thinkpoint

Addressing assessment issues: Value of using a two-component accountability assessment system

by Stuart Kahl, Ph.D.

Introduction

In "Re-Balancing Assessment: Placing Formative and Performance Assessment at the Heart of Learning and Accountability," authors Hofman, Goodwin, and Kahl summarize the currently oft-mentioned concerns about educational assessment and then propose a new formula for assessment systems that truly eliminates any barriers between instruction and assessment. They propose a twocomponent accountability assessment system.

One component, not new, is end-of-year, on-demand testing involving both machine-scorable item formats and human-scored tasks – constructedresponse questions and short performance tasks.

The second component consists of Curriculum-Embedded Performance Assessments (CEPAs). The addition of the second component could lead to efficiencies in implementing the first, including shorter tests, matrix-sampling, etc.



This Think Point document was produced by the Michigan Assessment Consortium (MAC) to address a recognized issue or challenge in educational assessmentand to contribute to an ongoing conversation about how best to serve teachers and learners through high-quality assessment practices. The ideas expressed here are presented for information and reflection only and do not necessarily represent an official position of the MAC or its board. Let's talk! Send feedback to info@michiganassessmentconsortium.org

What is a CEPA?

A CEPA is a multi-day, perhaps multi-week, instructional unit consisting of a series of instructional and assessment activities, some of which lead to student work that can be evaluated for formative purposes and some of which yield student work that can be scored for local summative and external accountability purposes. The paper describes how a state might develop and "phase in" the CEPA component. Figure 1 provides a brief example of a CEPA in science. A fully developed CEPA would identify the relevant content standards and learning targets, and include additional guidance for instruction and assessment, as well as scoring rubrics and sample student work.

As mentioned, there are many concerns about current assess-

ment systems. There are also concerns, myths, and misconceptions about performance assessment. The paper's authors believe that all of these are addressed by CEPAs as they describe them and their development and implementation. The table in Figure 2 identifies the various concerns and issues and explains how they are mitigated by CEPAs.

Figure 1: Sample CEPA—Heat Transfer

Activity 1

Students individually or in small groups research methods of heat transfer. They discuss what they have learned about conduction, convection, and radiation (student-guided learning).

Activity 2

Teachers check student understanding of methods of heat transfer via ungraded quizzes, interviews, or class discussion (formative assessment evidence gathering, feedback, and adjustment).

Activity 3

In small groups, students design and conduct an experiment to determine which of two fabrics better protects against the winter cold. Materials required include tin coffee cans of different sizes (with lids), two different fabrics (e.g., plastic and wool), fasteners, thermometers (thermal probes), timers, and hot water (performance activity).

Activity 4

Students individually write up a formal lab report of their experiment (graded summative product).

Activity 5

Teachers, via questioning, lead class discussion of how methods of heat transfer played a role in the design and implementation of the research (formative assessment reflection and reinforcement).

Activity 6

Students individually research how a home heating system works and write a paper describing a home heating system and how different methods of heat transfer are involved (graded summative product).

Figure 2: CEPA Solutions to Concerns with Current Assessment and Accountability Systems

Issue/Concern	How Proposed Use of CEPA Addresses Concern
Current efficient account- ability tests negatively impact local instruction and assessment, focusing both on low level knowledge.	CEPAs (instructional units) attend to both foundational knowledge/ skills and deeper learning. Planned formative assessment address- es both. Performance tasks address deeper learning, requiring the application of foundational knowledge and skills to higher order tasks.
Current instruction is teacher-driven with students playing a passive role.	Educational reformers are calling for significant changes in how teachers and students spend their time. Formative assessment and curriculum-embedded performance assessment, done right, accomplish that.
Student engagement is lacking and motivation to learn is low.	CEPAs involve students in engaging, real-world tasks both in and out- side the classroom. CEPAs can allow choice on the part of students with respect to roles they play in group and individual activities.
Teacher-made tests are sometimes of low quality, and teachers' capacity to evaluate student work is sometimes inadequate. Thus, performance assess- ment results are unreliable.	Gathering evidence of student learning for both formative and summative purposes is planned in CEPAs. The tools/measures are tried-and-true, having been reviewed and tested during the devel- opment process. The use of the CEPAs and associated supporting materials by teachers will enhance their assessment skills. The initial training and the scoring auditing we envision will build teacher capac- ity, as will ongoing teacher collaboration. Experience in many state assessment programs has shown that teachers can be trained to score consistently.
"Formative assessment" has been misinterpreted to mean frequent use of grad- ed quizzes and tests.	CEPA directions clearly differentiate between formative assessment evidence gathering and summative assessment tasks. An appropriate balance of the two should be a CEPA feature.
There is too much testing.	Actual testing time for state accountability testing is not too much. Putting instruction and learning on hold in order to prepare for those tests is problematic. Also, research has shown that over-testing that occurs pertains to the frequent local use of external interim assess- ments. The use of CEPAs as proposed could reduce or eliminate the need for such assessments. <i>(Continued next page)</i>

