



November 2016

Do Formative Assessments Influence Student Learning?: Research on i-Ready and MAP

Alanna Bjorklund-Young and Carey Borkoski

Research Fellows

Introduction

In an era that is seemingly saturated with standardized tests of all stripes, it is easy to forget their varying functions. Formative assessments, or assessments that are given while the learning is taking place, can serve as valuable tools for increasing students learning if they give educators insights into what students know or don't know. Teachers can subsequently address gaps in understanding or misperceptions, while continuing to build on the knowledge students have demonstrably gained. It is through such feedback and changes to the teacher's actions that tests improve student learning, not through giving the test alone (Black, 2015; Heitink et al., 2016). The formative tests must therefore impart timely and relevant information that the teachers can interpret and deploy (Heitink et al., 2016).

Summative assessments, in contrast, mainly serve a different purpose. They are primarily retrospective, i.e., designed to determine whether students learned the material that has supposedly been taught already. State summative assessments (i.e. state standardized tests) can also serve as accountability tools with which to evaluate the efficacy of districts, schools, and teachers. Because student scores are usually received months after the tests are administered, teachers cannot use the information to adjust classroom instruction. That task rests with formative assessments.

With this important role of formative assessments in mind, two popular formative assessments, i-Ready and Measures of Adequate Progress (MAP), purport to provide teachers with timely and relevant information about their students' learning as those students also prepare for summative tests. Indeed, the publishers of both assessments claim that their assessments precisely measure students' learning, and thus enable teachers to monitor students' learning targets to meet end-of-year learning goals measured on state summative tests. Both i-Ready and MAP are computer-adaptive; they are easy to administer and provide rapid response data.

Therefore, if i-Ready and MAP give teachers valid information about students' learning that is relevant and informative and also aligned with state summative tests, then they can serve as an important resource for helping teachers modify their teaching. What do we know about the research base of these tests? Specifically, have researchers established that these tests give teachers relevant information about their students' learning?

The Validity Research on i-Ready and MAP That We Have

Researchers have conducted numerous analyses of these two assessments. Unfortunately, parties associated with the publishers of the assessments have authored the studies, which inevitably calls objectivity into question. For example, Curriculum Associates, which owns i-Ready, hired Education Research Institute of America (ERIA) to evaluate the tool. More problematic still, MAP, which is published by Northwest Evaluation Association (NWEA), used its own in-house researchers to conduct validity research. In both cases, the resulting research has not been published in peer-reviewed journals. Completely impartial, peer-reviewed research is obviously preferable. In the absence of such research, we report Curriculum Associates' and NWEA's findings.

Before considering the research findings on i-Ready and MAP, it is useful to think about how researchers establish test validity. Validity means that the assessment measures what it says it does, and therefore that the tests may be used for the purpose the publishers claim—in this case, that the formative assessments can be accurately used to inform teacher instruction in preparation for state or common core summative tests. Crocker and Algina (1986) define three main types of validity studies: (1) content validity, which establishes the alignment between the test questions and the content it is intended to assess (e.g. this can be established by comparing i-Ready and MAP assessments to state standards)¹; (2) criterion-related validity, which establishes the relationship between the test score and an outcome or measure that has already been validated (e.g. this can be established by comparing i-Ready and MAP scores to the state or common core summative pre-test scores or other practice materials created by the summative test makers); and (3) construct validity, which establishes the extent to which the test measures what it purports to measure (e.g. this can be established by comparing i-Ready and MAP scores and student state or Common-Core-aligned summative test scores themselves). Existing research on the validity of i-Ready and MAP tests takes the third approach.

One way that researchers can compare i-Ready and MAP to state standardized or Common Core tests (i.e. PARCC and Smarter Balanced Assessments, or SBA) is by looking at the correlation between the two tests. Correlation is a measure of how two things move together.² The closer these findings show the correlation between i-Ready or MAP and the standardized test is to one, the more the test scores move together. Researchers often use a benchmark of 0.70 to indicate “strong correlation.”³ Curriculum Associates has shown that i-Ready is strongly correlated with the New York State (NYS) test, the SBA, and with the PARCC—with correlations between i-Ready and the NYS tests ranging from 0.74 - 0.86⁴ (ERIA, 2016a); correlations between i-Ready and SBA ranging from 0.82 - 0.85 (ERIA, 2016c); and correlations between i-Ready and PARCC ranging from 0.77-0.84 (ERIA, 2016b). NWEA researchers similarly found strong correlations between MAP and state and both Common Core tests—with correlation between MAP and state tests ranging from 0.70 - 0.92 (NWEA, 2016a-m); correlation between MAP and SBA ranging from 0.80 - 0.89 (NWEA,

¹ See, for example, Martone and Sireci (2009) for more information on this approach.

