Texas Education Review

An Overview of the Issues in Evaluating Special Educators: Current Challenges and Recent Developments

Gavin W. Watts, MEd
The University of Texas at Austin

An Overview of the Issues in Evaluating Special Educators: Current Challenges and Recent Developments

Gavin W. Watts, MEd
The University of Texas at Austin

Teachers have always been evaluated in some way; yet, current roles in the classroom have shifted to more scientific and data based approaches to teacher evaluation. These specialized approaches to teaching and evaluation have raised questions about using a single or uniform tool to assess educators. Special educators are particularly impacted by teacher evaluations since their roles require unique instructional and behavioral management skills. In addition to the range of technical and socio-emotional competencies, special educators are responsible for ensuring that all of their students' learning and academic needs are met. In this backgrounder, the author first provides a brief history of special education teacher evaluation processes, then considers the complexities of special educators' roles and responsibilities, and finally, interrogates special education educator evaluation within a new policy context.

A Brief History of Teacher Evaluation

In the early part of the twentieth century, teacher evaluations in the United States were grounded in mostly personal qualities (e.g., confidence, grooming, or charisma). A few decades later, Marzano, Frontier, & Livingston (2011) proposed the idea of incorporating student learning into these evaluations in order to determine the effectiveness of individual teachers and schools. These measures established assessment procedures focusing on a teacher's use of specific instructional strategies and behaviors. Soon after, a scientific approach to schooling emerged which utilized standardized testing at greater rates, as well as supported democratic values and social development as main initiatives. By the 1950s, teacher evaluations became more formalized and were conducted by way of industry assessments (i.e., inventories, checklists) (Medley, Coker, & Soar, 1984). The 1960s through the early 1970s saw the rise of clinical supervision as a functional method for evaluating teachers (Medley, Coker, & Soar, 1984). By the 1980s, Madeline Hunter's (1982) seven-step framework for lesson planning became the content of pre- and post- observation conferencing and the basis for evaluating and determining teacher growth and effectiveness. The components of Hunter's supervisory conferences include: identifying and explaining research-based instructional behaviors; assisting in the identification of ineffective components within lessons; encouraging teachers to revise their approaches; promoting the continued growth of excellent teachers; and overall, aiming to remove subjective opinion from the evaluation process and instead relying on objective evidence (Hunter, 1982).

This shift in teacher evaluation methodology from the start of the century to the mid-1980s illustrates a growth towards more scientific and systematized approaches to teacher evaluations. Despite these trends, there continues to be a lack of research evidence to support specific evaluative practices, such as effective methods of assessing teachers' performance and their relationship to student learning outcomes. Wise, Darling-Hammond, McLaughlin, and Bernstein (1985) found that many of the teacher supervision and assessment methods utilized in

schools were based upon standard reflective practices and lacked effectiveness in fostering growth in teaching practices due to their rigid nature. These findings were supported by a study of teacher evaluation practices of the largest 100 school districts' in the US (Ellett & Garland, 1987). The results of this national survey revealed several concerns about the quality and types of teacher evaluation systems. Specifically, Ellett and Garland (1987) identified a disconnect between teacher evaluation methodologies and professional development, the reliability of system approaches, and the establishment of professional standards and training processes. Ten years later, Loup, Garland, Ellet & Rugutt (1996) replicated the study and similar results were found, suggesting that improvements in these systems had not yet been made.

   With an aim to address several of the concerns identified in teacher evaluation research, Danielson's (1996) *Enhancing Professional Practice: A Framework for Teaching* identified 76 components of quality educators and provided a language and system to address instructional improvement through self-assessment and reflection. This system was comprehensive for incorporating various aspects of the teaching process, from planning to implementation to reporting student outcomes (Danielson, 1996). Danielson's (1996) framework also incorporated a measurement system in the form of a rating scale based on levels of proficiency (i.e., unsatisfactory, basic, proficient, and distinguished). As such, a more responsive evaluation system began emerging in the late 1990s, which reflected a growing research interest and new findings in the area of "teacher effectiveness," and acknowledged the impact of teachers on student learning (Danielson, 2013; Danielson & McGreal, 2000).