Figure 2: CEPA Solutions to Concerns with Current Assessment and Accountability Systems (Continued from previous page)

Issue/Concern	How Proposed Use of CEPA Addresses Concern
There is inconsistency between what is tested for accountability and what is being taught.	State tests cover state standards; so should instruction. If the incon- sistency is because the state tests cover only lower level knowledge and skills and local instruction also addresses higher order skills (not often the case), the two-component approach proposed for account- ability assessment, with CEPA performance tasks tapping deeper knowledge, addresses this problem well. Also, because CEPAs are curriculum-embedded, if the curriculum is aligned with standards, then the CEPAs will be as well.
There is no teacher owner- ship in the accountability assessment.	As proposed, CEPAs are initially drafted by teachers. Teachers can be involved in the selection of the state-approved CEPAs to be used in their schools, and they are totally responsible for the implementation of them, including the scoring of their students' work, which is audited by the state.
Teacher-developed performance tasks are of low quality.	The CEPAs used for accountability assessment, while drafted by teachers, undergo the same committee review and piloting steps as other state assessment tools to assure alignment to standards and technical quality. These CEPAs serve as good models for other CEPAs teachers develop, and states should provide training in the development and use of CEPAs.
Performance assessment is unreliable.	Often, the reliability of a single performance task is compared to that of a test of 50 or more multiple-choice items. The proposed approach to the CEPA component for accountability is for student work from multiple CEPAs to be counted toward accountability results. That could result in the CEPA component alone being as reliable as a 50- item multiple-choice test. In combination with the on-demand compo- nent, even one CEPA student work product would suffice for a reliable total score, but the intent is for multiple high quality CEPA summative measures to account for much more of the total score.
It takes too long for ac- countability assessment results to be delivered to the schools.	CEPAs, instructional units, address important curricular standards and should replace other units covering those standards. They are not an add-on. The performance tasks and other summative measures within CEPAs take the place of end-of-unit tests associated with the units the CEPAs replace.
Scoring performance assessments is expensive.	Student work from CEPAs is scored by teachers as part of their regular instructional programs. Costs of auditing processes for accountability purposes, including some central scoring done on a sampling basis, can be offset by savings from shortening the on-demand tests or by the use of matrix-sampling techniques for the on-demand component.

Issue/Concern	How Proposed Use of CEPA Addresses Concern
The contribution of per- formace components to accountability results is too little to be worth the time, effort, and expense.	The proposed system calls for the CEPA component to count significantly toward accountability results.
Achievement gaps are not diminishing fast enough, if at all.	Formative assessment as implemented in conjunction with CEPAs represents the instructional process that research shows can be es- pecially effective in enhancing student learning. It has been shown to be more effective with disadvantaged and underachieving students, thus capable of reducing achievement gaps. Also, the engaging, highly motivating activities in CEPAs are intended for all students. This is in contrast to enrichment tasks in traditional curriculum materials that are typically reserved for only the highest achieving students.
U.S. students are falling far- ther and farther behind the students in other nations in terms of achievement and are unprepared for college and careers.	This expressed concern often pertains to higher order skills which various international tests are purported to measure. Performance tasks within CEPAs tap higher order skills. At the same time, CEPAs do not shortchange the foundational knowledge and skills that must be applied by students to succeed on the higher order tasks.

