

# INTERIM ASSESSMENTS IN BRIEF

---

BY JOAN HERMAN



THE CENTER ON  
STANDARDS &  
ASSESSMENT  
IMPLEMENTATION

WestEd 

Copyright © 2017 The Regents of the University of California

The work reported herein was supported by grant number #S283B050022A between the U.S. Department of Education and WestEd with a subcontract to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). The findings and opinions expressed in this publication are those of the authors and do not necessarily reflect the positions or policies of CRESST, WestEd, or the U.S. Department of Education.

## INTRODUCTION

---

The Every Student Succeeds Act (ESSA) recognizes the importance of balanced assessment systems, the need for interim measures to monitor the progress of low performing schools, and the value of testing audits to improve system efficiency and effectiveness. ESSA also offers states the flexibility to “use a series of statewide interim assessments during the course of the academic year that result in a single summative assessment score” in place of a single annual test. Interim assessments thus have a strong potential role to play in ESSA planning.

Typically initiated at the local level, interim assessments represent an intermediate point between end-of-year accountability testing and on-going classroom and formative assessment and are intended to provide school- and/or district-level educators and administrators with the data they need to support all students achieving college and career ready standards (see Herman, 2016).<sup>1</sup> Districts and schools across the country have invested heavily in these types of tests: a Stanford University study estimates that districts spend an average of \$17.40 per student to implement interim assessments, and these estimates do not include the substantial human, social, and financial capital needed to support effective use (Topol, Olson, Roeber, & Hennon, 2013).

Built on the CSAI collection on interim assessments, this brief considers six questions related to their use.<sup>2</sup>

1. What are interim assessments?
2. What is their purpose? Whom do they serve?
3. What does research say about the effectiveness of interim assessments?
4. How does a district or school decide whether to use interim assessments?
5. What criteria are important in selecting or developing interim assessments?
6. What supports should be in place to promote the effective use of interim assessments for improvement?

## 1. WHAT ARE INTERIM ASSESSMENTS?

---

A commonly accepted definition describes interim assessments as “medium scale assessments falling between formative and summative assessment that serve to (1) evaluate students’ knowledge and skills relative to a specific set of academic goals, typically within a limited time frame, and (2) are designed to inform decisions at both the classroom and beyond the classroom level, such as the school or district level” (Perie, Marion, Gong, & Wurtzel, 2007, p. 1). Crane (2008) adds that interim assessments are administered

---

<sup>1</sup> Herman (2016) is located at <http://www.csai-online.org/resources/comprehensive-standards-based-assessment-systems-supporting-learning>.

<sup>2</sup> For a collection of resources on interim assessments, visit <http://www.csai-online.org/collection/2932>.

periodically over the course of the school year, typically under the purview of the school or district, and scores are aggregated for use at multiple levels—for example, classroom, school, and district.

## 2. WHAT IS THEIR PURPOSE? WHOM DO THEY SERVE?

---

There is an overall agreement on three general purposes that interim assessments can serve: instruction and curriculum planning, evaluation (e.g., of various programs or instructional approaches), and prediction of end of year proficiency in order to identify and take action on students at risk of failure (Herman, Osmundson, & Dietel, 2010; Perie et al., 2007). Herman, Osmundson, and Dietel (2010) add communication as another important purpose of interim assessments, such as in signaling what should be taught. Most authors agree that interim assessments should serve instructional purposes. They also advise that although interim assessments may serve a variety of purposes, they must be specifically designed and evaluated relative to the specific use and users for which they are intended.

Further, although ESSA offers states the flexibility to roll scores from interim assessments into an end of year summative score to replace annual accountability testing, there are any number of challenges in doing so—for example, challenges related to securing statewide agreement on a single standardized system; security and logistical burdens; implications for common curriculum; meeting ESSA requirements for precision, validity and fairness; and time required for system consensus, development and validation (see Dadey & Gong, 2017).

The available research has focused largely on teachers' use of interim assessments for improving instruction, and these studies had generally shown that teachers use the data to identify students who need help and to reteach in areas of inferred weaknesses. However, it is also the case that the studies have shown that teachers do not get sufficient detail on specific weaknesses and need to do additional assessment—or rely on their own classroom evidence—to bridge gaps in student learning. District and school leaders also can use interim assessment results to monitor teachers' and schools' progress and to identify and provide interventions for students, teachers, or schools that appear to be struggling. Parents too may use interim assessment results to monitor their children's progress. (See, for example, Christman et al., 2009; Goertz, Oláh, & Riggan, 2009; Herman, 2016.)