   In 2001, the *No Child Left Behind Act* (NCLB) placed a firm emphasis on school accountability and the importance of developing high-quality educators. In what Weisberg, Sexton, Mulhern, and Keeling (2009) call *The Widget Effect*, school districts often assume that all teachers are equal in effectiveness, from classroom to classroom. Insight into the evaluation methods of twelve school districts within four different states raised several concerns about assessing educator differences: the limited time provided for assessment process, unqualified/untrained administrators leading the evaluations, and a lack of follow-up support and resources for professional development (Weisberg, Sexton, Mulhern, & Keeling, 2009). Without an evaluative system that can accurately identify educators who are in need of improvement, as well as a system for incentives and promotion, teaching effectiveness will be difficult to improve.

   The conflict between what researchers and practitioners know and have identified as quality teaching and the lack of accurate and credible teacher evaluation systems continue to beckon further research. Previous evaluation systems based upon direct observation methods have demonstrated issues in accurately gauging an educator's effectiveness (Weisberg, Sexton, Mulhern, & Keeling, 2009), although recent developments in methodological approaches have shown promise (Johnson, Ford, Crawford & Moylan, 2016a; forthcoming). In addition to direct observation approaches, newer initiatives have explored value-added measures based on student growth (Carey, 2004; Glazerman, Loeb, Goldhaber, Staiger, Raudenbush, & Whitehurst, 2010). This model has become highly recommended due to its focus on student growth, rather than student achievement, in relation to set standards. In fact, The Bill and Melinda Gates Foundation (2011; 2013) commissioned a study, which included 3,000 teachers across seven districts, to examine the role of value added measures and its effect on student learning. Although frequently endorsed, the support for this system has received its fair share of criticism, with critiques on the heavy reliance on standardized testing, the validity and reliability of comparing scores across years, and the decision-making process based upon outcomes from these measures (Buzick &

Laitusis, 2010; Corcoran, 2010). While most of this opposition focuses on general educators, little attention has been paid to the ways in which value-added measures fail to capture students with disabilities, many of whom do not display consistent and straightforward growth models (Lawson, 2014). This gap is even more important when considering the evaluation of special educators, since a value added evaluative system does not take into account the unique arrangement of special education instruction and data collection (which is more frequently structured at the small group and/or individual level) (Holdheide, Browder, Warren, Buzick, & Jones, 2012).

Currently, the United States is transitioning to newly developed teacher evaluation systems, per *Race to the Top* and *Every Student Succeeds Act* initiatives (U.S. Department of Education, 2009), to improve the effectiveness of schools through multiple means of assessing educators (McGuinn, 2012; Murphy, Hallinger, & Heck, 2013). However, these general aims do not directly address the more complex aspects and needs when evaluating non-general educators. Although research has identified teacher effectiveness as a consistent indicator for having the greatest impact on student achievement; the relations between teacher impact and student outcomes becomes less distinctive when considering special education (Blanton, Sindelar, & Correa, 2006). Consequently, the complexities of developing new systems that incorporate the unique dynamics of special education remains a current challenge for the field.

The Unique Issues and Challenges in Evaluating Special Educators

Much progress has been made in the support for, and development of new evaluation models for teachers, but special educators still lack effective feedback systems. Current models have been developed with little differentiation for addressing the unique needs of different teacher roles, instructional models, and content areas (Prince, Schuermann, Guthrie, Witham, Milanowski, & Thorn, 2009). These systems, which were developed for general education, do not support the individual focus and needs of special education students and educators (Holdheide, Browder, Warren, Buzick, & Jones, 2012). Without sensitive and effective teacher evaluation systems that include, or are adapted for the unique needs of special educators, the necessary feedback and suggestions for professional development will not be available for special educators. Therefore, the current models are less likely to assist in improving the educational outcomes of students with the most need. Additionally, special educators will be included in accountability mandates where their compensation is determined by the proposed performance-based systems. Thus, there is a need to assess and address the unique issues in evaluating special educators, which will lead to the development or differentiation of evaluation systems that are sensitive to these characteristics and needs.

In addition to the common duties and roles of general educators, special educators must be prepared for challenging and unique pedagogical needs, behavior management, and content knowledge demands (Bettini, Kimerling, Park, & Murphy, 2015). These responsibilities, along with organizational support and high-need work conditions, are frequently associated with challenges to recruiting and retaining special educators (McLesky & Billingsley, 2008). In some schools and districts with open special education positions, they are filled with unqualified or uncertified educators or even, not filled at all (Boe & Cook, 2006). Consequently, this creates an ineffective instructional setting where evidence-based teaching practices are absent for students who need the most support (Reschly, Holdheide, Smart, & Oliver, 2007). These challenges to recruit and retain special education educators also has negative long-term outcomes for the field

