

An Effective Comprehensive Assessment System:

The Case for Adding Cognitive Assessment

By Nancy Weinstein



Summary

Widespread reports of declining academic performance following the COVID-19 pandemic have ushered in a potentially new cycle in educational assessment. There is a strong desire to understand which students are behind, by how much, and how to accelerate student growth to make up for missed learning during school closures and the bumpy return to business-as-usual.

While assessments are a valuable tool to understand and support learning, there is no single, universal assessment that can meet the broad and diverse needs of teachers, school leaders, district administrators, parents, and students. Rather, schools benefit from a well-designed comprehensive assessment system that balances the importance of providing essential data to key stakeholders while not spending unnecessary time and resources on testing. Assessments must also extend beyond simply reporting students' current mastery level to guide teachers to offer differentiated instruction based on the assessment results.

This paper offers a very brief, recent history of school standardized assessments, outlines the components of a comprehensive assessment system, and **proposes a solution for a more integrated, effective, and equitable comprehensive assessment system using cognitive assessment**. Cognitive assessments offer the unique capacity to objectively and efficiently pinpoint the reasons for variability in performance across traditional assessments, thus making data interpretation straightforward for key stakeholders. Cognitive assessments offer the additional benefits of alignment to instructional strategies that extend beyond a reteaching of curriculum and instead focus on modality and learning approach, providing teachers with actionable next steps to personalize instruction based on the underlying reasons for learning gaps. Perhaps most importantly, cognitive assessments provide an essential check on traditional language-based assessments. At a time when the educational community searches for solutions to address implicit bias, cognitive assessments offer the opportunity to see student strengths beyond literacy skills that disproportionately hinder students from lower socio-economic backgrounds and they mitigate the risk of teachers underestimating students based on their zip code or appearance. Included in this report are concrete examples of how cognitive data can provide key insights to understand and support performance differences between classroom and standardized tests, variability in student performance across subjects, and strengthen student engagement.

Brief Recent History of Standardized Assessment

Standardized testing of academic achievement (reading, writing, math) has been the norm in US public schools since the early 1900s (retrieved from the [NEA website](#), September 30 2022). However, public policy debate about the purpose and application of standardized test results has led to cycles of underuse, overuse and controversy. The current era of debate was ushered in by the passage of No Child Left Behind (NCLB) in 2001 which was the beginning of standards-based reform movement and the widespread use of summative assessments to universally measure students' academic knowledge at year-end (Perie et. al). The primary purpose of NCLB assessment was accountability, i.e. objectively evaluating school performance, rewarding high performing schools, highlighting schools for intervention, and

providing transparency on school performance for all stakeholders. The new level of transparency and the direct linkage of federal funding to summative assessment scores turned what was once a mere necessary annual school ritual into a high stakes annual event.

Backlash from teachers and parents was significant. Concerns included the inconsistency of state measurements, fairness of tying funding to a universal standard without a recognition of differences in communities, and the emphasis on “teaching to the test” over deeper learning. That backlash led to the eventual adoption of the Common Core Standards in 2012, an originally state-led initiative intended to provide universal educational benchmarks that would focus on developing critical-thinking skills over rote learning.

For a variety of well-reported reasons, including renewed backlash from teachers, parents, and some state leaders, Common Core was abandoned in most states and the federal Every Student Succeeds Act (ESSA) was adopted in 2015. ESSA was, in effect, a compromise solution to move forward from 15 years of nationwide debate. ESSA mandates annual state testing for students in grades 3 through 8 and once in high school with each state given the flexibility to offer its choice of normed, summative assessment. As a single end of year, normed assessment is insufficient to support teaching, learning, and decision making across stakeholders, states subsequently adopted a patchwork of recommendations and mandates for additional assessments throughout the school year to evaluate student performance, predict student outcomes on year-end assessments, and support instruction. Each district has the responsibility to manage and apply its own unique collection of assessment data. Policy experts responded to the assessment patchwork and variability in effective use by encouraging districts to adopt a comprehensive assessment system to ensure assessments are chosen with clear purpose, and assessment data is applied for the greatest benefit.