## 3. WHAT DOES RESEARCH SAY ABOUT THE EFFECTIVENESS OF INTERIM ASSESSMENTS?

Despite the popularity of interim assessments in current district practice, available evidence does not document a strong positive effect on student achievement (Cordray, Pion, Brandt, & Molefe, 2012; Konstantopoulos, Miller, van der Ploeg, & Li, 2016; Konstantopoulos, Miller, & van der Ploeg, 2013). In a lone study showing strong positive effects (Carlson, Borman, & Robinson, 2011), interim assessments were part of a data driven reform effort that provided support at the teacher, school, and district levels.

---

Moreover, there is fledgling evidence to suggest that interim assessments may be more effective for improving the performance of low ability, elementary school students (Konstantopoulos, Li, Miller, & van der Ploeg, 2016).

The most consistent, positive evidence about the effect of interim testing comes from laboratory learning research, which shows that interim testing improves the subsequent learning of initially tested, textual material and that testing of prior material improves the learning of subsequent material (Wissman, Rawson, & Pyc, 2011).

## **4. HOW DOES A DISTRICT OR SCHOOL DECIDE WHETHER TO USE INTERIM ASSESSMENTS?**

---

Districts and schools should consider a number of questions before they make the decision to purchase and use interim assessments (Perie et al., 2007):

- What do we want to learn from this assessment? What purpose will it serve?
- Who will use the information gathered from this assessment?
- What action steps will be taken as a result of this assessment?
- What professional development or support structures need to be in place to ensure the action steps are taken and are successful?
- How will student learning improve as a result of using this interim assessment? Will the interim assessment improve student performance more than an alternative option or investment?

In large part, the questions suggest that districts, in collaboration with a range of stakeholders, ought to develop a theory of action on how the use of interim assessments is expected to improve student learning—for example, who will get what kinds of student reports (individual level, class/teacher level, grade level), and how the results in these reports will be used to accelerate learning for students who are struggling, provide extra help for teachers who need it, or improve or initiate special programs or interventions. Then districts can select or develop interim assessments and create implementation plans that support the theory of action. Also among key considerations in the decision to use interim assessments is whether the district or school has sufficient organizational, technical, and financial capacity to support the system and has, or can mount, necessary infrastructure, including educator professional development, to sustain success. Without them, interim assessments are unlikely to be successful.

## 5. WHAT CRITERIA ARE IMPORTANT IN SELECTING OR DEVELOPING INTERIM ASSESSMENTS?

---

Herman and Baker (2005) discuss six criteria that can be used to evaluate the quality of interim assessments: alignment, diagnostic value, fairness, technical quality, utility, and feasibility. These are important characteristics in making purchasing decisions or in designing and developing interim assessments. Local systems and districts or schools should demand and consider evidence of each. (See also Perie et al., 2007.)

- **Alignment with purpose.** The overarching alignment question is whether a given assessment is aligned with the purpose(s) it is intended to serve by providing users the data they need to take anticipated action (see theory of action above). For example, if the assessment is intended to serve instruction or curriculum planning purposes, will it provide technically adequate data at the needed grain size to take action? If it will be used to evaluate the comparative effectiveness of different programs or instructional strategies, does the content and sequence of the assessments match that of the programs or strategies being compared? If the assessment's primary purpose is to identify struggling students who are at risk of not reaching proficiency at the end of year, then is there evidence that assessment results predict end of year scores and/or are highly related to other indicators of students being at risk? An assessment's alignment with its intended purpose permeates consideration of all the other criteria as well.
- **Alignment with district and/or school standards and learning goals.** Close alignment between an assessment and standards is the *sin qua non* in the use of assessment to support learning. If an assessment does not target what students are expected to learn, then results are not very useful in informing policy, programming, or instructional decision-making to improve student achievement. Yet the accomplishment and evaluation of such alignment is easier said than done. In a nutshell, the assessment must measure the full range of content and reflect the full range of cognitive complexity and depth of knowledge reflected in the standards, that is, be as intellectually rigorous in content applications.<sup>3</sup> The need for such rigor is one reason why some authors advise that interim assessments should not be composed of only multiple choice items (Perie et al., 2007).

Well documented strategies exist for evaluating the alignment between state/local standards and assessment.<sup>4</sup> These strategies typically also evaluate the quality of the items, another essential characteristic, as well as their link to specific content standards. It is important that alignment be independently verified, rather than simply accepting vendor evaluations.