(Billingsley, 2004). Still, the greatest impact of attrition occurs in schools and classrooms where student learning is compromised. Despite the support that suggests special education educators have limited instructional time due to increasing caseloads (Carlson, Brauen, Klein, Schroll, & Willig, 2002; McLeskey, Tyler & Flippin, 2004), special educators also have limited organizational input and are often given undefined and underutilized roles in inclusive or collaborative models of instruction (Gehrke & Murri, 2006). Although research generally supports collaboration between teachers, the ability to assess the individual impact of a given special education teacher on student achievement is difficult when special education educators share in the planning, coordination, and implementation of instruction with general educators. Furthermore, due to the varying roles of special educators, their unique responsibilities may change from classroom to classroom and school to school. Special educators are expected to possess a set of skills and knowledge that is unique to the needs of their students, and therefore, quite different from those of general educators. Evaluation systems focusing primarily on academic outcomes would overlook the instruction and acquisition of social skills, functional and life skills, and behavior management/support (Council for Exceptional Children, 2012a). This further complicates the evaluation process and creates a greater demand for flexible and sensitive measurement systems that account for these contrasting proficiencies.

In examining the current evaluation methods, systematic barriers to measurement also exist due to the unique nature of special education students and teachers. First, considerations must be taken to address evaluation methods that are based upon general student achievement, as students' learning trajectories may differ depending on their unique disabilities or needs. Next, students with disabilities are often assessed according to their legally mandated accommodations, modifications, and alternate standards, as outlined in their *Individualized Education Plan (*IEP). Research on the extent to which student accessibility is considered when developing teacher evaluation systems based on student performance on standardized tests remains scarce. Finally, high rates of student mobility (Ashby, 2010) and teacher attrition (Goldring, Taie, & Riddles, 2014) are also barriers that need to be addressed within the evaluation procedures for special educators. In sum, these multiple concerns make evaluation systems that require year-to-year and longitudinal comparisons of student and/or teacher progress extremely challenging.

Implications

Although the challenges in evaluating special educators are many, research in special education have identified several recommendations to support the development of evaluation systems that will support the field. These recommendations include taking into account the unique roles of special educators and the related evidence-based practices that are required to promote the learning for students with disabilities (Council for Exceptional Children, 2012b; National Council on Teacher Quality, 2012). In assessing the learning of this population of students, simply looking at student growth will not reveal the overall effectiveness of special educators (Council for Exceptional Children, 2012b). Furthermore, future systems must support professional development in order to ensure that special educators are well-prepared to meet the diverse needs of their students, and additionally, to deter the attrition of effective special educators, an issue that has long plagued the field (Billingsley, 2004; Council for Exceptional Children, 2012b). Moving forward, as reforms continue to adapt to findings from new research initiatives in teacher evaluation methods, so must the considerations for how these systems impact the learning of students with disabilities.

_____

Gavin W. Watts is a doctoral student in early childhood special education and behavioral disorders within the Department of Special Education at The University of Texas at Austin. His current research interests include no-cost behavioral interventions, cross-age tutoring models for promoting early numeracy, and evaluation systems in teacher preparation.

References

Ashby, C. M. (2010). *K-12 education: Many challenges arise in educating students who change schools frequently. Report to Congressional Requesters. GAO-11-40*. US Government Accountability Office.

Bettini, E., Kimerling, J., Park, Y., & Murphy, K. M. (2015). Responsibilities and instructional time: Relationships identified by teachers in self-Contained classes for students with emotional and behavioral disabilities. *Preventing School Failure: Alternative Education for Children and Youth*, *59*(3), 121–128. doi:10.1080/1045988X.2013.859561

Bill & Melinda Gates Foundation. (2011). *Learning about teaching: Initial findings from the Measures of Effective Teaching project*. Bellevue, WA: Author. Retrieved from www.gatesfoundation.org/college-ready-education/Documents/preliminary-findings-research-paper.pdf

Bill & Melinda Gates Foundation (2013). *Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project's three-year study*. Retrieved from http://www.metproject.org

Billingsley, B. S. (2004). Special education teacher retention and attrition: A critical analysis of the research literature. *The Journal of Special Education*, *38*(1), 39–55. doi:10.1177/00224669040380010401

Blanton, L. P., Sindelar, P. T., & Correa, V. I. (2006). Models and measures of beginning teacher quality. *Journal of Special Education, 40*, 115–127. doi:10.1177/00224669060400020201

Boe, E. E., & Cook, L. H. (2006).   The chronic and increasing shortage of fully-certified teachers in special and general education. *Exceptional Children, 72*(4), 443–460.