² For points of reference, a correlation of 0 indicates that two test scores do not move together at all; a correlation of -1 means that the two tests scores move in opposite directions, i.e. one test score always increases while the other test score always decreases; and a correlation of 1 indicates that the tests always move together in the same direction.

³ For example, the Center on Response to Intervention at the American Institutes for Research sets a criteria of 0.70 or higher. Retrieved at: <http://www.rti4success.org/resources/tools-charts/screening-tools-chart/screening-tools-chart-rating-system>

⁴ There is a range in the correlations because each grade and subject has a different correlation.

2015); and correlation between MAP and PARCC ranging from 0.73 - 0.91 (NWEA, 2016n-p). Thus, all of the correlations between i-Ready or MAP and standardized tests are strong.

Another way that researchers can compare i-Ready and MAP to state standardized or Common Core tests is by looking at how often the formative assessments were able to accurately predict proficiency on the summative assessment (i.e. the state or Common Core test); this measure can ideally be used by teachers to predict how well their students will perform on the summative assessments. To do so requires a ratio of the number of correct predictions (i.e. correctly predicting that a student is proficient or correctly predicting a student is not proficient) to total predictions (i.e. all students—students correctly predicted and students not correctly predicted). Here as well, researchers show that both formative assessments have high accurate-prediction rates. For example, i-Ready accurately predicts the proficiency of 88% and 87% of students on average on the NYS tests in math and reading respectively (EIRAa); 84% and 83% of students on average on the SBA in math and ELA respectively (ERIA, 2016c); and 83% and 81% on average on the PARCC in math and ELA respectively (ERIA, 2016b). The MAP accurately predicts the proficiency of between 76% and 93% of students on state tests (NWEA, 2016a-m); 88% and 84% of students on average on the SBA in math and ELA respectively (NWEA, 2015); and between 76% and 91% of students on the PARCC (NWEA, 2016n-p).

Accurately predicting proficiency, however, is not as informative a measure as we might wish. Proficiency is only a binary outcome—a student is either proficient or not—but test scores often have four or more outcomes. Therefore, a more rigorous and informative statistic is how accurately the formative assessment predicted a child’s actual score on the summative assessment. Only i-Ready research provides information on this. For example, i-Ready accurately predicts students’ final scores on: the NYS test for 68% and 66% of students on average for math and reading respectively (EIRA, 2016a); the SBA for 61% and 60% of students on average for math and reading respectively (EIRA, 2016d); and the PARCC for 58% and 55% of students on average for math and reading respectively (EIRA, 2016e). These results are less impressive than accurate prediction of proficiency. This is unsurprising, because correctly predicating one of four scores is more difficult than correctly predicting one of two scores (i.e. proficient or not proficient).

The Validity research on i-Ready and MAP That We Need

There is another hurdle to overcome: even i-Ready’s or MAP’s accurate prediction of students’ standardized summative test scores is not informative enough to determine whether these formative assessments are valid for the purposes of changing teacher instruction. For this, we would need to know how predictive the formative assessments are on a more granular level. After all, if teachers want to use the tests to change their teaching practices, they certainly need more information than whether or not their students are proficient or what their students’ final scores on the standardized test will likely be. For this, teachers need to know *why* their students receive the scores they do. They need to know which specific skills their students know or do not know.

Therefore, for research to show that i-Ready and MAP are valid for standardized tests in a way that will affect student learning, i-Ready and MAP must both establish that their tests assess the state standards (i.e. have content validity) and also that their tests are highly predictive of students’ scores

on the standardized tests for each state standard. Unfortunately, none of the current research on i-Ready or MAP provides any information on content validity or the validity of these tests at the sub-item or standards level. We cannot assess if these tests are in fact useful assessments for the purpose of increasing student achievement.