There are obvious tensions in achieving the full alignment between standards and assessment, in

---

<sup>3</sup> See US Department of Education (2007), *Standards and Assessment Peer Review Guidance* at <https://www2.ed.gov/policy/elsec/guid/saaprguidance.pdf>.

<sup>4</sup> See, for example, Web Alignment Tool at <http://wat.wceruw.org/index.aspx> or the National Center for Assessment's Assessment Quality Evaluation Methodology at <http://www.nciea.org/aqem-resources/>.

there is a reciprocal relationship between the range of content standards that can be represented and the number of assessment items that are required. Similarly there typically is a reciprocal relationship between items' intellectual depth and student response time—items requiring deeper understanding typically take more time. Thus greater breadth and depth of coverage requires more testing time.

- **Alignment with curriculum sequence.** Interim assessments are most useful for instruction when the content of the assessment is tailored to the sequence of classroom curriculum—addressing either standards and learning objectives that have already been covered to identify gaps in learning and/or standards and learning objectives that are upcoming. This alignment issue also means that interim assessments work best when there is uniform curriculum sequence and pacing across the district or school.
- **High quality items.** High quality assessment items are clearly aligned with important standards and learning objectives, address important and accurate content, and are clearly written to be easy to understand. For constructed response items, scoring rubrics reflect important aspects of knowledge and/or skill. Optimally, the item and their context are engaging for students.
- **Diagnostic value.** Interim assessments are useful for instructional improvement to the extent they provide users with actionable, diagnostic data about students, curriculum and/or program strengths and weaknesses, and identify sources of student difficulty. Most typically, diagnostic information is derived from students' subscale scores, which describe how students perform on clusters of related items. For example, an elementary math test may provide an overall score and subscale scores for various content strands such as number, operations, algebraic thinking, or in rare cases, may even provide standard by standard results. However, the more detailed the diagnostic information, the longer the test time, because creating a reliable subscale requires that students respond to a number of items in each area reported.

Teachers may feel they can discover important diagnostic information from the specific items that students miss and/or from their wrong answer choices. By analyzing these items, teachers may find hints about the source of student difficulties that can be confirmed (or not) by other information they have about student learning (e.g., classroom assessment, observations). But the fact of the matter is that students' responses to one or two items does not provide reliable information.

Teachers also can glean diagnostic information from students' responses to constructed-response and performance assessment tasks to the extent that the responses provide a window into students' thinking and/or possible misunderstandings. For this reason, experts recommend that for interim assessments to serve instructional purposes, the items should provide for "qualitative insights about understandings and misconceptions, not just a numerical score" (Perie et al., 2007).

Some vendors claim that their test scores—actually the scale scores—indicate where students lie on

an overall skill progression and thus can be used to pinpoint what skills students have mastered and which they have yet to attain. However, savvy districts and schools would be wise to ask for the evidence supporting such claims.

- **Fairness.** A fair assessment provides all students the opportunity to show what they know, has the same meaning for all students, and does not contain obstacles or insensitive representations that may confound some students' ability to respond. Qualitative reviews are typically conducted of test items to assure that they do not contain stereotypes or negative images of some groups, do not contain contexts that may be differentially familiar to some students (e.g., an item about a regatta), or language demands that are not relevant to the content being assessed (e.g., a math word problem that uses unnecessarily complex constructions and vocabulary).

There are also a variety of empirical analyses that can be conducted to evaluate fairness—for example, DIF analysis, tests of comparability of predictive relationships between interim test scores, and end of year proficiency levels for various subgroups.

The availability of appropriate accommodations for students who need them, such as English learners and students with disabilities, is also a fairness issue. Optimally, the same accommodations should be available for district or school interim tests that are available for end-of-year state tests and should match the accommodations available in instruction.

- **Technical quality.** A variety of indicators of technical quality should be available in technical manuals for purchased tests and computed for locally developed ones. Reliability coefficients reveal the consistency or coherence of the scores; the higher the coefficient, the more the score represents a stable attribute and the less error in it. Conversely, the lower the reliability, the more noise in the score – and the less it represents a stable attribute. Reliability coefficients of .9 and above are common for end-of-year tests and .8 is usually considered a minimum. Reliability is a concern not only for an interim test's total score but also for any subscale scores, which often presents a problem for reporting at more detailed, diagnostic scores, that is, the scores are not sufficiently reliable to provide sound information. Similarly, teachers often want to see the specific items that students did well or poorly on so that they can infer skill strengths or weaknesses. However, student responses to single items do not provide reliable information. Available evidence suggests that at least 5-10 items are necessary for a stable report.