Standardized Assessment and COVID-19

The COVID-19 pandemic seems to have ushered in a renewed interest in the value of educational assessments and how they can best fulfill their overriding mission to improve academic outcomes for all students. As a result of near universal school closures in Spring 2020, teaching and learning was highly variable and ESSA-mandated state testing was effectively suspended. Students returned to school in fall 2020 and there were widespread concerns about student academic performance. Normed assessments in winter and spring 2022 confirmed widespread suspicions that students were underperforming historical norms as a result of the school closures (retrieved from the [NCES](#), November 1, 2022). As such, assessment is now receiving a new wave of attention, as educators look to normed assessments to identify what gaps exist and for whom, so they can identify solutions to make up for lost learning time.

It is heartening to see the collective ecosystem renew its appreciation of normed assessments as a valuable tool to understand and address student learning needs, rather than as a weapon for punishment and control. However, one must acknowledge how the pandemic also magnified the flaws in schools’ patchwork of assessments. We see this as an opportunity for the industry to re-think the existing components of a comprehensive assessment system. This report seeks to describe those components, evaluate where they might not be meeting their intended objectives, and offer a proposal to make the current system more equitable and effective in meeting the diverse needs of all stakeholders, teachers, school and district administrators, and parents and students, by the addition of universal cognitive assessments.

Existing Elements of a Comprehensive Assessment System

As a starting point, we attempt to clarify the language around assessment which can be inconsistent, particularly when assessments are designated to serve multiple purposes for a variety of stakeholders. To minimize this confusion, we suggest using the assessment naming conventions proposed by Measures that Matter (Perie, et. al. 2007). We next analyze how well existing assessments are meeting objectives of their primary stakeholders before suggesting opportunities for improvement.

Measures that Matter classifies assessments into three broad categories:

Summative

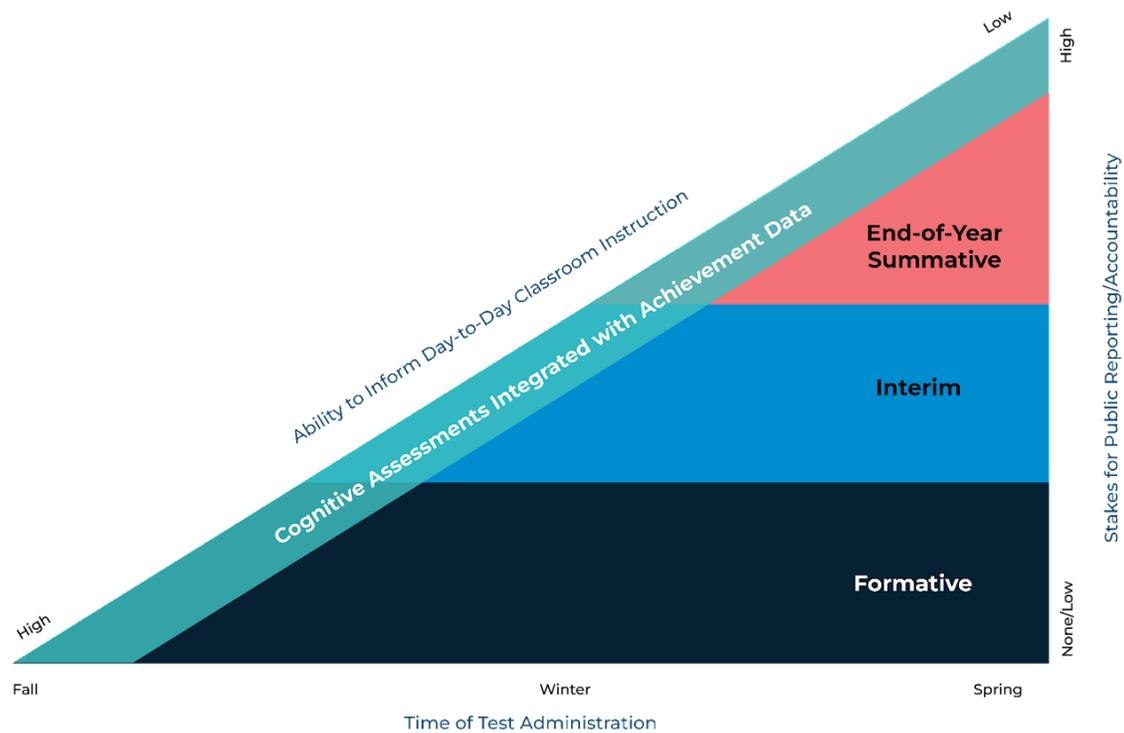
These are standards-based assessments generally given one time at the end of a school year to identify student achievement strengths and weaknesses. These are normed, state assessments designed to provide objective feedback on performance relative to same-aged peers statewide and nationally. The format is typically multiple choice or short answer for ease of grading and aggregating data and takes students longer than a formative or interim assessment.

