

Running Head: CHALLENGES OF SPECIAL EDUCATION EVALUATION

Issues in Evaluating Special Education Teachers: Challenges and Current Perspectives

Evelyn S. Johnson, Laura A. Moylan, Angela Crawford and Jeremy W. Ford

Boise State University

2016

Author Note

This research was supported by the Institute of Education Sciences, award number R324A150152 to Boise State University. The opinions expressed are solely those of the authors.

Correspondence concerning this article should be addressed to Dr. Evelyn Johnson,

evelynjohnson@boisestate.edu

Citation: **Johnson, E. S.**, Crawford, A., Moylan, L. A., & Ford, J. W. (2016). Issues in evaluating special education teachers: Challenges and current perspectives. *Texas Education Review*. 4, (1), 71-83.

Abstract

Observation systems can provide teachers with information about how to improve their instructional practice and can lead to improved student outcomes. However, most observation systems have not been designed to address issues specific to special education. An effective special education teacher evaluation system must measure and provide targeted, corrective feedback on instructional practice, rely on the use of raters and observers with content specific expertise and be correlated with individualized student growth measures. In this article, we explore the question of whether objective measures of special education teaching can be created and implemented in a valid and fair way that yields useful and reliable results, and examine issues related to the content of the observation, requirements for raters, and the type of feedback that will be required to support instructional change in the context of a recently funded research project, Recognizing Effective Special Education Teachers (RESET).

Keywords: special education teacher evaluation, observation systems, evidence-based practices

Issues in Evaluating Special Education Teachers: Challenges and Current Perspectives

Teacher quality is the most important variable in improving student outcomes (Goldhaber, 2016). Specifically, the quality of instruction provided by the teacher is the most important school based influence on children's academic skills (Crawford, Zucker, Williams, Bhavsar & Landry, 2013), but we know that teachers vary significantly in their impact on student learning (Chetty, Friedman & Rockoff, 2012). To improve instructional quality, state and district education policy makers are increasingly turning to teacher observation systems. While this focus on improving teacher quality is promising, current observation tools have been criticized for being too heavily focused on managerial aspects of the classroom (Crawford, Zucker, Williams, Bhavsar & Landry, 2013); for being too generic with respect to content areas (Hill & Grossman, 2013); for not providing specific feedback to teachers, which has been shown to lead towards greater gains in instructional improvement (Biancarosa, Bryk, & Dexter, 2010); and for not being relevant across a significant number of content areas, including special education (Johnson & Semmelroth, 2014; Jones & Brownell, 2014).

Students receiving special education represent approximately 12% of the K-12 population (U.S. Department of Education, 2015). As is the case with most major education reforms, teacher observation systems have developed without the inclusion of special education teachers or the students they serve. Subsequently, there are a number of unanswered questions about how best to proceed with developing an observation tool that will realize the potential of improved instructional quality for students with disabilities (SWD). Students served through special

education typically have the most intense instructional needs, and require specially designed instruction. Meeting the needs of this group of students is extremely challenging and requires teachers who are highly skilled. Unfortunately, SWD are more often served by a special education teaching force that is highly subject to attrition and turnover, which compromises the educational services that SWD receive (Billingsley, 2004; Boe, Erling, Cook, Lynn & Sunderland, 2008; Connelly & Graham, 2009). The shortage of highly trained special education teachers is a national issue, with nearly every state and U.S. territory reporting special education as a critical shortage area for over 20 years (U.S. Department of Education, 2015). This negatively impacts student outcomes. Nationally, as few as 30% of SWD have been able to meet performance standards (Odom, 2009). An observation system designed to provide special education teachers with specific feedback on their implementation of practices that have been demonstrated to lead to significant and meaningful gains for SWD offers one way to improve the special education teaching force and improve outcomes for SWD.