Buzick, H. M., & Laitusis, C. (2010). Using growth for accountability: Measurement challenges for students with disabilities and recommendations for research. *Educational Researcher, 39*(7), 537–544. doi:10.3102/0013189X10383560

Carey, K. (2004). The real value of teachers: Using new information about teacher effectiveness to close the achievement gap. *Thinking K-16, 8*(1). Retrieved from http://www.edtrust.org/sites/edtrust.org/files/publications/files/Spring04_0.pdf

Carlson, E., Brauen, M., Klein, S., Schroll, K. & Willig, S. (2002). *Study of Personnel Needs in Special Education: Key Findings*. Retrieved from http://fer.dig.coe.ful.edu/spense/KeyFindings.pdf.

Corcoran, S. P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice* (Education Policy for Action Series). New York, NY: Annenberg Institute for School Reform at Brown University.

Council for Exceptional Children. (2012a). CEC initial level special educator preparation standards. Retrieved from

http://www.cec.sped.org/sitecore/shell/Controls/Rich%20Text%20Editor//~/media/Files/Standards/Professional%20Preparation%20Standards/Initial%20Preparation%20Standards%20with%20Elaborations.pdf

Council for Exceptional Children (2012b). The Council for Exceptional Children's Position on Special Education Teacher Evaluation. *Council for Exceptional Children 2012 Policy Manual, 4*(3), L-9. Retrieved from http://www.cec.sped.org/~/media/Files/Policy/CEC%20Professional%20Policies%20and%20Positions/Position_on_Special_Education_Teacher_Evaluation_Background.pdf

Danielson, C. (1996). *Enhancing professional practice: A framework for teaching.* Alexandria, VA: Association for Supervision and Curriculum Deveklopment.

Danielson, C. (2013). *The framework for teaching evaluation instrument, 2013 Edition* (2nd ed.). Princeton, NJ: Danielson Group.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Princeton, NJ: Educational Testing Services.

Ellett, C.D., & Garland, J. (1987). Teacher evaluation practices in our largest school districts: Are they measuring up to 'state-of-the-art' systems? *Journal of Personnel Evaluation in Education*, *1*(1), 69–92.

Gehrke, R. S., & Murri, N. (2006). Beginning special educators' intent to stay in special education: Why they like it here. *Teacher Edcuation and Special Education, 29,* 179–190. doi:10.1177/088840640602900304

Glazerman, S., Loeb, S., Goldhaber, D., Staiger, D., Raudenbush, S., & Whitehurst, G. (2010). *Evaluating teachers: The important role of value-added*. Brookings Institution. Retrieved from http://www.brookings.edu/research/reports/2010/11/17-evaluating-teachers

Goldring, R., Taie, S., and Riddles, M. (2014). *Teacher attrition and mobility: Results from the 2012–13 teacher follow-up survey* (NCES 2014-077). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Holdheide, L., Browder, D., Warren, S., Buzick, H., & Jones, N. (2012). *Summary of using student growth to evaluate educators of students with disabilities: Issues, challenges, and next steps.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved from http://isbe.net/peac/pdf/using_student_growth_summary0112.pdf

Hunter, M. C., (1982). *Mastery Teaching.* El Segundo, CA: TIP Publications.

Johnson, E. S., Ford, J. W., Crawford, A., & Moylan, L. A. (2016a). *Recognizing Effective Special Education Teachers (RESET) Technical Manual.* Boise State University, Boise, ID.

Johnson, E. S., Ford, J. W., Crawford, A., & Moylan, L. A. (forthcoming). Issues in evaluating special education teachers: Challenges and current perspectives. *Texas Education Review, 4*(1). Retrieved from www.txedrev.org

Lawson, J. (2014). Value-added modeling: Challenges for measuring special education teacher quality. *InterActions: UCLA Journal of Education and Information Studies*, *10*(1). Retrieved from http://escholarship.org/uc/item/9r67085n

Loup, K.S., Garland, J.S., Ellett, C.D., & Rugutt, J.K. (1996). Ten years later: Findings from a replication of a study of teacher evaluation practices in our 100 largest school districts. *Journal of Personnel Evaluation in Education, 10*(3), 203-226.