Formative

These classroom-specific just-in-time assessments provide teachers and students ongoing feedback on academic knowledge. They are typically much shorter than interim assessments and designed to measure the specific skills the teacher or school deems most important. They range in format and include teacher-created essays, projects, and multi-step problems or can be compiled from large question banks aligned to standards. Formative assessments offer opportunities to understand depth of knowledge in ways that multiple choice and short response items cannot. Conversely, the results can be difficult to aggregate and grading can include varying degrees of subjectivity.

Interim

These standards-based assessments are generally shorter than summative assessments, longer than formative assessments, and are often administered 2-3 times during the school year to inform teachers and administrators of academic strengths and weaknesses so adjustments can be made prior to year-end summative assessments. These assessments are usually district-created or nationally-normed and questions are usually objective, multiple choice or short answer, so they can be efficiently graded and aggregated across classes and grades.



Adapted from the Graphic Originally Developed by Perie, Marion, Gon, and Wurtzel (2001)

As one might imagine, these different types of assessments serve very different needs for stakeholders. In Table 1 we summarize the purpose for which a key stakeholder needs an assessment, what actions can ideally be taken based on the assessment results, and the corresponding category of assessment typically used. We also identify the primary stakeholder. That is not to say that there are no other important stakeholders. Rather it is to clearly prioritize who needs the data most to fulfill an overriding objective.

Table 1: Overview of Assessment

Objective	Actionable Next Steps	Primary Stakeholder	Assessment Category/Timing
Short-term Planning	<ul style="list-style-type: none"> • Arrange classes based on mastery • Set achievement goals 	School Leader	Summative (from prior year)/Interim
Progress Monitoring	<ul style="list-style-type: none"> • Adapt instruction • Provide learning strategies • Support teacher efficacy 	Classroom Teacher	Formative, Interim
Learner Agency	<ul style="list-style-type: none"> • Help student adapt how they learn • Teach students to self-advocate 	Student or Parent	Formative
Early Warning	<ul style="list-style-type: none"> • Identify at-risk students • Provide interventions 	School Leader	Interim, Formative

**Strategic
Multi-Year
Decision Making**

- Identify trends
- Make changes to curriculum and instruction

District Leader

Summative

Risks of an Inadequately Designed Assessment System

We proffer that the objectives outlined in Table 1 are near universal, but that delivering on actionable next steps by the primary stakeholder is highly inconsistent. The reason for the inconsistency can be attributed to a number of factors including, though not limited to:

- Poor quality (or perceived) of assessment data so stakeholders do not want to act on it;
- Limited actionable next steps provided by assessment creators;
- Inadequate training on how to understand the data and take actionable next steps;
- Lack of clarity on the key stakeholder; and
- Insufficient accountability for the key stakeholder.

In Table 2, we highlight the primary problems that arise when districts do not have a well-designed comprehensive assessment system. Most districts struggle with at least one of these three major problems which negatively impact teaching and learning.

Table 2: Risks of an Ill-Defined Comprehensive Assessment System

Problem	Negative Impact
Too many assessments to meet needs	<ul style="list-style-type: none"> • Unnecessary time spent testing at the expense of learning activities • Backlash against “too many assessments” from teachers, parents, and students
Too few assessments to gather necessary data	<ul style="list-style-type: none"> • Inadequate data to address objectives • Key stakeholder does not take action or takes sub-optimal actions (i.e. students who need supports don’t receive them)
Data exists but it is not used effectively	<ul style="list-style-type: none"> • Teachers do not adapt their teaching (i.e. students who need supports don’t receive them) • Student outcomes do not improve

Assessment Performance since COVID-19

Despite a rocky history for acceptance of summative assessments in US public schools, there was widespread acknowledgment that summative assessments accurately captured the learning gaps that most stakeholders perceived following COVID-19. Objective measures identified which students lost the most learning, by how much, and in which subjects. While student performance was disappointing, on a positive note, one can say that there seemed to be collective acceptance that standardized assessments played a valuable role in identifying key data trends and supported administrators in both short and long-term decision making.