Designing observation systems for special education teachers is not an easy task. As Hill and Grossman (2013) have noted, if observation systems are to achieve the goal of supporting teachers in improving instructional practice, they must: 1) be subject specific, 2) involve content experts in the process of observation, 3) provide feedback that is both accurate and usable in the service of improving instruction, and 4) produce observation scores that align with student test score data to bring expectations for teaching and learning into agreement. Few instruments are available that meet these requirements for general education teachers, and the application of these principles within a special education context brings added challenges. First, special education teachers are responsible for providing instruction across a number of subject areas (e.g. reading, math, writing), which requires unique evaluation instruments for each of these

content areas across grade levels. Second, principals are often the primary evaluators of their teaching staff, yet principals typically do not have the expertise and knowledge to provide specific feedback in special education (Derrington & Campbell, 2015; Frost & Kersten, 2011). Third, special education teachers work with students who require specially designed instruction that is *individualized* depending on student need. Special education teachers therefore, must be well versed in numerous evidence-based practices (EBP) and be cognizant of various disability types to effectively plan and implement effective instruction (Odom, Brantlinger, Gersten, Horner, Thompson, & Harris, 2005), and therefore, observation systems must be able to capture a broad range of EBPs that are adapted based on individual student needs. Finally, defining student achievement through one universal measure, or even through a set of accepted, predetermined measures, poses methodological problems for SWD (Baker, Barton, Darling-Hammond, Haertel, Ladd, Linn, et al., 2010). A variety of student measures may be required to achieve the alignment between teacher and learner expectations necessary to reduce the conflicting messages that special education teachers receive about instructional improvements (Hill & Grossman, 2013), yet the inclusion of a variety of measures within one observation system will be difficult.

In summary, there are numerous challenges to developing effective observation systems, and these challenges are exacerbated in their application to special education teachers. In light of these challenges, it is reasonable to ask, Can objective measures of special education teaching be created and implemented in a valid and fair way that yields useful and reliable results? We believe so. In this article, we review the complexities of special education teacher observation, describe a pilot observation tool called *Recognizing Effective Special Education Teachers* (RESET), and outline continued steps in its development to address this critical question.

Recognizing Effective Special Education Teachers (RESET)

The RESET project is a four-year research project, funded by the Institute for Education Sciences (IES). The goal of RESET is to create a special education teacher observation tool designed to reliably evaluate instructional practice, to provide specific and actionable feedback to special education teachers about the quality of their instruction and ultimately, to improve outcomes for SWD (Johnson, Ford, Crawford & Moylan, 2016). The conceptual framework guiding RESET is that a targeted, well-defined, observation tool that incorporates clearly explicated criteria linked to EBPs in special education will target teacher attention to those instructional practices that have been demonstrated to result in improved student outcomes. RESET will evaluate the extent to which EBPs are implemented with fidelity, provide explicit feedback to the teacher on the specific components of instructional practices, and measure the impact on student outcomes. It is important to note that the current focus of RESET is on evaluating the EBPs identified for use for students with high-incidence disabilities.

Currently at the end of the first of four years of this project, we have developed a framework for the organization of RESET, and have developed initial drafts of 12 rubrics aligned with EBPs in special education instructional practice. In the remaining years of this project, we will continue to develop rubrics that align with EBPs for SWD, and conduct a number of studies to assess RESET's psychometric properties and its utility in achieving the described objectives. In this article, we describe how we are addressing the issues framed by Hill and Grossman (2013) as we develop the observation protocols. Specifically, we explain how we have designed a *subject-specific* observation instrument that provides concrete guidance on desirable teaching practices for SWD. We then detail next steps regarding raters, feedback and connections to student outcomes.

RESET framework

The use of EBPs is imperative in special education if we hope to improve the outcomes of SWD (Cook, Tankersley, & Landrum, 2009; Gersten, Vaughn, Deshler, & Schiller, 1997; Odom et al., 2005). Although standards for identifying EBPs have been articulated by the What Works Clearinghouse (U.S. Department of Education, 2013) and by the Council for Exceptional Children (Council for Exceptional Children, 2014), there have been few efforts to systematically identify EBPs for SWD, which complicates the decisions about what should be observed within a special education classroom. The National Autism Center's National Standards Project (www.nationalautismcenter.org) is perhaps the best example of a sustained and comprehensive effort to systematically review the research on EBPs against a set of agreed upon standards and to share them with practitioners. The checklists of EBPs and accompanying implementation modules (see for example, AFIRM team, 2015) that have been developed as a result of these efforts provide clear guidance to teachers working with students with autism on which practices have a strong evidence base and how to implement them.

For other disability categories, a similarly accessible clearinghouse is not quite so readily available. Although various attempts to identify EBPs for students with high incidence disabilities have been made over the years, (see for example, Cook et al., 2009), the research on EBPs is still surprisingly difficult to navigate and organize into a comprehensive set of instructional practices. The intervention tool charts and reviews provided by the National Center on Intensive Intervention (www.intensiveintervention.org), are a welcome initial resource to provide practitioners with information on effective practice, but much remains to be done to synthesize the vast research base into a set of manageable, practitioner friendly resources.