Similarly, indicators of errors of measurement provide an estimate of the precision of the scores—or how much error is contained in the score. The concept is based on the idea that a student's score on any test, called the observed score, is an estimate of the student's true score, which is what a student hypothetically would score if the measurement was perfect. Obviously, the lower the error the better. The standard error is used to compute a confidence interval, which shows the range over which a student's true score actually lies, based on the observed score. Experts believe that any interpretation or use of scores should take account of the confidence interval. Sometimes the

confidence interval for a score can cross proficiency levels, so that a student may score proficient based on the observed score, yet based on the confidence interval, could also fall into the non-proficient category, and thus subsequent action should be based on the student's borderline status.

- **Utility.** A useful assessment provides users with the usable information they need to serve the purpose of the assessment. Assuming the assessment is well aligned with its intended purpose (see the first criterion on alignment with purpose), this utility criterion largely focuses on the comprehensibility and usability of the assessment reports: Are they timely? Do they provide the level of detail that users need? Are they easy to understand? Are they actionable?

Some assessment systems are linked to instructional resources, such that they recommend specific resources that will be helpful to students based on their scores. Savvy districts and schools will ask for evidence of how such links are made and evidence of the effectiveness of the resources.

- **Feasibility.** A feasible assessment is one which can be purchased, administered, scored, analyzed, and used within available local constraints—financial, technological, operational, and human capacity. It is worth its cost in dollars, time, and effort.

## 6. WHAT SUPPORTS SHOULD BE IN PLACE TO PROMOTE THE EFFECTIVE USE OF INTERIM ASSESSMENTS FOR IMPROVEMENT?

---

The guidance below draws on both research specific to the use of interim assessments and the more general data use literature.<sup>5</sup> The bottom line from the research is that investment in interim assessments is not likely to bring substantial benefits for student learning unless there is concomitant investment in the human and organizational capital needed to well implement and use the assessments. It is useful to think about both the district and school conditions that support effective use and the processes and actions that may enable those conditions. Most all are overlapping and inter-related.

### Antecedent Conditions

- **Teacher understanding and trust.** Marshall (2006) puts this as the number one antecedent to establish a foundation for successful use of interim assessments and lays out seven steps to reducing teacher anxiety and resistance and establishing a productive, “no blame” environment for honest analysis and continuous improvement.
- **District/school culture encouraging data use.** The norms and values of the district and school support expectations for evidence based decision-making and commitment to continuous improvement at all levels (see for example, Hamilton et al., 2009; Wohlstetter, Datnow, & Park, 2008). The district and/or school has a well-articulated vision and theory of action that is linked to

---

<sup>5</sup> For a collection of resources on data use, visit <http://www.csai-online.org/collection/2307>.

broader reform programs and initiatives (Crane, 2010; Hamilton et al., 2009). There are incentives for data use and continuous improvement (Coburn & Turner, 2011; Wohlstetter, Datnow, & Park, 2008).

- **Leadership support.** It has long been known that leadership support through both exhortation and participation are key to implementing any new practice. For example, Research for Action (Christman et al., 2009) highlights the role of principal engagement in the effectiveness of interim assessments. (See also Lachat & Smith, 2005, and Coburn & Turner, 2011.)
- **Teacher knowledge and content expertise.** Teachers' assessment literacy and their knowledge of content and how learning develops are critical for the productive use of data for improvement (see for example, Coburn & Turner, 2011). Data coaches and full time teacher leaders in math and reading also were found helpful for supporting effective use (Christman et al., 2009).
- **Collective responsibility.** Part of a school culture encouraging data use and continuous improve is educators having collective responsibility for their students' learning (Christman et al., 2009).
- **User friendly data systems.** To make use of data, educators need easy access to accurate data, through systems that store data and have the ability to the merge, disaggregate, and otherwise analyze multiple sources of data and provide user-friendly results. (For example, see Coburn & Turner, 2011; Lachat & Smith, 2005). The rapid turnaround of results also is stressed by a number of authors.
- **Clear, agreed upon standards and grade-by-grade learning expectations.** These are the foundation for common assessments across each grade and course and for collective responsibility, analysis, and improvement (see, for example, Marshall, 2006).
- **High quality tests.** High quality tests are well aligned with district/school learning expectations in both content and cognitive demand and provide accurate results (see prior section).

