



Using Student Surveys as a Measure of Teaching Effectiveness

David English, Jackie Burniske,
Debra Meibaum, and Lisa Lachlan-Haché
American Institutes for Research

JUNE 2016

Using Student Surveys as a Measure of Teaching Effectiveness

June 2016

David English, Jackie Burniske,
Debra Meibaum, and Lisa Lachlan-Haché
American Institutes for Research



1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Copyright © 2016 American Institutes for Research. All rights reserved.

Contents

	Page
Introduction.....	1
Background.....	1
Benefits and Limitations Regarding the Use of Student Surveys in Teacher Evaluation	2
Student Surveys Approved by Ohio Department of Education.....	7
Implementation of Student Surveys.....	9
Using Student Perception Data Effectively	11
Summary.....	11
References.....	12
Additional Resources	15

Introduction

The 2016–17 budget bill for Ohio includes provisions that modify the alternative framework for teacher evaluation beginning in the 2015–16 school year. Districts may still use the original teacher evaluation framework, but for those electing to use the alternative framework, one or any combination of the following components shall comprise 15% of each teacher’s evaluation: student surveys, teacher self-evaluations, peer review evaluations, student portfolios, or a district-determined component. This document provides guidance regarding the use of student surveys, including benefits and limitations of their use, summaries of student surveys approved for use by the Ohio Department of Education, and guidance related to implementing surveys and using student perception data effectively.

Background

Students spend more time observing classroom learning conditions than any outside observer (MET Project, 2012). Solicitation of student perceptions regarding teachers has a long history in the United States. As early as 1896, students in Grades 2–8 in Sioux City, Iowa provided input on effective teacher characteristics (Follman, 1995). It therefore might not be counterintuitive that, as of 2015, at least [23 states required or encouraged the use of student perception surveys](#) as one of multiple measures of teacher performance (Center on Great Teachers and Leaders, 2015).

In reviewing 154 articles written over a 74-year time period, Aleamoni (1999) identified 16 myths regarding students’ ratings of their instructors. Important findings reinforced across multiple studies include the following:

- Student ratings demonstrate consistency across school years for the same instructor.
- Students do not automatically rate instructional skills highly for those teachers who also scored highly on constructs associated with “popularity,” such as demeanor with students.
- Student ratings are not highly correlated with grades received from respective teachers.

Follman (1995, 1992) conducted a literature review of teacher effectiveness ratings in public schools and concluded that student surveys are reliable, that students are able to discriminate between effective and ineffective teaching, and that teacher ratings by students are no more impacted by the “halo effect,” or leniency toward teachers, than ratings by adult observers.

Goe, Bell, and Little (2008), in a later review of the literature, concluded that existing studies “provide convincing evidence that student ratings of teaching are worth considering for inclusion in teacher evaluation systems” (p. 40). They cautioned, however, that “the reliability and validity of the student survey depends to some extent on the instrument used, how it is developed, how it is administered, and the level of detail it attempts to measure” (p. 40).

Recent results from the rigorous Measures of Effective Teaching (MET) project lend additional credibility to the use of student perception data; they indicate that student ratings are predictive

of achievement results and that results are reliable to the extent that the instrument is well constructed and purposeful sampling methods are used (MET Project, 2012).

Marzano and Toth (2013), in considering the results of the MET project and preceding research, resolutely concluded that student perception surveys “can and should” (p. 75) be used as one of multiple measures in teacher evaluation.

Benefits and Limitations Regarding the Use of Student Surveys in Teacher Evaluation

Research indicates that there are both benefits and limitations to using student surveys in teacher evaluations, as is true with many evaluation measures. In Table 1, findings from research on using student surveys as a measure of teacher effectiveness are grouped into four major categories: (1) use of students as raters, (2) reliability and validity of student ratings, (3) predictability of student achievement results, and (4) implementation issues. In making decisions on the use of student surveys as a component of a comprehensive teacher evaluation system, policymakers should consider both the benefits and limitations.

