

Using Classroom Observations in the Evaluation of Special Education Teachers

Nathan D. Jones, Courtney A. Bell, Mary Brownell, Yi Qi, David Peyton, Daisy Pua, Melissa Fowler, Steven Holtzman*

POLICY ISSUE

While teacher evaluation systems and policies have evolved nationwide during the past decade, the use of structured classroom observation tools remains nearly universal. Districts commonly use the scores from observations to inform both human capital decisions and teacher improvement efforts. Districts are also likely to use the same observation tools across all of their teachers, regardless of their instructional assignment. We know very little about the tradeoffs of this choice. This paper examines one of the most popular observation systems in the country—Charlotte Danielson’s Framework for Teaching (FFT)—and asks the question: “How well do common observation systems capture the full range of teaching skills and behaviors that special educators need to use to teach effectively?”

STUDY DATA AND FINDINGS

To address this question, the authors examined more than 200 video lessons from a sample of 51 elementary and middle school special education teachers in Rhode Island. Lessons were then evaluated by expert raters and scored using two observation systems—the FFT and the Quality of Classroom Instruction (QCI), the latter of which is believed to more closely reflect the kinds of instructional practices valued by the special education community. By comparing the different dimensions of practice aligned with either the generalized (FFT) or specialized (QCI) observation tool, the authors were able to assess differences in the ways the practice of special education teachers may be rated using these various lenses.

Overall, the authors find that FFT has only limited use in assessing the quality of instructional practices most valued in special education. Highlighted findings include:

- No teachers received “distinguished” marks on any FFT components for any of the lessons, despite receiving higher scores on the QCI. Regardless of how strong teachers’ lessons looked on the QCI, their FFT scores were unlikely to be above proficient.
- The low scoring patterns were especially pronounced on FFT’s instructional components (as opposed to classroom climate). Less than 1% of lessons scored at the proficient level on the Instruction Domain of the FFT, while nearly 40% of lessons received the highest rating on QCI.

KEY FINDINGS

There are marked differences between how Special Education Teachers are rated depending on which observation tool is being used. When using Charlotte Danielson’s widely implemented Framework for Teaching observation tool, the study found:

- Not a single component of lessons taught by special educators received a “Distinguished” rating
- Less than 1% of lessons were rated as “Proficient” on the Instruction Domain

These findings are in *contrast* to ratings, for the same lessons, when using an observation protocol designed to look specifically at explicit instruction practices prioritized for teaching students with disabilities.

These findings have implications for state and local leaders when it comes to:

- High stakes decisions about teacher performance;
- The type and quality of feedback provided to educators, and;
- Implementation of instructional practices designed to explicitly and systematically teach students (e.g., science of reading)

*For full author list and associated affiliations, please refer to the working paper at wheelockpolicy.org.

POLICY IMPLICATIONS

These findings provide the first comprehensive empirical evidence of how a popular observation system functions for special education teaching. There are critical implications for school, district, and state leaders as they consider the impact of these observation systems on students and teachers.

Most notably, the findings suggest that special education teachers may not be receiving the type of evidence-based feedback that they need to improve their practice; and, in fact, they may instead be encouraged to implement instructional strategies counter to best practices for serving students with disabilities. Teachers of students with disabilities simply cannot afford to go down improvement paths that do not align with their students' needs, as could be the case if observation systems incorrectly diagnose the areas where teachers need to improve. Policymakers should consider these findings in making decisions about which observation tools are being used, for whom, and for what purpose.

Additionally, these results have consequences beyond special educators. Ideally, all teachers should have expertise in a range of instructional practices that they might deploy based on their students' needs. If the FFT is not capturing teacher-led instructional practices (e.g., explicit, systematic instruction), which have been shown to be particularly beneficial when learners are first encountering new content, ideas, or skills, this may be equally problematic for general educators. This may be most relevant right now for policymakers considering statewide or district-based initiatives to encourage a stronger adoption of practices aligned with the science of reading.

More broadly, the research suggests that we should all, researchers and practitioners alike, more flexibly consider the definition of effective teaching. It should not be defined only as what's good for most but what might be necessary to prioritize if we envision a system where all students are receiving the instruction they each need, particularly those who have been most marginalized by our system.

ADDITIONAL RESOURCES

Observing Special Education Toolkit (<https://sites.bu.edu/setleaders/>): A companion set of resources to help principals and other administrators effectively observe special education teachers and provide meaningful, evidence-aligned feedback

Morris-Mathews, H., Stark, K. R., Jones, N. D., Brownell, M. T., & Bell, C. A. (2021). Danielson's Framework for Teaching: Convergence and divergence with conceptions of effectiveness in special education. *Journal of Learning Disabilities, 54*(1), 66–78.

Gilmour, A. F., & Jones, N. D. (2020). Policies that define instruction: A systematic review of states' and districts' recommendations for evaluating special educators. *Educational Researcher, 49*(9), 645–655

FULL REPORT

For the complete working paper, visit wheelockpolicycenter.org.

This research was supported in part by a grant from the Institute of Education Sciences (Award # R324A150231)

OUR MISSION

The Wheelock Educational Policy Center (WEPC) conducts and disseminates rigorous, policy-relevant education research in partnership with local, state, and federal policymakers and stakeholders to improve educational opportunities and holistic outcomes for underserved students.

www.wheelockpolicycenter.org
wheelockpolicy@bu.edu



Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