The emerging assessment story is, however, not all positive. There is still widespread confusion about how classroom teachers should intervene for at-risk students and adapt instruction for the larger than usual variability they are seeing in their learners. If classroom teachers are struggling to understand how to fill learning gaps, it is a clear indicator that [formative and interim] assessments for progress monitoring, learner agency, and early warning are falling short.

This failure was, in fact, predicted by Perie et. al. as they explained that interim assessments, despite being effective for predictive and evaluative purposes, “generally fail in providing rich diagnostic feedback regarding student thinking. That is, few provide any information on why a student answered an item incorrectly or how best to provide corrective feedback.” Logically one might suppose that formative assessments could and do provide that information.

Unfortunately, while formative assessments can and do provide teachers some of those actionable next steps it is not nearly at the universality teachers need. Perhaps in decades past with smaller classes and more homogeneous student bodies it was possible for formative assessments to meet classroom needs much of the time. However, in today’s world of highly variable student achievement (particularly post-pandemic), larger class sizes, and a higher percentage of less experienced teachers, existing formative assessments are inadequate for what most teachers need. Yes, teachers have plenty of data on gaps, but they lack the resources to interpret those results and convert them into the strategies and interventions students need most.

We suggest that normed cognitive assessments can affordably and efficiently fill the gaps left by formative and interim assessments. Normed cognitive data enables teachers to interpret their achievement data, efficiently adapt instruction, provide strategies to individual students, and offer more structured interventions when needed. Using cognitive data, teachers are also well-suited to develop learner agency by helping students to recognize their struggles, use strategies, and self-advocate when necessary.

Cognitive Assessment: The Actionable Next Step

Unlike achievement assessments that measure academic learned knowledge and acquired skills, cognitive assessments measure learning capacities in the domains of complex reasoning (verbal, abstract, and spatial), executive functions (attention, working memory, and flexible thinking), memory (visual and verbal), and speed (visual processing and visual motor). These diverse capacities are essential for learning across contexts and together are the most

powerful single predictor of academic achievement, whether measured by grades or standardized academic achievement tests (Bergold, Sebastian & Steinmayr, Ricarda, 2018). Perhaps more importantly than their predictive capability, assessments of distinct cognitive skills provide clear insight into the most likely reasons for gaps in learned knowledge. As such, they can be readily aligned to strategies to provide teachers with insight into student thinking and the best way to provide corrective feedback.

Historically cognitive assessments have required one-to-one administration by a certified clinician, creating significant cost, time, and resource-based limitations on their usage. Fortunately, a recently developed adaptive online cognitive assessment can be used as an affordable, scalable addition to a comprehensive assessment system. The one-hour, online self-administered, self-paced group assessment from MindPrint Learning is a perfect example. MindPrint, like other strictly cognitive assessments, has the flexibility to be administered at any point in the year since it is not designed to measure acquired knowledge at a particular point in time. This allows schools to schedule around stricter testing windows offered by normed achievement assessment providers. Reports can be delivered within a few school days to give teachers actionable next steps to support students' performance.

While the validity and reliability of cognitive assessments is widely accepted and used in special education, the use in general education environments is fairly limited. As such, we provide specific examples of common perplexing achievement results teachers struggle to interpret from current assessments and explain how cognitive data can effectively provide insights where existing assessments leave classroom teachers guessing for answers. Furthermore, that insight provides long-term value for lifelong learning well beyond the current academic year.

Below are four common examples of challenges that teachers observe, but might misunderstand the root cause of a student's performance. The student's cognitive skills provide the most likely reason and the best way to support the student in the classroom. If a student is properly supported in school, this should lead to long-term improvement in the student's self-efficacy skills:

Example 1: Student Does Well on Quizzes but Struggles on Chapter Tests

Cognitive Insight	Likely Reason	Actionable Next Step	Long-term Value
Student has weaker than average memory skills	Student had difficulty retaining the breadth of material for a longer test and/or forgot what they previously learned	Encourage spaced repetition of all material throughout the chapter to ensure long-term retention	Student learns to space studying and avoid long-term learning gaps across subjects
Student has a weaker reasoning skill	Student didn't learn deeply but gaps are missed on brief quizzes and become obvious on in-depth tests	Re-teach using student's strongest reasoning skill to ensure mastery	Students avoids long-term learning gaps and has deeper understanding
Student has weaker executive functions	Student can focus for shorter quizzes but has difficulty sustaining focus for longer assessments	Require student to show and check all work to catch avoidable mistakes and take regular breaks	Student learns to self-monitor and manage heavier workloads in higher grades