Summary

Ultimately, the promise of CEPAs is that they provide a more motivating, robust, and balanced way to measure student learning. If we believe the maxim that what you test is what gets taught, then these new measures hold the promise of driving many positive changes throughout the system—from better engaging students to supporting deeper learning to encouraging new classroom practices to supporting greater teacher collaboration. Although better measures alone won't address all of the challenges facing schools, the authors believe a new formula for measuring student success may be what is most needed to put our nation's schools on a path that breaks through performance ceilings and creates a generation of highly motivated students engaged in deeper learning.



STUART KAHL, PH.D., is an independent consultant and former founder and CEO of Measured Progress, Inc.

Dr. Kahl has earned a bachelor of arts degree in mathematics and masters in education degree from the Johns Hopkins University. His doctorate from the University of Colorado was in curriculum and instruction and focused on research and evaluation methodology and mathematics education.

Reflection Points

- 1 How would you define "performance assessment"? What are its important characteristics?
- 2 What do you see as obstacles to the implementation of performance assessment at a significant level both locally and in statewide accountability assessment?
- 3 Are CEPAs the best way to address deeper learning in instruction and assessment? If not, what approach(s) would be better?
- 4 In what ways would the significant use of CEPAs change classroom instruction?

To learn more

The Michigan Assessment Consortium website contains a rich library filled with sharable assessment resources as well as online learning modules designed to increase assessment literacy among all education stakeholders. Explore these resources, then spread the word! Visit www.michiganassessmentconsortium.org.



Resources for further study

Documents

- Hofman, P., Goodwin, B., Kahl, S. (2015). Re-Balancing Assessment: Placing Formative and Performance Assessment at the Heart of Learning and Accountability Denver, CO: McREL International. mcrel.org/cepa-white-paper
- Kahl, S. (2017). Replace, Don't Add On: Giving Performance Assessment a Chance. Dover, NH: Measured Progress, Inc. (also published in eSchool News, 4/24/2017 as "How an edtech innovation is giving performance assessment new life"). measuredprogress.org/wp-content/uploads/2018/09/Replace-Dont-Add-On.pdf

 Marion, S. & Shepard, L. (2010). Let's not forget about opportunity to learn: Curricular supports for innovative assessments. Dover, NH: Center for Assessment. nciea.org/sites/default/files/publications/Marion_Shepard_Curricular_units_042610.pdf

Pecheone, R. L. & Kahl, S. R. (2010). Through a Looking Glass: Lessons Learned and Future Directions for Performance Assessment.

Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education. scale.stanford.edu/system/files/through-looking-glass-lessons-learned-and-future-directios-performance-assessment.pdf

Programs

- Performance Assessment of Competency Education (PACE) pilot program (still operating) receiving waivers from USDOE under NCLB and ESSA and using teacher-developed and scored performance tasks for accountability in subset of federally-mandated target grades. education.nh.gov/assessment-systems/pace.htm
- Ohio Performance Assessment Pilot Program (OPAPP) pilot program (ended) funded by Race to the Top and involving teacher-developed and scored DYADS, each consisting of an instructional activity and an assessment activity.

education.ohio.gov/Topics/Learning-in-Ohio/Ohio-Performance-Assessment-Pilot-Project-OPAPP

- Literacy Design Collaborative (LDC) community of practice with partners in 50 states (including state departments of education in Arkansas, Georgia, Kentucky, Louisiana, and Colorado) and hundreds of participating schools and districts developing and using instructional ladders (modules) leading toward writing tasks and reflecting authentic disciplinary literacy. Idc.org
- Quality Performance Assessment (QPA) program of the Center for Collaborative Education in Boston, with participating schools and districts mostly in New England states and which provides processes and tools that leverage complex, multi-step performance assessments as an essential tool of educational equity and teaching. The RI and NH state departments of education are QPA partners. cce.org/work/quality-performance-assessment

Policy

- Every Student Succeeds Act (ESSA) requirements for multiple measures (defined to include performance and portfolio assessments) and for quick turnaround of results; option of interim assessments with results to be aggregated; and innovative as sessment pilot option with waivers from regular state testing for participating students.
- Rhode Island Diploma System guided by The Council on Elementary and Secondary Education 2016 Secondary Regulations and requiring students to complete two Diploma Assessments chosen by the district or school (student portfolios, exhibitions, senior projects and/or comprehensive course assessments).



Michigan Assessment Consortium | 1980 North College Road | Mason, Michigan 48854 www.michigan assessmentconsortium.org