To begin the design of RESET, we conducted exhaustive literature reviews attempting to synthesize the research into a set of organizing principles that could be translated into rubrics to be employed across a variety of contexts and content areas. The result of our literature review has led to the organization of RESET into three main subscales reflecting critical aspects of special education: instructional practices, content area instruction, and individualization. Each of the subscales is briefly described below.

Instructional practices. Although students with high-incidence disabilities reflect a very heterogeneous group to which no one instructional model can be recommended, there are some common principles that underlie effective intervention programs (Swanson & Deshler, 2003; Vaughn & Swanson, 2015). The three main categories of instructional practices for which we have found substantial empirical support include explicit instruction (e.g., Archer & Hughes, year; Brophy & Good, 1986; Christenson, Ysseldyke, & Thurlow, 1989; Gersten, Schiller & Vaughn, 2000; Rosenshine & Stevens, 1986; Swanson, 1999), cognitive strategy instruction (e.g. Graham & Harris, 1989; Montague, 1992; Montague & Dietz, 2009; Swanson & Sachs-Lee, 2000) and peer-assisted learning (or reciprocal teaching) techniques (e.g. Delquadri, Greenwood, Whorton, Carta & Hall, 1986; Fuchs, Fuchs, Mathes & Simmons, 1997; Mathes, Howard, Allen & Fuchs, 1998; McMaster, Fuchs & Fuchs, 2007; Rosenshine & Meister, 1993). Meta-analyses of instructional components across content areas and intervention studies consistently support the use of these three instructional strategies for remediating the academic difficulties that students with high incidence disabilities encounter. Additionally, interventions that use a combination of these approaches tend to produce the largest effect sizes (Rosenshine & Meister, 1993; Swanson, 1999). This suggests that special education teachers should provide SWD with instruction organized around these research-validated instructional principles.

Once we identified these categories of instructional practices, we reviewed and consolidated the descriptions across studies to develop a component list for each instructional practice. Our goal was to prioritize practices with clearly identified components that are empirically validated yet flexibly designed to match various contexts and student populations (Harn, Parisi, & Stoolmiller, 2013; Odom, Fleming, Diamond, Lieber, Hanson, Butera et al, 2010).

Creating the items for each of the rubrics will be an iterative process. In the initial drafts, we have relied on careful reviews of the extant literature to develop each of the items in our rubrics. However, even when a practice has sufficient support to be called evidence-based, it can be difficult to identify the *specific* elements that comprise that EBP. Instructional practices often consist of multiple elements, and are implemented within a dynamic, complex environment (Swanson & Deshler, 2003). It is difficult to know which of the many individual elements are the *key* ingredients that lead to successful student outcomes. For example, in an instructional sequence that is comprised of eleven steps, is each step critical? Should they be weighted equally? Are there important interactions between special education teachers and their students that can be difficult to capture in an observation system? Are different instructional elements more or less important depending on the specific needs of the student?

The consideration of fidelity and how it is assessed is important as we develop observational tools to evaluate instructional practice. For example, if we weight each of the specific elements of an instructional practice as equal, we may encourage teachers to engage in some practices that are unnecessary (and therefore take time away from critical elements), or we may underemphasize practices that are critical. Most of the research on evidence-based instructional practices does not weight steps or identify those that are *crucial* versus those that

are *good but not essential*. Attempts to do so have yielded results isolating the effect of just one factor, explicit practice (Swanson & Deshler, 2003; Swanson & Hoskyn, 2001). Clearly, explicit practice cannot be the only element of a well-designed instructional lesson, but the research to date has not provided clear guidance on what instructional components to emphasize. In an evaluation system that may ultimately be tied to high stakes decisions about teachers, it will be important to better understand key elements of various EBPs so that we direct a special education teacher's efforts to the practices that are likely to have the most positive impact on student outcomes. As we continue with the development of RESET, we will conduct numerous studies that examine the predictive utility of each of the individual components on student outcomes so that we can emphasize those that seem to be the most influential in improving student outcomes. If specific elements of instructional practice do not add significantly to a predictive model, we can revise our observation rubrics to create an evaluation system that is flexible and responsive to the context and that focuses on essential elements of a practice (Harn et al., 2013).

