



## REVIEW OF *FIXING CLASSROOM OBSERVATIONS*

*Reviewed By*

Jennie Whitcomb

University of Colorado Boulder

January 2014

### Summary of Review

*Fixing Classroom Observations: How Common Core Will Change the Way We Look at Teaching* is an advocacy document. It asserts that current classroom observation rubrics are not aligned with Common Core standards and have too many cumbersome criteria; thus, observers are overloaded, give too many high ratings, and seldom give productive feedback. To remedy these problems, the report proposes two “must-have” changes to observation rubrics: (1) pay more attention to lesson content; and (2) pare observation rubrics down to make them more focused and clear. These “must haves” may or may not address some problems of classroom observations, but there is good reason to conclude that they won’t provide much benefit. The report includes no research-informed argument to support its claim that new observation rubrics improve implementation of new teacher evaluation systems by fixing inadequate observer training, insufficient monitoring of rater calibration, and lack of time or low skills in providing instructional feedback. Tools that help observers focus on lesson content may guide substantive improvements, but the report does not offer a strong rationale for doing so. Streamlined instruments and curriculum orientation may also hold some promise, but are unlikely to seriously address core problems surrounding teacher evaluations.

**Kevin Welner**

*Project Director*

**William Mathis**

*Managing Director*

**Erik Gunn**

*Managing Editor*

---

**National Education Policy Center**

School of Education, University of Colorado  
Boulder, CO 80309-0249  
Telephone: (802) 383-0058

Email: [NEPC@colorado.edu](mailto:NEPC@colorado.edu)  
<http://nepc.colorado.edu>

Publishing Director: Alex Molnar



*This is one of a series of Think Twice think tank reviews made possible in part by funding from the Great Lakes Center for Education Research and Practice. It is also available at <http://greatlakescenter.org>. The material as long as NEPC and its author(s) are credited as the source. For inquiries about commercial use, please contact NEPC at [nepc@colorado.edu](mailto:nepc@colorado.edu).*

# REVIEW OF *FIXING CLASSROOM OBSERVATIONS*

*Jennie Whitcomb, University of Colorado Boulder*

---

## I. Introduction

Many states and school districts are in the midst of implementing two extensive reform mandates: Common Core and new teacher evaluation systems. Common Core standards have been promoted and adopted on the grounds that clearly framed high expectations will raise the level of rigor in US schools.<sup>1</sup> The overhaul of teacher evaluation systems, driven largely by federal Race to the Top competitive-grant criteria, is predicated on the notion that teacher appraisal needs to be more systematic and rigorous, and that it needs to explicitly take student performance into account.<sup>2</sup> The fundamental assumption driving both efforts is that student learning, as measured on standardized test scores, is the most important indicator of school and teacher effectiveness.

In the report *Fixing Classroom Observations: How Common Core Will Change the Way We Look at Teaching*, TNTP extends its recent efforts to reshape teacher evaluation systems.<sup>3</sup> By partnering with Student Achievement Partners,<sup>4</sup> a non-profit dedicated to providing supports for schools and teachers to implement the Common Core, TNTP seeks to influence the next generation of classroom observation rubrics by creating new observation criteria focused on whether Common Core content is being taught and by streamlining observation rubrics to make them easier for observers to use.

Both Common Core standards and new teacher evaluation systems exert pressure to standardize teaching practice in particular ways. This review does not take up whether the Common Core is the right or best set of content standards for US children and youth, but acknowledges this is contested terrain.<sup>5</sup> With regard to improving classroom observation tools, the report highlights several documented challenges those implementing systematic, high-quality classroom observations must address.<sup>6</sup> Its two recommendations have potential to improve implementation of Common Core and teacher evaluation systems, but there is no research evidence that the strategies suggested will “fix classroom observations.” Improving observer training, monitoring calibration, and providing necessary time in school schedules to both conduct adequate observations and provide sound feedback may hold greater promise for improving teachers’ instruction and students’ learning experiences.

## II. Findings and Conclusions of the Report

*Fixing Classroom Observations* posits that many states and school districts are rolling out new teacher evaluation systems that will not succeed in improving instruction because they have not been updated to reflect the Common Core. Moreover, the report asserts many observation systems have not resolved two practical challenges of classroom observations: (1) principals or observers struggle to give high-quality, specific feedback and (2) too many teachers receive high ratings. The report argues these problems with observation systems threaten the success of the Common Core. Thus, the report advocates updating observation tools to pay greater attention to what content is taught and to focus on a smaller number of observable teaching practices.