Marzano, R. J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

McGuinn, P. (2012). *The state of teacher evaluation reform.* Washington, DC. Retrieved from http://www.americanprogress.org/wp-content/uploads/2012/11/McGuinn_TheStateofEvaluation-1.pdf

McLeskey, J., & Billingsley, B. S. (2008). How does the quality and stability of the teaching force influence the research-to-practice gap?: A perspective on the teacher shortage in special education. *Remedial and Special Education, 29*(5), 293–305. doi:10.1177/0741932507312010

McLeskey, J., Tyler, N., & Flippin, S. S. (2004). The supply of and demand for special education teachers: A review of research regarding the nature of the chronic shortage of special education teachers. *The Journal of Special Education, 38*(1), 5–21. doi:10.1177/00224669040380010201

Medley, D. M., Coker, H., & Soar, R. S. (1984). *Measurement-based evaluation of teacher performance: An empirical approach.* New York: Longman.

Murphy, J., Hallinger, P., & Heck, R. H. (2013). Leading via teacher evaluation: The case of the missing clothes? *Educational Researcher, 42*, 349–354. doi:10.3102/0013189X13499625

National Council on Teacher Quality. (2012). *Making effectiveness matter.* Retrieved from http://www.nctq.org/p/publications/docs/YearbookBrief_Area5.pdf

Prince, C. D., Schuermann, P. J., Guthrie, J. W., Witham, P. J., Milanowski, A. T., & Thorn, C. A. (2009). *The other 69 percent: Fairly rewarding the performance of teachers of nontested subjects and grades.* Washington, DC. Retrieved from http://www.cecr.ed.gov/guides/other69Percent.pdf

Reschly, D. J., Holdheide, L. R., Smartt, S. M., & Oliver, S. M. (2007). *Evaluation of LBS-1 teacher preparation in inclusive practices, reading, and classroom organization-behavior management.* Springfield: Illinois State Board of Education.

U.S. Department of Education. (2009). *Race to the Top program executive summary.* Washington, DC: Author. Retrieved from www2.ed.gov/programs/racetothetop/executive-summary.pdf

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* New York, NY: The New Teacher Project. Retrieved from http://widgeteffect.org/downloads/TheWidgetEffect.pdf

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *The Elementary School Journal*, *86*(1), 61–121.

Suggested Readings

Abedi, J., & Faltis, C. (2015). Teacher assessment and the assessment of students with diverse learning needs. *Review of Research in Education*, *39*(1), vii–xiv. doi:10.3102/0091732X14558995

Benedict, A. E., Thomas, R. A., Kimerling, J., & Leko, C. (2013). Trends in teacher evaluation: What every special education teacher should know. *TEACHING Exceptional Children*, *45*(5), 60–68.

Billingsley, B. S. (1989). Evaluating special educators: Context and considerations for evaluation design. *Journal of Personnel Evaluation in Education*, *2*(3), 229–238. doi: 10.1007/BF00127457

Elliott, S. N., Roach, A. T., & Kurz, A. (2014). Evaluating and advancing the effective teaching of special educators with a dynamic instructional practices portfolio. *Assessment for Effective Intervention*, *39*(2), 83–98. doi:/10.1177/1534508413511491

Hill, H. C., Kapitula, L., & Umland, K. (2011). A validity argument approach to evaluating teacher value-added scores. *American Educational Research Journal*, *48*(3), 794–831.

Johnson, E., & Semmelroth, C. L. (2014). Special education teacher evaluation: Why it matters, what makes it challenging, and how to address these challenges. *Assessment for Effective Intervention*, *39*(2), 71–82. doi:10.1177/1534508413513315

Kyriakides, L., Demetriou, D., & Charalambous, C. (2006). Generating criteria for evaluating teachers through teacher effectiveness research. *Educational Research*, *48*(1), 1–20. doi: 10.1080/00131880500498297

National Council on Teacher Quality. (2012). *State of the states 2012: Teacher effectiveness policies*. Retrieved from http://www.nctq.org/dmsStage/State_of_the_States_2012_Teacher_Effectiveness_Policies_NCTQ_Report

Steinbrecher, T. D., Selig, J. P., Cosbey, J., & Thorstensen, B. I. (2014). Evaluating special educator effectiveness addressing issues inherent to value-added modeling. *Exceptional Children*, *80*(3), 323–336. doi:10.1177/0014402914522425

U.S. Department of Education. (2011). *Developing effective teacher evaluation systems*. Retrieved from http://www.ed.gov/Teacher-Evaluation-Systems

Warger, C. L., & Aldinger, L. E. (1987). Teacher evaluation: The special case of the special educator. *NASSP Bulletin*, *71*(500), 54–62. doi:10.1177/019263658707150009