Content Areas. While the instructional practices employed by special education teachers are critical to support information processing (Swanson & Deshler, 2003), a focus on instructional practice alone would fail to recognize the critical aspect of evaluating the content that is being presented to SWD. Student performance in reading for example, can be significantly impacted by both the quality of instruction as well as the quality of the content organization and presentation (Carnine, Silbert, Kame'enui & Tarver, 2009; Johnson & Boyd, 2013; Moats & Foorman, 2003), and a focus exclusively on instructional practices at the expense of content could lead to inaccurate evaluations of teacher performance.

The most common content areas in which students with high incidence disabilities receive individualized instruction services are reading, writing, math and social/emotional skills (Cortiella, 2015). With content areas as diverse as this, and considering that special education teachers work across grades P-12, content rubrics that reflect a broad range of areas across grade levels will need to be developed. Across the academic areas, the literature base is most well-developed for reading. Therefore, we began our work developing content specific rubrics for RESET with reading.

Reading. The National Reading Panel report outlined the Big 5 in Reading that has served as an organizing framework for understanding and researching reading intervention for the last 15 years (NICHD, 2000). However, a surprisingly small number of studies examining the effect of intensive reading intervention exclusively for SWD is available (Vaughn & Swanson, 2015). Therefore, to develop the reading rubrics, we drew on the reading research that identifies best practices for organizing reading instruction for students at risk for and those with SWD. Our current set of reading rubrics are organized into the following areas: phonological awareness, letter sound correspondence and sounding out words, multi-syllabic decoding and word analysis, vocabulary, reading for meaning, and comprehension strategies (Moylan, Johnson, Ford & Crawford, 2016). To design each rubric, we consulted multiple sources that outline the way that the content for each of these areas should be presented. Our goal was to reflect the best practices within each of the specific reading areas rather than to create checklists of a number of programs. For example, when teaching letter-sound correspondence, there are principles regarding the sequencing of letters to be taught, the structure of effective decoding lessons, and the composition of practice and discrimination activities that allow students to work to mastery (Carnine et al., 2009; Moats, Glaser, & Tolman, 2011). These principles are reflected in our

rubric. As we continue with the development of RESET, we will examine the alignment of performance on content area rubrics with outcomes relevant to that area to determine the validity of the rubric and to make revisions as needed.

Challenges with content rubric development. Content areas other than reading are not as well developed and therefore, creating rubrics that depict best-practices will be challenging. Neither the math nor writing instruction literature has been synthesized into an organizing framework in the way that the Big 5 in Reading has allowed. This raises the question of how best to construct a set of content specific frameworks within each of these content areas that will support the goal of improved instruction in these areas. To begin to tackle this issue, we are conducting syntheses of the research and consulting with content area experts to frame the rubrics in ways that align with current understandings. Once created, the validation of these rubrics will pose additional challenges, as many special education teachers are not well trained to provide instruction in either area (Brownell, Sindelar, Kiely & Danielson, 2010), and in our current data set of instructional video captured across more than 40 special education classrooms nationally, we have very few observations that include math or writing instruction that aligns with current EBPs in these areas.

Individualization. A defining characteristic of special education is that SWD have learning needs that are substantially different from those of general education students (Cook & Schirmer, 2003; Fuchs & Fuchs, 1994). Although a hallmark of instruction for SWD, individualized intervention has been seriously understudied (Vaughn, Denton & Fletcher, 2010). Individualization is conceptualized somewhat differently across the research, making the organization and specification of rubric criteria difficult. Instructional grouping (mostly related to size of instructional group), frequency and duration, and aligning the focal areas to student

needs (e.g. focus on phonological awareness and decoding for students with dyslexia) are the primary ways that researchers have described ways to individualize intervention (Vaughn et al., 2010). However, an emerging evidence base examining treatment by aptitude effects indicates that when specific types of instructional practices are aligned with student profiles based on cognitive or information processing evaluation, treatment by aptitude interactions can be significant (Fuchs et al., 2014).