In making a case for “evolving” classroom observations, the report argues there are four problems with current classroom observation tools. First, **observers are asked to do too much** (p. 2); according to the report, some observation frameworks ask observers to rate teachers on aspects of their practice not observable in a classroom, such as a teacher’s professionalism. Second, they contend **rubrics are too complex** (p.2). That is, rubrics have too many criteria, detailed gradations, or both in performance for an observer to keep in mind. Third, **ratings are often inflated and inaccurate** (p.2); citing recent news coverage of several states’ initial implementation of teacher evaluation systems, the report claims overall teacher evaluation ratings are too high and that classroom observations contribute to these inflated ratings. Fourth, **observations don’t focus enough on feedback** (p.4). Relying on responses from a survey of “America’s Best Teachers” conducted by the New Teacher Project,<sup>7</sup> the report claims teachers do not receive constructive feedback to guide improved practice and that there is little accountability within districts to provide quality feedback.

As a response to the problems of classroom observations, the report offers two “must have” changes to bring focus and coherence to classroom observation tools. The report claims these two “must have” changes will improve the fairness of ratings, enhance the quality of feedback teachers receive, make using the rubrics less burdensome for principals, and give state and district leaders more accurate pictures of teacher performance.

**Must Have No. 1: Assess What’s Being Taught, Not Just How It’s Taught.** The report argues that current instruments focus primarily on instructional delivery and student engagement, resulting in insufficient attention to what content is taught. The report suggests the next generation of observation tools should rate teachers on whether they choose the “right content to teach” (p. 4). It advocates giving more weight in an observation tool to whether the teacher asks students to engage in grade-appropriate intellectual work articulated in Common Core standards. To guide observers, the report promotes new reference materials currently in development by TNTP that will provide anchors of lesson tasks, assessments, and sample student work so observers may rate the alignment of a lesson’s content along with instructional delivery.

**Must Have No. 2: Put Observation Rubrics on a Diet.** The report promotes pruning observation rubrics, thereby making them less comprehensive and more focused on “a

small number of essential components of a successful lesson” (p. 6). Recommending 5 to 10 essential items, the report suggests collapsing or combining items that covary—items that tend to cluster together and therefore may be measuring the same aspect of practice. The report also advocates for greater precision in rubrics, thus allowing observers to find evidence for ratings in an efficient manner.

### III. Report’s Rationale for Its Findings and Conclusions

The report’s critique of current teacher observation tools and recommendations for improvement are not grounded in a theoretical or conceptual framework. To support its claims for why classroom observation rubrics need to change, the report relies on prior TNTP reports (e.g., *Widget Effect*, *Teacher Evaluation 2.0*, *Perspectives of Irreplaceable Teachers*), anecdotal examples, and articles in the popular press (e.g., news coverage about teacher ratings in Florida, Michigan, and Tennessee). The only empirical study cited is the Bill & Melinda Gates Foundation 2012 research report *Gathering Feedback for Teaching*, which reported on the Measures of Effective Teaching (MET) project’s large-scale comparison of multiple classroom observation instruments.<sup>8</sup> In sum, with the notable exception of the MET study, the report relies primarily on conventional wisdom to support its claims regarding both the problems with observation rubrics and recommended changes.

### IV. Report’s Use of Research Literature

The report does not anchor its recommendations in the research literature. The only research study cited is the Bill & Melinda Gates Foundation 2012 report *Gathering Feedback for Teaching*.<sup>9</sup> While *Gathering Feedback* is a significant study of observation tools, the report ignores all other research on the design of research-based classroom observation tools, including peer-reviewed publications related to the MET study.<sup>10</sup> It ignores the role of human judgment in making accurate classroom observations that distinguish strong from weak teaching.<sup>11</sup> In addition, a central claim of the study is that quality feedback matters in teacher’s professional development. As important as this may be, no research is invoked on clinical or instructional feedback and its relationship to improved practice.

### V. Review of the Report’s Methods

*Fixing Classroom Observations* is not a research study. TNTP labels this report an “Issues Analysis Report.” As such it is an advocacy document crafted to make a case for TNTP’s soon-to-be-released prototype rubric for classroom observations. The report relies primarily on conventional wisdom or anecdotal examples to persuade readers that new

tools are needed to support Common Core implementation. New tools may be helpful, but better implementation of extant tools and of teacher evaluation systems might be a more promising path.