### Processes and Supports for Foundational Conditions and Productive Use

- **Involve multiple stakeholders.** Multiple stakeholders, particularly school level educators, should be involved in all assessment selection, policies, and practices of use. If teachers and school leaders are involved in decision-making, they are more likely to understand and "own" the decision (Crane, 2010).
- **Align and clarify, as needed, learning goals, curriculum, and assessment.** If interim assessments are intended to not only stimulate conversation and action about improving learning but also to provide important information for doing so, learning goals, curriculum, and assessment must be closely and meaningfully aligned and common within and across schools (Abrams & McMillan, 2013; Marshall, 2006).

- **Support teacher content and assessment literacy capacity with professional development and on-site expertise.** Data coaches and content experts (ELA and math) can help teachers develop the assessment literacy and content-pedagogical knowledge they need to use data and take action to improve students' learning (e.g., see Christman et al., 2009; Wohlstetter, Datnow, & Park, 2008). Wohlstetter and colleagues further note that schools vary in their data use competency, and commitment and implementation plans should be differentiated accordingly.
- **Establish conditions and routines for the collective analysis of data and action planning based on the data.** Nearly all authors note the importance of the collaborative use of data. Schools need to create the time for teachers to come together around the data and have specific routines, protocols and skilled facilitation to guide data interactions. Coburn and Turner (2011) note that routines specify who comes together, when they come together, and over what data to support action-oriented data conversations. Research for Action (Christman et al., 2009) lays out a number of characteristics of effective grade group meetings, including the involvement of the principal and teacher leaders.
- **Involve students and parents in the data use.** Students need to take responsibility for their own learning and the value of parent involvement is well established. (See, for example, Hamilton et al., 2009; Marshall, 2006; Stiggins & Chappuis, 2013.)
- **Follow-up, evaluate, and improve.** Just as data use is aimed at continuous improvement, so, too, should the use of interim assessments be subject to evaluation and its implementation and consequences monitored (Crane, 2010). Further, as Research for Action (Christman et al., 2009) observes that if practitioners' sense-making does not lead them to seek and develop new and robust instructional interventions, if these interventions are not actually implemented or not implemented well, or if their effectiveness is not assessed, then teaching and learning is not likely to improve.

## CONCLUSION

---

Interim assessments can be an important strategy for improving student learning, but achieving that goal requires the sensitive orchestration of a number of factors. Among these are an understanding of the potential purpose(s) of interim assessments and of the role that they are intended to play in a given district and/or school, a thoughtful decision about whether they are the right strategy for the setting, and a careful selection process. Purchasing interim assessments and even professional development for them is only part of what is needed to assure their effective use. The human, social, and technological infrastructure needed for success is substantial as is the close involvement and commitment of all stakeholders who are intended users.

## REFERENCES

---

- Abrams, L., & McMillan, J. (2013). The instructional influence of interim assessments: Voices from the Field. In R. W. Lissitz (Ed.), *Informing the practice of using formative and interim assessment: A systems approach*. Charlotte, NC: Information Age Publishing, Inc.
- Topol, B., Olson, J., Roeber, E., & Hennon, P. (2013). *Getting to higher-quality assessments: Evaluating costs, benefits, and investment strategies*. Palo Alto, CA: Stanford Center for Opportunity Policy in Education (SCOPE). Available at <https://edpolicy.stanford.edu/sites/default/files/publications/getting-higher-quality-assessments-evaluating-costs-benefits-and-investment-strategies.pdf>
- Carlson, D., Borman, G. D., & Robinson, M. (2011). A multi-state district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis*, 33(3), 378-398. Retrieved from <http://journals.sagepub.com/doi/abs/10.3102/0002831212466909>
- Christman, J. B., Neild, R. C., Bulkley, K., Blanc, S., Liu, R., Mitchell, C., & Travers, E. (2009). *Making the most of interim assessment data. Lessons from Philadelphia*. Philadelphia, PA: Research for Action. Retrieved from [http://8rri53pm0cs22jk3vvqna1ub-wpengine.netdna-ssl.com/wp-content/uploads/2015/10/Making\\_the\\_Most\\_of\\_Interim\\_Assessment\\_Data.pdf](http://8rri53pm0cs22jk3vvqna1ub-wpengine.netdna-ssl.com/wp-content/uploads/2015/10/Making_the_Most_of_Interim_Assessment_Data.pdf)
- Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, 9(4), 173-206.
- Cordray, D., Pion, G., Brandt, C., Molefe, A., & Toby, M. (2012). *The impact of the Measures of Academic Progress (MAP) Program on student reading achievement*. (NCEE 2013-4000). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from [http://www.ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL\\_20134000.pdf](http://www.ies.ed.gov/ncee/edlabs/regions/midwest/pdf/REL_20134000.pdf)
- Crane, E. (2008). *Interim assessments practices and avenues for state involvement: TILSA SCASS interim assessment subcommittee*. Washington, DC: Council of Chief State School Officers. Retrieved from [http://www.ccsso.org/Documents/2008/Interim\\_Assessment\\_Practices\\_and\\_2008.pdf](http://www.ccsso.org/Documents/2008/Interim_Assessment_Practices_and_2008.pdf)
- Crane, E. (2010). *Building an interim assessment system: A workbook for school districts*. Washington, DC: Council of Chief State School Officers. Retrieved from [http://www.ccsso.org/Documents/2010/Building\\_an\\_Interim\\_Assessment\\_2010.pdf](http://www.ccsso.org/Documents/2010/Building_an_Interim_Assessment_2010.pdf)
- Goertz, M. E., Oláh, L. N., & Riggan, M. (2009). *From testing to teaching: The use of interim assessments in classroom instruction*. Retrieved from <http://www.cpre.org/testing-teaching-use-interim-assessments-classroom-instruction-0>
- Gong, B., & Dadey, N. (2016). *Using interim assessments in place of summative assessments? Consideration of an ESSA Option*. Washington, DC: Council of Chief State School Officers. Retrieved from <http://www.ccsso.org/Documents/ASR%20ESSA%20Interim%20Considerations-April.pdf>
- Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J. & Wayman, J. (2009). *Using student achievement data to support instructional improvement. IES Practice Guide*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from [https://www.ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm\\_pg\\_092909.pdf](https://www.ies.ed.gov/ncee/wwc/Docs/PracticeGuide/dddm_pg_092909.pdf)