Not only is the construct of individualization defined across a number of variables in the research, it is also difficult to observe without having information about each student's specific needs. We might be able to observe a teacher differentiating a lesson, and we may be able to identify with a high degree of reliability those students whose needs are not being met by the instructional lesson by their response, but in order to give special education teachers specific feedback about their ability to effectively individualize instruction based on the needs of their student, we will need greater specification about this process, and we will likely need to include evaluation methods that go beyond observation. Assessing individualization will likely require the inclusion of teacher artifacts that help the evaluator understand on what basis the decisions to individualize were made, and how the specific adaptations are expected to meet student needs. While it seems logical that an individualized education plan (IEPs) might provide the type of data to inform this process, reviews of IEPs suggest that they are highly variable in quality and many do not contain sufficient information about the relevance of the IEP goals and instructional plans to the students' needs (La Salle, Roach, & McGrath, 2013). Even if IEPs are the right artifact to help evaluate individualization, is it feasible to collect and review the IEPs of SWD a teacher? Or can teachers provide a brief description of how they engage in this process in a way that we can standardize across observations? If we standardize the process, which dimensions of

individualization (e.g. time, duration, frequency or individualization based on cognitive processing profiles) should be included? We are currently piloting the use of a template in which special education teachers will document how they individualize for their students to determine whether we can capture this critical process through an observation and artifact review process.

Raters and Feedback

One of the promises of observation systems is that they will provide individualized and specific information about a teacher's instructional practice that will promote individual improvement among teachers. Through the observation and feedback loop, teachers could be encouraged to be more self-reflective, to engage in conversations with instructional leaders and fellow teachers about effective practices, and to gain specific information about their own practice that could allow them to improve (Taylor & Tyler, 2012). Most existing observation protocols however, are generic with respect to content area and are designed to be used across all teachers, across all grade levels (Hill & Grossman, 2013). While generic descriptions of instructional practice might lead to greater reliability across raters, they compromise the specificity of the feedback provided and do little to reflect the specialized nature of special education instruction that teachers will need in order to improve their instructional practices (Johnson & Semmelroth, 2014). Assuming the purpose of observation systems is to improve instruction, it is critical that instructional practices are reflected to a level of detail that will allow special education teachers to respond (Grossman, Compton, Igra, Ronfeldt, Shahan, & Williamson, 2009). Greater specification in rating systems however, will also require raters with a deep understanding of the EBPs they are observing.

In the structure of schools, principals are typically in a position to evaluate their teaching staff, and in fact, the Great Teacher and Leaders Center recommend that principals are involved

in the evaluation of their staff (Holdheide, 2013). But principals cannot have expertise in all subject areas and surveys indicate that most principals do not have the specialized knowledge required to reliably and effectively evaluate and provide feedback to special education teachers (Frost & Kersten, 2011). Rater expertise is critical. Studies examining the differences in results across raters indicate that administrators differentiate more among teachers than peer raters, and that the reliability of ratings is compromised when only one observer participates (Ho & Kane, 2012). Numerous studies examining the reliability of observation systems have indicated that a minimum of three raters and three observations of a teacher are required to achieve acceptable levels of reliability (Hill, Charalambous, & Kraft, 2012; Johnson & Semmelroth, 2015; Kane & Staiger, 2012). These requirements pose significant implementation challenges. The reliability and validity of the evaluation may be compromised if raters who are not special education experts are used, however it may not be feasible to adhere to the findings regarding rater qualifications currently reported in the research.

With regard to feedback, the process by which this is delivered to special education teachers will be important. Promising results from studies of feedback and coaching based on teacher observations have demonstrated that both can positively affect student outcomes (Allen, Pianta, Gregory, Mikami & Lun, 2011; Taylor & Tyler, 2012). This is encouraging; however, the process of providing feedback will require administrative and logistical support to ensure that the coaching component of supporting special education teachers in improving practice is not compromised given the competing demands of schools. Some districts have created systems in which highly effective teachers are temporarily removed from the classroom to serve as instructional coaches to provide feedback to teachers (Steinberg & Sartain, 2015). This model may not make sense in a field like special education which suffers from critical shortages,

making the removal of effective special education teachers from the classroom a difficult decision to rationalize.

As we continue with the implementation of RESET we will need to test and refine many aspects related to the qualification and training of raters, as well as towards developing a greater understanding of the process of providing feedback. Simply providing teachers with observational data is not sufficient to change behavior (Crawford et al., Joyce & Showers, 2002). Through RESET's observation and feedback loop, special education teachers are expected to improve their ability to implement EBPs. For this feedback loop to effectively support instructional change, the results from an observation using the RESET protocol must yield reliable information that is explicit enough for teachers to understand what changes they need to make to effectively implement EBP (Crawford, et al, 2013), and this will require raters who are skilled in the practices they are observing who can provide meaningful feedback across a fairly wide range of content areas and EBPs.