The best information currently available about whether either of these formative assessments affects student learning is from a randomized control trial (RCT) that found no effect of MAP (both the test and additional teaching resources provided by the publishers of MAP) on reading achievement for 4th and 5th graders in Illinois (Cordray et al., 2013). The two-year study consisted of 32 elementary schools in 5 school districts in Illinois. Half of the schools were randomly assigned to implement MAP in 4th grade and the other half were randomly assigned to implement MAP in 5th grade. The study investigated whether the MAP program affected reading achievement after the second year of implementation. The results show that, overall, the MAP program did not have a statistically significant impact on students' achievement in either grade. The study also investigated whether the MAP resources, which consist of training, consultations, and web-based materials, were used by teachers as planned. The results show that the MAP program was implemented with moderate fidelity. Finally, the study investigated whether MAP teachers used more differentiated instruction than did non-MAP teachers, which would provide evidence that teachers were changing their instructional practices as a result of information they received on the MAP test. The results show that MAP teachers were not more likely to use differentiated instruction than teachers in the control group. Thus, this study provides no evidence that the use of MAP increases student learning.

The lack of a research base on i-Ready and MAP as means for improving student learning is both surprising and disappointing given their widespread use as well as their cost. To be clear, the negative findings of a single study should not be taken as conclusive. Rather, they illustrate how just important it is for states and districts to understand precisely what research suggests about these two tests, and where we have important, unanswered questions that deserve peer-reviewed, external research studies commensurate with the widespread use of these assessments.

Sources:

- Black, P. (2015). Formative assessment—an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22(1), 161-177.
- Cordray, D. S., Pion, G. M., Brandt, C., & Molefe, A. (2013). The Impact of the Measures of Academic Progress (MAP) Program on Student Reading Achievement. *Society for Research on Educational Effectiveness*.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.
- Educational Research Institute of America (ERIA). (2016a). *i-Ready diagnostic New York State validity study* (518). Bloomington, IN: Curriculum Associates, LLC.

- Educational Research Institute of America (ERIA). (2016b). *i-Ready and PARCC: Findings from Independent Research Linking the i-Ready Diagnostic and 2015 PARCC Assessments*. Bloomington, IN: Curriculum Associates, LLC.
- Educational Research Institute of America (ERIA). (2016c). *i-Ready and the Smarter Balanced Assessments: Findings from Independent Research Linking the i-Ready Diagnostic and Smarter Balanced Assessments*. Bloomington, IN: Curriculum Associates, LLC.
- Educational Research Institute of America (ERIA). (2016d). *i-Ready Diagnostic Smarter Balanced Consortium (SBAC) Validity Study (519)*. Bloomington, IN: Curriculum Associates, LLC.
- Educational Research Institute of America (ERIA). (2016e). *i-Ready Diagnostic Partnership for Assessment of Readiness for College and Careers (PARCC) Validity Study (521)*. Bloomington, IN: Curriculum Associates, LLC.
- Heitink, M. C., Van der Kleij, F. M., Veldkamp, B. P., Schildkamp, K., & Kippers, W. B. (2016). A systematic review of prerequisites for implementing assessment for learning in classroom practice. *Educational research review*, 17, 50-62.
- Martone, A., & Sireci, S. G. (2009). Evaluating alignment between curriculum, assessment, and instruction. *Review of Educational Research*, 79(4), 1332-1361.
- Northwest Evaluation Association (NWEA). (2015). *Linking the Smarter Balanced Assessments to NWEA MAP Assessments*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016a). *Linking the Alaska AMP Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016b). *Linking the Arizona AzMERIT Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016c). *Linking the Georgia Milestones Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016d). *Linking the Kansas KAP Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016e). *Linking the Kentucky K-PREP Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016f). *Linking the Minnesota MCA-III Assessments to NWEA MAP Tests*. Portland, OR: NWEA.
- Northwest Evaluation Association (NWEA). (2016g). *Linking the Nebraska NeSA Assessments to NWEA MAP Tests*. Portland, OR: NWEA.

- Northwest Evaluation Association (NWEA). (2016h). Linking the North Carolina EOG State Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016i). Linking the New York State NYSTP Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016j). Linking the Ohio State Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016k). Linking the Pennsylvania PSSA Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016l). Linking the Texas STAAR Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016m). Linking the Virginia SOL Assessments to NWEA MAP Tests. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016n). Linking the PARCC Assessments to NWEA MAP Tests for Colorado. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016o). Linking the PARCC Assessments to NWEA MAP Tests for Illinois. *Portland, OR: NWEA.*
- Northwest Evaluation Association (NWEA). (2016p). Linking the PARCC Assessments to NWEA MAP Tests for New Mexico. *Portland, OR: NWEA.*