- Herman, J. (2016). Comprehensive standards-based assessment systems supporting learning. Los Angeles, CA: The Regents of the University of California. Retrieved from [http://www.csai-online.org/sites/default/files/resources/4666/CAS\\_SupportingLearning.pdf](http://www.csai-online.org/sites/default/files/resources/4666/CAS_SupportingLearning.pdf)
- Herman, J., & Baker, E. (2005). Making benchmark testing work. *Educational Leadership*, 63(3), 48-54.
- Herman, J. L., Osmundson, E., & Dietel, R. (2010). *Benchmark assessment for improved learning*. (AACC Policy Brief). Los Angeles, CA: University of California. Retrieved from [http://www.csai-online.org/sites/default/files/resource/imported/R1\\_benchmark\\_polbrief\\_Herman.indd.pdf](http://www.csai-online.org/sites/default/files/resource/imported/R1_benchmark_polbrief_Herman.indd.pdf)
- Konstantopoulos, S., Li, W., Miller, S. R., & van der Ploeg, A. (2016). Effects of interim assessments across the achievement distribution: Evidence from an experiment. *Educational and Psychological Measurement*, 76(4), 587-608.
- Konstantopoulos, S., Miller, S., & van der Ploeg, A. (2013). The impact of Indiana's system of interim assessments on mathematics and reading achievement. *Educational Evaluation and Policy Analysis*, 35, 481-499.
- Konstantopoulos, S., Miller, S. R., van der Ploeg, A., & Li, W. (2016). Effects of interim assessments on student achievement: Evidence from a large-scale experiment. *Journal of Research on Educational Effectiveness*, 9(sup1), 188-208.
- Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk*, 10(3), 333-349.
- Marshall, K. (2006). *Interim assessments: Keys to successful implementation*. Retrieved from <https://marshallmemo.com/articles/Interim%20Assmt%20Report%20Apr.%2012,%2006.pdf>
- Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). *The role of interim assessments in a comprehensive assessment system: A policy brief*. Washington, DC: Aspen Institute. Retrieved from <http://www.achieve.org/files/TheRoleofInterimAssessments.pdf>
- Stiggins, R., & Chappuis, S. (2013). Productive formative assessment always requires local district preparation. In R. W. Lissitz (Ed.), *Informing the practice of using formative and interim assessment: A systems approach* (pp. 237-247). Charlotte, NC: Information Age Publishing, Inc.
- Wissman, K.T., Rawson, K.A., & Pyc, M.A. (2011). The interim test effect: Testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review*, 18(6), 1140-1147.
- Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement*, 19(3), 239-259. Retrieved from <http://www.tandfonline.com/doi/full/10.1080/09243450802246376?scroll=top&needAccess=true>