## VI. Review of the Validity of the Findings and Conclusions

The report does not build a research-informed case for its conclusions; as such, the validity of the report's conclusions is questionable.

### The Report's Analysis of the Problems with Classroom Observation Tools

The problems the report describes are an admixture of limitations of current observation rubrics as well as problems that result from rushed or poorly executed implementation of more demanding teacher evaluation systems. Many researchers familiar with developing and implementing observation rubrics and teacher evaluation systems recognize and have written about these issues.<sup>12</sup> The recommendation for new observation tools does not necessarily follow from the problems identified.

The report tends to rely on broad generalization followed by anecdote. For example, when describing the problem of observers being asked to rate too much in a classroom visit, the report supports the claim with the assertion that observers are asked to rate off-stage teaching practice, such as communication with parents or professionalism, on each classroom visit. No empirical evidence is provided to suggest how common or widespread this practice is, which runs counter to practical advice on how to build a sound teacher evaluation system.<sup>13</sup> Is this a problem of observation rubrics or poor evaluation system implementation?

The time and cost associated with training observers to rate accurately is pivotal, and failure to do so may contribute to inflated scores.<sup>14</sup> While the report cites newspaper accounts of inflated scores that occurred in several states' first round of implementing new evaluation systems, the key question again is whether score inflation is a problem of rubrics or, instead, poor evaluation system implementation.

The claim that teachers do not receive adequate feedback based on observations is grounded in a TNTP survey of 117 teachers who have been recognized through elite teacher awards. The report states, "only half of respondents agreed or strongly agreed with the statement, *I get regular, constructive feedback on my teaching*" (p.3). While this finding may be one many teachers agree with, will a new rubric fix the problem that many administrators do not have time, and some do not have instructional coaching skill, to provide regular, constructive feedback?

Thus, in seeking to create a case for the problems of observation tools in teacher evaluation systems, the practical problems identified are likely to be affirmed by those in schools and districts. There is, however, no research-informed argument offered to

support the report's claim that new observation rubrics will fix the problems of inadequate observer training, insufficient monitoring of rater calibration, and lack of time or low skills in providing instructional feedback.

One problem the report highlights is indeed a substantive and methodological one: the complexity of rubrics and how that complexity affects an observer's ability to score accurately. Relying on the MET project's *Gathering Feedback* study, the report declares that "beyond a handful of indicators, observers have too much trouble keeping the competencies and indicators distinct and become overloaded, generally rating most components similarly" (p.2). However, this assertion overlooks relevant caveats or qualifications made in *Gathering Feedback*, which states,

We need to determine whether the small number of principal components discerned by these instruments reflects a general fact about teaching—that is, teaching competencies do come in a small number of packages or clusters—or whether individual observers can only keep track of a few sets of competencies in their minds simultaneously.<sup>15</sup>

Further research is warranted to settle whether clustering of ratings is more likely due to covariation or observers' cognitive overload.

The report's specific claim that "any rubric with more than 10 scored elements is likely to see significant covariation among the standards" is not supported by any citations of empirical research. However, this is a rhetorically punchy statement. The number of items on a rubric often reflects how comprehensive a tool it is. There may be good reasons for having a longer tool. When observation rubrics are used as a *learning tool*, it is often desirable for the tool to represent the full range of research-based teaching behaviors and the complexities of teaching practice. A clinical coach does not attempt a comprehensive observation during each visit, but rather focuses on selected dimensions over a series of observations. When observation rubrics are used as an *evaluative tool*, a smaller number of well-chosen and well-specified observable practices is often more appropriate; such a tool is more likely to lead to reliable and valid ratings. No research-based finding is offered that leads one to conclude that the number of elements in a rubric should be capped at 10.<sup>16</sup>

## The Report's Recommendations to Fix Classroom Observations

The report's first recommendation—to attend more to a lesson's content in observation rubrics—reflects a commonsense view of how to enact change, in this case Common Core implementation. The implied rationale is that our observation tools—what we look for—should reflect what we value, what we want to see improved, or both. Sidestepping debates about the substantive and intellectual merit of Common Core standards or the politics of their adoption, it is sensible to devote attention in a rubric to a lesson's content, and this is a noted weakness in many of the instruments currently available. While attention to content does matter, the proposed tools may serve more to pressure teachers into

accepting and implementing the Common Core standards rather than to improve the intellectual challenge of classroom instruction.<sup>17</sup>

The Common Core standards specify intellectual performances students will be able to demonstrate at different grade levels. As such, the Core presses teachers to make wise choices about both content and instructional practices. Making better use of extant research-based tools may be more practical and prudent than developing new observation

*The report does not build a research-informed case for its conclusions.*

rubrics. For example, the MET project used both generic and content-focused tools and found that a number of teaching practices that guide students to do demanding intellectual work were infrequently observed (e.g., problem-solving [CLASS], effective discussion [FFT], intellectual challenge [PLATO], richness, [MQI] and investigation [UTOP]).<sup>18</sup> Improving instructional feedback that supports teachers to enact complex teaching practices may be a more promising strategy to leverage change inside classrooms, where the real action of teaching and learning take place.