Table 1. Benefits and Limitations of Student Surveys

Category	Benefits	Limitations
Use of students as raters	<p>Students have extensive daily contact with teachers, resulting in unique perspectives and ratings of teacher behaviors (Follman, 1995, 1992; Peterson, Wahlquist, & Bone, 2000; Worrell & Kuterbach, 2001).</p> <p>Students are the direct recipients of instruction and have more experience with their teachers than other evaluators (Follman, 1992).</p> <p>Student ratings are consistent from year to year (Aleamoni, 1999).</p> <p>Students discriminate between effective teaching behaviors and warm, caring, supportive teacher-student interactions (Peterson et al., 2000; Aleamoni, 1999).</p> <p>Secondary students can discriminate between effective and ineffective teachers (Follman, 1992; Worrell & Kuterbach, 2001).</p> <p>Student responses distinguish between teachers; students may rate one teacher high and another low, based on the quality of teaching that the student experiences (Ferguson, 2010; MET Project, 2012).</p>	<p>Student raters have a lack of knowledge of the full range of teaching requirements and responsibilities, such as curriculum, classroom management, content knowledge, and professional responsibilities (Follman, 1995, 1992; Worrell & Kuterbach, 2001; Goe et al., 2008).</p> <p>Students will respond with candid judgments only when they can be assured that survey results are anonymous and, to that end, surveys should not contain comment boxes, per students' belief that their handwriting will be recognized (Popham, 2013).</p>

Category	Benefits	Limitations
<p>Reliability and validity of student ratings</p>	<p>Student ratings are reliable and are significant predictors of other observable teaching practice ratings (Peterson et al., 2000; Worrell & Kuterbach, 2001).</p> <p>Elementary and secondary student raters are no more impacted by validity concerns, such as halo and leniency effects, than adult raters (Follman, 1995, 1992).</p> <p>Elementary and secondary students are as reliable as older, adult raters in rating teaching behaviors (Follman, 1995, 1992; Worrell & Kuterbach, 2001).</p> <p>Elementary students, including preschoolers as young as 4 years old, can rate reliably (Follman, 1995).</p> <p>Appropriately administered, well-constructed instruments yield high-reliability results; subjective and correlational studies indicate positive validity of student rating results (Aleamoni, 1999).</p> <p>Secondary and older primary education students provide ratings of teacher behavior that are stable, reliable, valid, and predictive for teacher evaluation (den Brok, Brekelmans, & Wubbels, 2004).</p> <p>Student responses (in the Tripod survey) are reliable, valid, and stable over time at the classroom level (Ferguson, 2010).</p> <p>Student survey results (in the Tripod survey) are more likely to demonstrate consistency than classroom observations or achievement gain measures (MET Project, 2012).</p>	<p>Reliability and validity of student ratings depend on the content, construction, and administration of student rating instruments (Popham, 2013; Goe et al., 2008; Little, Goe, & Bell, 2009; Kyriakides, 2005; Aleamoni, 1999).</p> <p>The earliest age by which students can adequately rate their teachers is unresolved, and that must be considered when applying ratings by students who are below Grade 3 (Follman, 1995).</p> <p>To ensure reliability of a student survey, enough items must be included for each teaching construct measured and a representative sampling of classes and students must be surveyed (MET Project, 2012).</p> <p>Research has yielded widely inconsistent results regarding the impact of student rater demographics, student grades and other student rater characteristics on ratings of instructors (Follman, 1992; Aleamoni 1999).</p>

Category	Benefits	Limitations
<p>Predictability of student achievement results</p>	<p>Student ratings are more highly correlated with student achievement than principal ratings and teacher self-ratings (Kyriakides, 2005; Wilkerson, Manatt, Rogers, & Maughan, 2000).</p> <p>On average, teachers who receive the most favorable survey responses are those whose students demonstrate the most gains in achievement scores (MET Project, 2012).</p> <p>Student ratings of teachers align with student achievement; a teacher rated more highly by students in instructional effectiveness aligns with students achieving at higher levels in that teacher’s class (MET Project, 2010a; Crow, 2011).</p> <p>The students of teachers ranking in the top 25th percentile, based on student survey (Tripod) results, learned the equivalent of about 4.6 more months of schooling in math, over a school year, than students of teachers whose survey results were in the bottom 25th percentile (MET Project, 2012).</p>	<p>Student ratings are a moderate predictor of student achievement (Worrell & Kuterbach, 2001).</p> <p>Student ratings should not be the primary teacher evaluation instrument but, rather, should be included in a comprehensive teacher evaluation process (Goe et. al., 2008; Little et al., 2009; Peterson et al., 2000; Follman, 1995, 1992).</p> <p>For any teacher, survey data should be collected over multiple classrooms and across multiple school years before they are used for high-stakes decision making (Ramsdell, 2011).</p>

