Commission, Big Picture Learning, EL Education, and Digital Learning Badges, among many others linked and cited throughout this report.

⁴ Michael Fullan and Maria Langworthy, *A Rich Seam: How New Pedagogies Find Deep Learning*, Pearson, 2014.

⁵ Robert Halpern, *Youth, Education, and the Role of Society*, Harvard Education Press, 2013.

⁶ Andrew Miller, "[Using Assessment to Create Student-Centered Learning](#)," blog, Edutopia, September 2, 2015.

⁷ RAND Corporation and the Asia Society, *Measuring 21st Century Competencies: Guidance for Educators*, 2013, pp 32–37.

⁸ Alec Barron, "[Why do we use digital badges?](#)," Del Lago Academy's Competency X blog, 2/6/2017.

⁹ Angela Duckworth and David Yeager, "[Measurement Matters: Assessing Personal Qualities Other Than Cognitive Ability for Educational Purposes](#)," *Educational Researcher*, vol. 44, no. 4, 2015, p. 246.

¹⁰ Ibid.

¹¹ "Formative assessment is more of a mindset on using student thinking as the basis for teaching and learning rather than a quick checklist or a list of strategies." From David Wees, "[Formative Assessment: More than just an exit ticket](#)," blog, *The Reflective Educator* (undated).