The report's second recommendation—pare down observation rubrics—is grounded in the MET Project's finding that teaching competencies tend to cluster.<sup>19</sup> The report does not take up how one justifies which domains will be selected and which eliminated. Ultimately, such decisions rely on the purpose of these streamlined observation tools, and the report does not make clear whether the newly proposed tools will be primarily for formative, teacher learning purposes or for evaluative purposes. A streamlined tool that focuses on Common Core content may well leave out other important dimensions of teaching practice. This suggests some caution in assuming the new tool will be suitable for comprehensive teacher evaluation, particularly of novice or probationary teachers. Furthermore, there is always the danger that the culling process will oversimplify the complexity of teaching practice required to enact Common Core standards.

## **VII. Usefulness of the Report for the Guidance of Policy & Practice**

Schools and districts are in the early stages of implementing both the Common Core and ambitious teacher evaluation systems, and are thus encountering predictable, practical challenges. To “fix” the problems of teacher evaluation systems, the report proposes new streamlined observation tools and focusing on content.

However, the root problem may not be the observation tools. New tools will have little success fixing problems that are driven primarily by poor observer training, poor monitoring of observer quality, inadequate time in schools for observers and teachers to confer, or lack of support for teachers to learn new instructional practices and curriculum.



Whether the new observation tools proposed in the report will make a difference in Common Core implementation is up for grabs. Tools that help well-trained observers focus thoughtfully on lesson content may guide substantive improvements in instruction. In considering the status of annual teacher performance assessments Kennedy (2010) states, “documenting the intellectual meaning of teaching events remains the elusive final frontier in performance assessments.”<sup>20</sup> While Kennedy points to several promising assessment tools that attend to a lesson’s content, ultimately, the field has not yet settled whether content focus in an observation tool is an effective teacher evaluation method. Indeed, to evaluate lesson content through the review of artifacts such as student work samples may prove more illustrative and efficient. To improve instructional practice, the field already has available a number of research-based classroom observation instruments from the MET study. These carefully vetted tools include indicators of the intellectually demanding instructional practices that will be needed to implement Common Core standards.

A new instrument that does not solve the structural problems many schools and districts face in implementing observation in new teacher evaluation systems is unlikely to be very helpful. Moreover, in the report, the rationale for building an observation tool around content is not well established, nor is a research-informed argument made for why this new tool will be more successful than better implementation of high-quality observation rubrics already available. The report appears to be a sales-pitch for TNTP’s soon-to-be-launched observation tools focused on Common Core lesson content.