Category	Benefits	Limitations
Implementation issues	<p>Appropriate use of student ratings feedback by the teacher can result in an improved teaching and learning environment (Follman, 1995, 1992; Aleamoni, 1999).</p> <p>Student surveys present timely and specific feedback in ways that other measures, such as achievement results, do not (MET Project, 2012).</p> <p>Relevance and use can be expanded to nontested grades and subjects relatively easily (MET Project, 2012).</p> <p>Results of student ratings can be collected anonymously (Little et al., 2009; Worrell & Kuterbach, 2001).</p> <p>Student ratings require minimal training and are both cost- and time-efficient (Little et al., 2009; Worrell & Kuterbach, 2001).</p>	<p>Confidentiality concerns in regard to protecting the anonymity of student raters must be addressed (Popham, 2013; McQueen, 2001).</p> <p>Results of students' ratings may be misinterpreted and misused (Kyriakides, 2005; Aleamoni, 1999).</p> <p>Use of data by administrators for punitive purposes could result in teachers' lack of support for the student ratings (Aleamoni, 1999).</p> <p>For results to be meaningful to teachers, training should ensure that they understand each item and how their scores compare with those of other teachers (MET Project, 2012).</p>

Student Surveys Approved by Ohio Department of Education

In Ohio, a local education agency opting to use student surveys may choose from among four state-approved, off-the-shelf surveys or develop its own survey. If a district decides to adopt or develop its own survey, it is highly recommended that the district ensures the survey meets the following criteria:

- Aligns with Ohio Standards for the Teaching Profession
- Is grounded in research about teaching and designed to provide evidence of effectiveness of teachers’ practice for formative and/or summative purposes
- Meets appropriate standards of validity (consider validating the survey over a number of years before using it to drive consequences for educators)
- Meets appropriate standards of reliability

District-Created Student Perception Surveys in Davis School District, Utah

In Davis School District, Utah, the educator assessment committee created [student perception surveys](#) at the K–2, 3–6, and 7–12 grade bands. Each survey consisted of seven to 10 questions each, considerably shorter than most off-the-shelf products. Schools are required to administer surveys at least for informational purposes and also may use them toward final teacher evaluation scores. Teachers are advised to use survey results to help create their professional development plans and to inform goal setting. Results also are made available to supervisors who are instructed to incorporate results into end-of-year conferences. Prompts for the discussion of survey results are built into “preconference” forms.

Students rate teachers on a scale of “no,” “sometimes,” or “yes” (or 😞, 😐, 😊 for Grades K–2) on the following items (Hanover Research, 2013) indicated in Table 2.

Table 2. Items from Davis School District Student Perception Surveys.

Survey Item	High School	Elementary	K–2
I learn new things in this class.	X	X	
My class is a good place for learning.	X	X	X
This teacher treats me with care and respect.	X		
This is a good teacher.	X	X	X
I know what I am supposed to do in this class.	X	X	X
I understand the class rules.	X	X	
This teacher treats me fairly.	X	X	
I know how well I am doing in this class.	X		
I usually understand how to do my assignments.	X		
This teacher maintains class discipline.	X		
I like to come to this class.		X	X
I know how well I am learning in this class.		X	
My teacher is nice to me.		X	X
My teacher shows me how to do new things.			X
My teacher’s rules are fair.			X

The four state-approved, off-the-shelf surveys that districts may use are summarized below. All vendors offer online and paper-and-pencil versions of the surveys except where noted.

Tripod. During the past decade, the Tripod Project (www.tripodproject.org) has worked with more than 300,000 students in the United States, Canada, and China. Seven urban school districts used the Tripod survey instrument, developed by Harvard researcher Ron Ferguson, as part of the MET project funded by the Bill & Melinda Gates Foundation: Dallas Independent School District (Texas), Charlotte-Mecklenburg Schools (North Carolina), Denver Public Schools (Colorado), Hillsborough County Public Schools (Florida), Memphis Community Schools (Tennessee), Pittsburgh Public Schools (Pennsylvania), and New York City Public Schools (New York). Different versions of the survey instrument are available for Grades K–2, 3–5, and 6–12. Students respond to statements categorized under each of the “7 Cs”: care, control, clarify, challenge, captivate, confer, and consolidate. During multiple administrations, the survey items have been validated and refined to capture the essential elements of classroom-level teaching and learning. Both initial and recent findings of the MET project demonstrate that student perceptions as measured by the Tripod survey can be one of multiple measures that reliably contribute to a balanced view of teacher performance and effectiveness (Ferguson, 2010; MET Project, 2012). In addition to the student survey used in the MET project, Tripod includes optional parent and teacher surveys. See more details [here](#).