Example 2: Student Does Well on Classroom Tests but Poorly on Standardized Tests

Cognitive Insight	Likely Reason	Actionable Next Step	Long-term Value
Student has good reasoning but slow processing speed	Student has difficulty finishing all questions in allotted time to the best of his/her ability	Teach students how to pace themselves for longer tests or break-up over multiple sessions when possible.	Student develops time management strategies and is better able to manage heavier workloads in higher grades
Student has good reasoning but weaker flexible thinking	Student can answer straight-forward questions but has difficulty with “best answer” or questions asked in an unfamiliar format	Provide structured steps to analyze questions when unsure of the answer	Student grows more comfortable with ambiguity and strengthens complex problem solving
Student has good abstract reasoning and weaker verbal reasoning	Student struggles with directions and/or the amount of reading, even on Math and Science sections	Teach how to break down questions and use pictures to aid in problem solving	Student continues to develop STEM strengths even if weaker verbal skills

Example 3: Student Does Well in Reading but Poorly in Math (or vice versa)

Cognitive Insight	Likely Reason	Actionable Next Step	Long-term Value
Student has stronger verbal reasoning but weaker abstract reasoning (or vice versa)	Student more naturally understands what they read but has more difficulty when they need to interpret numbers, diagrams, or charts (or vice versa)	Provide students with opportunities to digest challenging concepts in their stronger skill (e.g. provide a text description of the diagram)	Student learns to transform content into their more natural format when working independently (e.g. describe the diagram in their own words)
Student has weaker attention or working memory	Student makes “careless errors” when tired; depending on the chapter topic and time of day student might underperform	Teach strategies to stay focused in class and how to check work	Student develops self-awareness of focus and learns to more effectively self-manage
Student’s memory skills are not as strong as reasoning skills	Student understands a lesson but forgets; this can be a bigger challenge in subjects that are more cumulative like math and foreign language	Provide students structured repetition to review core material regularly so it is not forgotten	Student learns to review key content regularly, so it is not forgotten and appreciates the difference between understanding and remembering

Example 4: Student is below proficiency standards and disengaged in class

Cognitive Insight	Likely Reason	Actionable Next Step	Long-term Value
Student is weak in abstract and/or verbal reasoning	Student is struggling to understand the content and might be embarrassed to ask for help or has “given up”	Provide student with small group support or one-to-one intervention to fill gaps and boost self-confidence	Student learns they can succeed but might need to ask for help
Student is exceptional in abstract reasoning and/or verbal reasoning	Student is bored or does not see the value; often student has a weaker attention or memory skill that causes frustration	Re-engage student by acknowledging their strength and finding opportunities to tie that strength to their interests	Student sees purpose in learning and develops interests that use their strengths
Student has weak attention	Student cannot focus in class for even short periods of time, particularly in subjects that are not of natural interest; student stops trying to focus and disengages	Provide student with ways to sustain focus including frequent breaks, stretching and meditation; provide class notes so the student does not fall behind	Student develops self-awareness of focus and learns to use strategies

As the examples above demonstrate, cognitive assessments can offer additional potential benefits that extend well beyond filling the gaps left by other assessments. The more students struggle, the greater the need for more frequent progress monitoring. However, when teachers and students are able to use strategies that improve learning, overall performance improves and the overall need for testing thus declines. Hence, not only are online cognitive assessments efficient in their own right (typically one hour, every three years), they have the potential to reduce overall student testing time across the comprehensive assessment system.

The inherent stability of cognitive skills also provides the opportunity to significantly improve learner agency above and beyond the capacity of formative assessments. When students understand their personal cognitive variability, they can make changes to their approach to learning when they are not succeeding, e.g. studying more frequently, pacing themselves on tests, and using visual models to support their understanding. Once students understand that their cognitive strengths and needs will carry over class-to-class and year-to-year, they are far more likely to invest the time and energy to understand and use the strategies that will boost their efficiency. This is in sharp contrast to student response to most achievement assessments where students rarely return to learn from their mistakes. Whereas students might not see the benefit of correcting achievement weaknesses if the class has moved on to the next topic, the relative stability of cognitive skills provides a strong incentive for students to understand and support them.