## Notes and References

---

- 1 Common Core Standards Initiative (2012). *Mission Statement*. Retrieved November 30, 2013, from <http://www.corestandards.org/>.
  - 2 United States Government Accountability Office (2013). *Race to the Top: States Implementing Teacher and Principal Evaluation Systems*. Report to the Chairman, Committee on Education and the Workforce, House of Representatives. Washington, DC: Author. Retrieved November 30, 2013, from <http://www.gao.gov/products/GAO-13-777>.
  - 3 TNTP and Student Achievement Partners (2013). *Fixing Classroom Observations: How Common Core Will Change the Way We Look at Teaching*. New York: Author. Retrieved November 30, 2013, from [http://tntp.org/assets/documents/TNTP\\_FixingClassroomObservations\\_2013.pdf](http://tntp.org/assets/documents/TNTP_FixingClassroomObservations_2013.pdf).
- See also the following TNTP Reports:
- TNTP (2010). *Teacher Evaluation 2.0*. New York: Author. Retrieved November 30, 2013 from <http://tntp.org/assets/documents/Teacher-Evaluation-Oct10F.pdf> ;
- For a review, see:
- Milner, H.R. IV (2010). *Review of "Teacher Evaluation 2.0."* Boulder, CO : National Education Policy Center. Retrieved November 30, 2013, from <http://nepc.colorado.edu/thinktank/review-teach-eval-TNTP>;
- Weisberg, D. Sexton, S. Mulhern, & J. Keeling, D. (2009). *The Widget Effect: Our National Failure to Acknowledge and Act on Teacher Effectiveness Differences*. New York: TNTP. Retrieved November 30, 2013 from [http://tntp.org/assets/documents/TheWidgetEffect\\_2nd\\_ed.pdf](http://tntp.org/assets/documents/TheWidgetEffect_2nd_ed.pdf);
- For a review, see:
- Pecheone, R. L. & Wei, R. C. (2009). *Review of "The Widget Effect: Our National Failure to Acknowledge and Act on Teacher Differences."* Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved November 30, 2013, from <http://nepc.colorado.edu/files/TTR-Pecheone-WIDGET.pdf>.
- 4 Student Achievement Partners (n.d.) Our purpose (web page). New York: Author. Retrieved November 30, 2013, from <http://www.achievethecore.org/about-us>.
  - 5 McDonnell, L. & Weatherford, M.S. (2013). Organized interests and the Common Core. *Educational Researcher*, 42(9), 488-497.
  - 6 Goe, L., Holdheide, L., & Miller, T. (2011). *A practical guide to designing teacher evaluation systems: A tool to assist in the development of teacher evaluation systems*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved November 30, 2013 from [http://www.ets.org/research/policy\\_research\\_reports/publications/report/2011/ivtg](http://www.ets.org/research/policy_research_reports/publications/report/2011/ivtg).
  - 7 TNTP (2013). *Perspectives on Irreplaceable teachers: What America's best teachers think about teaching*. New York: Author. Retrieved November 30, 2013, from [http://tntp.org/assets/documents/TNTP\\_Perspectives\\_2013.pdf](http://tntp.org/assets/documents/TNTP_Perspectives_2013.pdf).

8 Kane, T.J. & Staiger, D.O., et al. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observation with Student Surveys and Achievement Gains* (MET Project Research Paper). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved November 30, 2013, from [http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).

For a review, see:

Guarino, C. & Stacy, B. (2012). *Review of "Gathering Feedback for Teaching."* Boulder, CO: National Education Policy Center. Retrieved November 30, 2013, from <http://nepc.colorado.edu/thinktank/review-gathering-feedback>.

9 Kane, T.J. & Staiger, D.O., et al. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observation with Student Surveys and Achievement Gains* (MET Project Research Paper). Seattle, WA: Bill & Melinda Gates Foundation. Retrieved November 30, 2013, from [http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).

10 See, for example, the following research syntheses:

Gitomer, D. (Ed.) (2009). *Measurement Issues and Assessment for Teaching Quality*. Thousand Oaks, CA: Sage;

Gitomer, D. H., & Bell, C. A. (2013). Evaluating teaching and teachers. In K. F. Geisinger (Ed.), *APA Handbook of Testing and Assessment in Psychology* 3, 415–444. Washington, DC: American Psychological Association;

Goe, L., Bell, C., & Little, O. (2008) *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved November 30, 2013, from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>;

Kennedy, M. (2010). Approaches to Annual Performance Assessment. In M. Kennedy (Ed.). *Teacher Assessment and the Quest for Teacher Quality: A Handbook*, 225-250. San Francisco, CA: Jossey-Bass;

Porter, A., Youngs, P., & Odden, A. (2001). Advancement in teacher assessments and their uses. In V. Richardson (Ed.) *Handbook of Research on Teaching*, 4th ed., 259-297. Washington, DC: American Educational Research Association.

11 Strong, M., Gargani, J., & Hacifazlıoğlu, Ö. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education*, 62(4), 367-382.

12 Darling-Hammond, L. (2013). *Getting Teacher Evaluation Right: What Really Matters for Effectiveness and Improvement*. New York: Teachers College;

Hinchey, P. (2010). *Getting teacher assessment right: What policymakers can learn from research*. Boulder, CO: National Education Policy Center. Retrieved November 30, 2013, from <http://nepc.colorado.edu/publication/getting-teacher-assessment-right>;

Goe, L., Holdheide, L., Miller, T. (2011). *A Practical Guide to Designing Teacher Evaluation Systems: A Tool to Assist in the Development of Teacher Evaluation Systems*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved November 30, 2013, from [http://www.ets.org/research/policy\\_research\\_reports/publications/report/2011/ivtg](http://www.ets.org/research/policy_research_reports/publications/report/2011/ivtg);

Joe, J., Tocci, C., Holtzman, S., Williams, J. (2013). *Foundations of Observation: Considerations for Developing a Classroom Observation System That Helps Districts Achieve Consistent and Accurate Scores*. MET Project

Policy and Practice Brief. Seattle, WA: Bill & Melinda Gates Foundation. Retrieved November 30, 2013, from [http://www.metproject.org/downloads/MET-ETS\\_Foundations\\_of\\_Observation.pdf](http://www.metproject.org/downloads/MET-ETS_Foundations_of_Observation.pdf);

Popham, J. (2013). *Evaluating America's teachers: Mission possible?* Thousand Oaks, CA: Corwin.