Tripod Survey Items as Predictors of Achievement

In 2012, the MET project found that student surveys were more predictive of student achievement scores than classroom observations (MET Project, 2012). For the Tripod survey, students indicated their level of agreement with 36 descriptive items. The following five items, covering instructional and classroom management dimensions of teaching practice, were the most closely correlated with achievement:

- Students in this class treat the teacher with respect.
- My classmates behave the way my teacher wants them to.
- Our class stays busy and doesn't waste time.
- In this class, we learn a lot almost every day.
- In this class, we learn to correct our mistakes.

(MET Project, 2012)

Survey of Teacher Practices. The Survey of Teacher Practice (STeP), distributed by My Student Survey (<http://mystudentsurvey.com/>), was developed by researchers at Vanderbilt University and was initially validated using results from approximately 12,000 students across seven school districts, as part of the Georgia Race to the Top program in spring 2011. Developer-provided data demonstrate a high level of reliability and validity (Voight & Hanson, 2012), although no independent studies of the survey are available. Versions are available for Grades 4–5 and 6–12. The instrument consists of 55–60 questions that measure teacher performance along six constructs: presenter, manager, counselor, coach, motivational speaker, and content expert. See more details for [Grades 3–5](#) and [Grades 6–12](#).

iKnowMyClass. The iKnowMyClass survey was developed by Russell Quaglia at the Quaglia Institute for Student Aspirations (QISA) and is distributed by [Sage Publications](#). Versions of the survey are offered for Grades 3–5 (27 items) and Grades 6–12 (20–50 items) and measure student perceptions along the following constructs: engagement, relevance, relationships, class efficacy, cooperative learning environment, critical thinking, positive pedagogy, and discipline problems. Validation studies have been developer-led so far; thus, no independent studies are available at the QISA website. This survey is administered online only; there is no paper-and-pencil format. See more details [here](#).

Panorama Student Survey. The Panorama Student Survey, developed in collaboration with the Harvard Graduate School of Education under the leadership of Dr. Hunter Gehlbach, measures student perceptions of teaching and learning as well as perceptions of school climate and students' own strengths and weaknesses. The Panorama website (www.panoramaed.com) notes that it is a free, open-source survey. Student perceptions about the classroom and teaching are measured in two versions, for Grades 3–5 and 6–12, along the following scales: interest in subject, teacher press, expectations and rigor, supportive relationships, student engagement, classroom environment, and pedagogical effectiveness. The survey also includes open-ended responses. The Panorama Student Survey website includes a report on its piloting methods for establishing survey validity. There appear to be no independent studies confirming its validity at this time. See more details [here](#).

Implementation of Student Surveys

For student perceptions to be meaningfully used in high-stakes teacher evaluation, they must be one of multiple measures. It is important for districts to ensure that they are measuring aspects of instruction embedded in their particular teaching model by engaging appropriate expertise in writing, testing, and implementing surveys (National Council on Teacher Quality, 2013). Including teachers in the process of survey and item selection will allow them to see how items are aligned with the district's teaching framework and help assuage concerns that student surveys are otherwise “popularity contests.” This is the case for both district-developed surveys and off-the-shelf surveys, from which a number of survey items might be chosen. Districts should consider the trade-offs between increasing the number of questions in any one teaching construct covered by the survey, in order to improve reliability, and establishing a reasonable survey completion time. Fewer questions will lead to more thoughtful ratings by students (Popham, 2013). Reliability also must be ensured by surveying an adequate number of students, including those in multiple classrooms, for applicable teachers (MET Project, 2012; Ramsdell, 2011). Accommodations also should be considered for special populations of students to ensure accurate responses (MET Project, 2012).