Finally, cognitive assessments offer the important value of explaining the variability in student performance across the disparate assessments administered throughout the year. The underlying assumption herein is that stakeholders believe in the quality and reliability of the assessments. However, this is far from a given. Historically teachers have expressed

skepticism for assessments when they note inconsistency in individual student performance across interim and summative assessments, or when students' standardized assessment performance diverges from their formative classroom assessments or observations. When teachers question the validity and reliability of interim and summative assessments they are less likely to adapt their teaching practices and more likely to resist mandated changes in curriculum and instruction based on the results. As Perie confirms, if assessments are to be used for instructional purposes teachers must believe in the results. Cognitive assessments have the power to explain the reasons for student variability across assessments, thus increasing the willingness of all stakeholders to use assessment data to make well-informed decisions.

Assessment and Equity

One clear value of formative assessment over standardized assessment is the allowance for teacher judgment. Teachers can evaluate student mastery in ways that can never be captured on a standardized test with personal knowledge such as “what did they really mean”, was the student “having a bad day” or how much effort the student put in. However, these clear benefits of formative assessments undoubtedly also leave room for unintentional bias. When a teacher gives one student credit for effort, they are implicitly assuming another student might not have been as motivated. If one student informed them it was a “bad day” and the other did not, teachers are, intentionally or not, evaluating students with inconsistent data. Objective, normed summative and interim assessments provide a necessary check on the subjectivity of formative assessments. When the results of formative and normed assessments don't match, cognitive assessments provide the often needed “tiebreaker”.

Cognitive assessments also offer the opportunity to reveal talented students who are being lost in the system. While it could be easy to dismiss a student's higher than expected score on a normed assessment as luck or an outlier, cognitive assessments will suggest if the standardized score is a better reflection of that student's true capabilities. They shine a spotlight to explore why that student is underperforming on classwork. Teachers might easily attribute a student's poor performance or behavior to disengagement, failing to realize a student is struggling with executive functions or is “bored” given superior reasoning skills. Conversely, it can be convenient to dismiss a low standardized test score for a student with good grades, attributing the low score to test anxiety or a bad day. Integration with cognitive assessment will ensure that early warning indicators (example 2 above) are not being ignored.

Any discussion of reliability and validity of normed assessments cannot ignore the understandable concerns raised about racial and socio-economic bias in standardized assessments. We tackle one small aspect of that debate with a clear acknowledgment that normed achievement assessments tend to be highly correlated with reading skills, even when reading is not the target skill being measured. When tests of math knowledge require good reading skills, poor readers struggle, which disproportionately affects lower income students. This is a valid concern that should be addressed by normed achievement assessment providers, but cognitive skills can offer at least a partial solution. Students with weaker verbal reasoning and memory skills can be readily identified by cognitive assessments. We recommend that results of non-reading summative assessments should be interpreted with this consideration. In short, cognitive assessments can at least shine a light on high potential [STEM] students whose strengths are going unrecognized because of

reading challenges, a problem that affects students of all socioeconomic status though disproportionately students in lower income communities.

Summary

There is no quick fix for addressing the achievement declines that emerged during the COVID-19 pandemic, as school closures only exacerbated the pre-pandemic challenges around students' proficiency in reading and math. However, the pandemic has provided the opportunity to re-establish the value that a comprehensive assessment system can bring to teaching and learning. While existing assessments meet many needs, there are undeniably places where they fall short. This paper provides support for adding a cognitive assessment to schools' comprehensive assessment systems to make achievement results more understandable and actionable so that key stakeholders can more effectively act on their data. While cognitive assessments will not address all the problems with assessment, they hold the promise of filling many of the inadequacies both affordably and efficiently and with an overarching level of fairness.

About the Author



Nancy Weinstein is the Founder and CEO of MindPrint Learning. Prior to founding MindPrint, Nancy worked in industry at Bristol-Myers Squibb, The Walt Disney Company and Goldman Sachs. She has a BSE in Bioengineering from University of Pennsylvania and an MBA from Harvard. She is co-author of *The Empowered Student* and is a national speaker on using science of learning best practices to improve student outcomes. She is the Principal Investigator on two National Science Foundation grants.