13 Goe, L., Bell, C., & Little, O. (2008) *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved November 30, 2013, from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>.

14 For a discussion about the importance of proper training of raters, see:

Goe, L., Bell, C., & Little, O. (2008) *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality, 25. Retrieved November 30, 2013, from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>.

On objectivity in scoring, see:

Kennedy, M. (2010). Approaches to Annual Performance Assessment. In M. Kennedy (Ed.), *Teacher Assessment and the Quest for Teacher Quality: A Handbook*, 225-250. San Francisco, CA: Jossey-Bass, 243-245.

15 Kane, T.J. & Staiger, D.O., et al. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observation with Student Surveys and Achievement Gains* (MET Project Research Paper). Seattle, WA: Bill & Melinda Gates Foundation, 33. Retrieved November 30, 2013, from [http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).

16 For a discussion about the different purposes of observation tools and the importance that must be paid to “dual-purpose” tools., see:

Goe, L., Bell, C., & Little, O. (2008) *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality, 12-14. Retrieved November 30, 2013, from <http://www.gtlcenter.org/sites/default/files/docs/EvaluatingTeachEffectiveness.pdf>.

17 Mary Kennedy (2010) offers a helpful historical analysis of how teacher performance assessments, such as a classroom observation tools, have changed over time to accommodate shifts in what the general public values about good teaching. In the chapter's concluding discussion, she states:

The most significant problem remaining in most state and district assessments of teaching, including those based on Danielson's *Enhancing Professional Practice: Framework for Teaching*, is that they still do not give sufficient attention to the substantive and intellectual merits of classroom lessons themselves. Evaluators may be asked to judge the clarity of lesson plans, the quality of classroom interactions, the physical arrangement of the classroom, or interactions with parents, but they rarely are asked to evaluate the accuracy, importance, coherence, or relevance of the content that is actually being taught or to rate the level of intellectual work students are asked to do (p.246).

Kennedy, M. (2010). Approaches to Annual Performance Assessment. In M. Kennedy (Ed.). *Teacher Assessment and the Quest for Teacher Quality: A Handbook*, 225-250. San Francisco, CA: Jossey-Bass.

18 See:

Kane, T.J. & Staiger, D.O., et al. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observation with Student Surveys and Achievement Gains*. MET Project Research Paper. Seattle: WA: Bill & Melinda Gates

Foundation, 25. Retrieved November 30, 2013, from  
[http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).

19 Kane, T.J. & Staiger, D.O., et al. (2012). *Gathering Feedback for Teaching: Combining High-Quality Observation with Student Surveys and Achievement Gains*. MET Project Research Paper. Seattle: WA: Bill & Melinda Gates Foundation, 28-33. Retrieved November 30, 2013, from  
[http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).

20 Kennedy, M. (2010). Approaches to Annual Performance Assessment. In M. Kennedy (Ed.). *Teacher Assessment and the Quest for Teacher Quality: A Handbook*, .225-250. San Francisco, CA: Jossey-Bass, 246.

DOCUMENT REVIEWED:

**Fixing Classroom Observations: How  
Common Core Will Change the Way  
We Look at Teaching**

AUTHORS:

TNTP and Student Achievement Partners

PUBLISHERS/THINK TANKS:

TNTP and Student Achievement Partners

DOCUMENT RELEASE DATE:

November 2013

REVIEW DATE:

January 21, 2014

REVIEWER:

Jennie Whitcomb, University of Colorado  
Boulder

E-MAIL ADDRESS:

Jennie.Whitcomb@Colorado.edu

PHONE NUMBER:

(303) 735-3029

SUGGESTED CITATION:

Whitcomb, J. (2014). *Review of "Fixing Classroom Observations."* Boulder, CO: National Education Policy Center. Retrieved [date] from <http://nepc.colorado.edu/thinktank/review-fixing-classroom-observations>.