Gradual Implementation of Student Surveys in Pittsburgh Public Schools

As one of six school districts hosting the MET project, Pittsburgh Public Schools (PPS) began exploring the use of student surveys in 2009–10 through a gradual implementation process and in close consultation with the Pittsburgh Federation of Teachers. The process was as follows:

- 2009–10: Tripod survey piloted in 250 PPS classrooms.
- 2010–11: Tripod survey piloted for 50 teachers in PPS's Promise-Readiness Corps.
- 2011–12: Tripod survey administered districtwide for the first time, twice to students in more than 3,400 classrooms. Some teachers received results reports for the first time.
- 2012–13: 1,300 teachers received no-stakes-attached reports on student survey results. Along with the reports, teachers received recommended professional development steps according to performance level and resource listings. Principals received training on using report results to support teachers.
- 2013–14: Evaluations formally included student survey results for the first time, weighted at 15% of the overall rating. Personnel decisions were required to be based on multiple years of evaluation data. An independent study prepared for the Institute of Education Sciences in August 2014 concluded that classroom observation scores, value-added achievement measures, and Tripod scores were valid predictors of various teaching practices that were overlapping but not identical, as was PPS's intent in using multiple measures.

(Chaplin, Thompkins & Miller, 2014)

Administrators should remember that many students have never completed evaluative surveys and significant attention should be paid to the verbal and written instructions provided to student raters concerning the following areas (Popham, 2013):

- **How the ratings will be used.** Language should be included about the survey's role for both evaluation and professional development of teachers. Students will provide more honest answers if they recognize that teachers are encouraged to integrate feedback into their practice.
- **Meaning of the survey items.** Administrators of the survey should discuss the various survey items to provide clarity.
- **How to use the rating scale.** Whether it contains descriptive categories or numbers, the rating scale should be explicitly explained to students.
- **The importance of honesty.** To ensure honesty in responses, it is critical that students trust that their responses will be kept confidential. Districts should ensure that strong security protocols are in place and that these are clearly described to students. Administrators of the survey should assure students that their individual responses will not be viewed by teachers or other educators and that they will not have a negative or positive impact on their grades.

Consistently applied protocols that exclude teacher access to completed surveys should be instituted. Districts should include teachers in discussions of survey administration protocols, especially with regard to ensuring that results are paired with the proper teachers in data collection systems.

Using Student Perception Data Effectively

An important benefit of student surveys, beyond their evaluative utility, is their potential to inform professional learning for teachers. The MET project noted that student ratings yield timely and specific feedback for teachers in contexts that other evaluation measures cannot (MET Project, 2012, 2010a). Districts may consider collecting student survey data more than once during the school year to facilitate ongoing adjustments to teaching practice.

Providing a crosswalk between survey items and elements of the district’s teaching framework can increase the usefulness of results to teachers. Training for teachers should involve efforts to ensure understanding of the student survey items, the meaning of survey scores, and the actual survey administration process (MET Project, 2012). The data from the surveys can be used by districts to look across their schools, by schools to look across their classrooms, and by classroom teachers to look across their students. Results can provide rich data with which to target professional development and monitor teacher growth (Ramsdell, 2011; Crow, 2011).

Summary

Student survey instruments can provide valuable insight into the teaching and learning environment of a classroom when used as one of multiple measures of educator effectiveness. Students have the “deepest, broadest, and most veridical perception of their teacher” (Follman, 1992, p. 176) and, as such, student survey instruments can be a valuable component of a comprehensive teacher evaluation system.

References

- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education*, 13(2), 153–166.
- Barge, J. D. (2011, September). *Teacher and leader evaluation systems*. (Slide presentation). Presented at the Georgia RT3 Great Teachers and Leaders Program, Atlanta, GA.
- Center for Education Policy and Practice, Massachusetts Teachers Association. (2011). *Massachusetts Department of Elementary and Secondary Education educator evaluation regulations*. Boston, MA: Author.
- Center on Great Teachers and Leaders. (2015). *Databases on state teacher and principal evaluation policies* (Website). Washington, DC: Author. Retrieved from <http://resource.tqsource.org/stateevaldb/STEPEightKeys.aspx>
- Chaplin, D., Gill, B., Thompkins, A., & Miller, H. (2014). *Professional practice, student surveys and value-added: Multiple measures of teacher effectiveness in the Pittsburgh Public Schools*. Washington, DC: Institute of Education Sciences. Retrieved from https://ies.ed.gov/ncee/edlabs/regions/midatlantic/pdf/REL_2014024.pdf
- Crow, T. (2011). The view from the seats: Student input provides a clearer picture of what works in schools. *JSD*, 32(6), 24–30. Retrieved from <http://learningforward.org/docs/jsd-december-2011/crow326.pdf?sfvrsn=2>
- Davis School District. (2014). *Evaluate Davis: Licensed educator manual*. Retrieved from http://www.nctq.org/docs/Davis_Evaluation_Handbook_2014-2015.pdf
- Davis School District. (2009). *Educator Assessment System: Acknowledging and honoring quality performance*. Farmington, UT: Davis School District. Retrieved from <http://www.nctq.org/docs/72-08.pdf>
- den Brok, P., Brekelmans, M., & Wubbels, T. (2004). Interpersonal teacher behaviour and student outcomes. *School Effectiveness and School Improvement*, 15(3–4), 407–442.
- Ferguson, R. F. (2010). *Student perceptions of teaching effectiveness* (Discussion Brief). Cambridge, MA: National Center for Teacher Effectiveness and the Achievement Gap Initiative.
- Follman, J. (1995). Elementary public school pupil rating of teacher effectiveness. *Child Study Journal*, 25(1), 57–78.
- Follman, J. (1992). Secondary school students' ratings of teacher effectiveness. *The High School Journal*, 75(3), 168–178.

- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>
- Hanover Research. (2013). *Student perceptions surveys and teacher assessments*. Washington, DC. Retrieved from <http://dese.mo.gov/sites/default/files/Hanover-Research-Student-Surveys.pdf>
- Kyriakides, L. (2005). Drawing from teacher effectiveness research and research into teacher interpersonal behaviour to establish a teacher evaluation system: A study on the use of student ratings to evaluate teacher behavior. *Journal of Classroom Interaction*, 40(2), 44–66. Retrieved from <http://files.eric.ed.gov/fulltext/EJ768695.pdf>
- Little, O., Goe, L., & Bell, C. (2009). *A practical guide to evaluating teacher effectiveness*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://files.eric.ed.gov/fulltext/ED543776.pdf>
- Lopez, S. J., Agrawal, S., & Calderon, V. J. (2010, August). *The Gallup Student Poll technical report*. Washington, DC: Gallup, Inc. Retrieved from [http://www.gallup.com/services/177095/gallup-student-poll-technical-report.aspx?g_source=Lopez, S. J., Agrawal, S., & Calderon, V. J. \(2010&g_medium=search&g_campaign=tiles](http://www.gallup.com/services/177095/gallup-student-poll-technical-report.aspx?g_source=Lopez,%20S.%20J.,%20Agrawal,%20S.,%20&Calderon,%20V.%20J.(2010&g_medium=search&g_campaign=tiles)
- Marzano, R. J., & Toth, M. (2013). *Teacher evaluation that makes a difference*. Alexandria, VA: ASCD.
- McQueen, C. (2001). Teaching to win. *Kappa Delta Pi Record*, 38(1), 12–15.
- MET Project. (2012). *Asking students about teaching: Student perception surveys and their implementation* (Policy & Practice Brief). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/Asking_Students_Summary_Doc.pdf
- MET Project. (2010a). *Student perceptions and the MET project*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from http://k12education.gatesfoundation.org/wp-content/uploads/2015/12/Student_Perceptions_092110.pdf
- MET Project. (2010b). *Working with teachers to develop fair and reliable measures of effective teaching*. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from <https://docs.gatesfoundation.org/Documents/met-framing-paper.pdf>
- MET Project. (2010c). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. (Research paper). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved from <https://docs.gatesfoundation.org/Documents/preliminary-findings-research-paper.pdf>

- Murphy, P. K., Delli, L., & Edwards, M., (2004, Winter). The good teacher and good teaching: Comparing beliefs of second-grade students, preservice teachers, and inservice teachers. *The Journal of Experimental Education*, 72(2), 69–92.
- National Council on Teacher Quality. (2013). *Connect the dots: Using evaluations of teacher effectiveness to inform policy and practice*. Washington, DC: Author. Retrieved from http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report
- Opdenakker, M-C., Maulana, R., & den Brok, P. (2011). Teacher-student interpersonal relationships and academic motivation within one school year: Developmental changes and linkage. *School Effectiveness and School Improvement*, 23(1), 95–119.
- Peterson, K. D., Wahlquist, C., & Bone, K. (2000). Student surveys for school teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 135–153.
- Peterson, K. D., Wahlquist, C., Bone, K., Thompson, J., & Chatterton, K. (2001). Using more data sources to evaluate teachers. *Educational Leadership*, 58(5), 40–43.
- Popham, W. J. (2013). *Evaluating America's teachers: Mission possible?* Thousand Oaks, CA: Corwin.
- Ramsdell, R. (2011, December). *Enhancing teacher evaluation and feedback systems with Tripod student surveys*. (Slide presentation). Presented at the Teacher Evaluation That Works Conference, Cromwell, CT. Retrieved from http://www.azed.gov/highly-qualified-professionals/files/2012/01/tripod_ramsdell_2011-11-14_arizona_publish.pdf
- Tripod Project. (2011). *Tripod survey assessments: Multiple measures of teaching effectiveness and school quality*. Westwood, MA: Cambridge Education.
- Voight, A., & Hanson, T. (2012). *Summary of existing school climate instruments for middle school*. San Francisco, CA: REL West at WestEd. Retrieved from https://relwest.wested.org/system/resources/21/RELW_Memo_PUBLIC_SC3-3-1_20120918.508.pdf?1351537047
- Wilkerson, D. J., Manatt, R. P., Rogers, M. A., & Maughan, R. (2000). Validation of student, principal, and self-ratings in 360 degree feedback for teacher evaluation. *Journal of Personnel Evaluation in Education*, 14(2), 179–192.
- Worrell, F. C., & Kuterbach, L. D. (2001). The use of student ratings of teacher behaviors with academically talented high school students. *Journal of Secondary Gifted Education*, 12(4), 236–147.
- Wubbels, T., Levy, J., & Brekelmans, M. (1997). Paying attention to relationships. *Educational Leadership*, 54(7), 82–86. Retrieved from http://www.ascd.org/publications/educational-leadership/apr97/vol54/num07/Paying-Attention-to-Relationships.aspx?sm_au=iVVZQHLNFJrj3NMQ

Wubbels, T., & Levy, J. (Eds.). (1993). *Do you know what you look like? Interpersonal relations in education*. London, UK: Falmer Press.

Yazzie-Mintz, E. (2010). *Charting the path from engagement to achievement: A report on the 2009 High School Survey of Student Engagement*. Bloomington, IN: Center for Evaluation and Education Policy.

Additional Resources

The reader may find the following additional resources to be of value in examining the evaluation of teacher effectiveness.

Gallagher, C., Rabinowitz, S., & Yeagley, P. (2011). *Key considerations when measuring teacher effectiveness: A framework for validation teachers' professional practices*. (AACC Report). San Francisco and Los Angeles, CA: Assessment and Accountability Comprehensive Center. Retrieved from http://www.michigan.gov/documents/mde/AACC_report_Key_Considerations_When_Measuring_Teacher_Effectiveness_352994_7.pdf

Goe, L., Holdeheide, L., & Miller, T. (2014). *A practical guide to designing comprehensive teacher evaluation systems: A tool to assist in the development of teacher evaluation systems* (Rev. ed.). Washington, DC: Center on Great Teachers and Leaders. Retrieved from <http://www.gtlcenter.org/sites/default/files/docs/practicalGuideEvalSystems.pdf>

Holdheide, L. R., Goe, L., Croft, A., & Reschly, D. J. (2010). *Challenges in evaluating special education teachers and English language learner specialists* (Research & Policy Brief). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from <http://www.gtlcenter.org/sites/default/files/docs/July2010Brief.pdf>

Learning Point Associates. (2010 May). *Evaluating teacher effectiveness: Emerging trends reflected in the State Phase 1 Race To The Top applications*. Naperville, IL: Author. Retrieved from <https://www.wested.org/wp-content/uploads/lpa-evaluating-effectiveness.pdf>

Memphis City Schools. (2011). *Memphis Teacher Effectiveness Initiative*. Memphis TN: Author. Retrieved from http://nctq/docs/TEI_Overview.pdf

Potemski, A., Baral, M., & Meyer C. (2011). *Alternative measures of teacher performance*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://www.gtlcenter.org/sites/default/files/docs/TQ_Policy-to-PracticeBriefAlternativeMeasures.pdf

ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW
Washington, DC 20007-3835
202.403.5000

www.air.org

Making Research Relevant

LOCATIONS

Domestic

Washington, D.C.
Atlanta, GA
Austin, TX
Baltimore, MD
Cayce, SC
Chapel Hill, NC
Chicago, IL
Columbus, OH
Frederick, MD
Honolulu, HI
Indianapolis, IN
Metairie, LA
Naperville, IL
New York, NY
Rockville, MD
Sacramento, CA
San Mateo, CA
Waltham, MA

International

Egypt
Honduras
Ivory Coast
Kyrgyzstan
Liberia
Tajikistan
Zambia