Alignment to Student Outcomes

The purpose of any teacher evaluation system is to improve outcomes for students. Yet observation systems, even those that are extensive and comprehensive, have reported low to moderate correlations with student outcomes (Connor et al., 2014; Kane & Staiger, 2012). For example, Connor et al. (2014) conducted a comprehensive evaluation of literacy instruction, examining instructional practices, classroom contexts, and content. Their evaluations took an average of 85-90 minutes to complete, using highly trained raters, and yet the correlations of teacher evaluation to student outcomes were reported as low to moderate. Similarly, Kane and Staiger (2012) reported low to moderate correlations of teacher practice to student outcomes in the Measures of Effective Teaching (MET) study. A variety of explanations for the low

correlations have been offered, but ultimately the weak reported relationship suggests how challenging it can be to develop a system that captures the elements of instructional practice that have the strongest effect on student outcomes.

For special education teacher evaluation, an additional concern is to integrate outcome measures that are more relevant and sensitive to changes in student performance than state standardized assessments. In the development of RESET, we are examining a variety of standardized measures that are widely used to assess meaningful outcomes for students with disabilities. To develop a common metric and means of investigating the impact of special education teacher performance on student outcomes, we plan to convert student performance to effect size measures, and then determine which elements of each EBP most strongly predict changes in student performance. Through this process, our goal is to develop an observation tool that focuses special education teachers on the implementation of EBPs that positively impact student growth.

Conclusions

The challenges of special education teacher evaluation through observation systems are significant, but students with disabilities need and deserve access to high quality instruction. An observation system that is focused on supporting special education teachers' implementation of EBPs has the potential to improve educational opportunities for students with disabilities. The design of such a system requires a theoretical framework that aligns well with the research on best practices for SWD, which suggests that instructional practice, content, and individualization assessed by meaningful student outcome measures are critical elements of effective special education. There are a number of challenges to be addressed in the design, and certainly in the

implementation of such a system, but if we can successfully navigate these challenges, we hope to improve practice and ultimately, improve outcomes for students with disabilities.

References

- AFIRM Team. (2015). *Visual supports*. Chapel Hill, NC: National Professional Development Center on Autism Spectrum Disorders, FPG Child Development Center, University of North Carolina. Retrieved from <http://afirm.fpg.unc.edu/visual-supports>
- Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A. Y., & Lun, J. (2011). An interaction-based approach to enhancing secondary school instruction and student achievement. *Science*, 333(6045), 1034-1037.
- Archer, A. L., & Hughes, C. A. (2010). *Explicit instruction: Effective and efficient teaching*. Guilford Press:
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Ravitch, D., et al. (2010). Problems with the use of student test scores to evaluate teachers. *Economic Policy* (Vol. 278, pp. 1–29). Economic Policy Institute. Retrieved from http://epi.3cdn.net/b9667271ee6c154195_t9m6iij8k.pdf
- Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *The Elementary School Journal*, 111, (1), 7-34.

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

Boe, Erling, E., Cook, Lynne, H., & Sunderland, Robert, J. (2008). Teacher turnover: Examining exit attrition, teaching area transfer, and school migration. *Exceptional Children*, 75(I), 7–31.

Brophy, J., & Good, T.L. (1986). Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.). New York: Macmillan.

Brownell, M. T., Sindelar, P. T., Kiely, M. T., & Danielson, L. C. (2010). Special education teacher quality and preparation: Exposing foundations, constructing a new model. *Exceptional Children*, 76 (3), 357-377.

Carnine, D. W., Silbert, J., Kameenui, E. J., & Tarver, S. G. (2010). *Direct Instruction Reading, Fifth Edition*. Merrill: Boston, MA.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Whitmore Schanzenbach, D., & Gayan, D., (2011). How does your kindergarten classroom affect your earnings? Evidence from Project Star, *The Quarterly Journal of Economics*, 126(4), pages 1593-1660.

Christenson, S. L., Ysseldyke, J. E., & Thurlow, M. L. (1989). Critical instructional factors for students with mild handicaps: An integrative review. *Remedial and Special Education*, 10(5), 21-31.

- Connelly, V., & Graham, S. (2009). Student teaching and teacher attrition in special education. *Teacher Education and Special Education, 32*(3), 257–269.
doi:10.1177/0888406409339472
- Connor, C. M., Morrison, F. J., Schatschneider, C., Toste, J., Lundblom, E., Crowe, E. C., & Fishman, B. (2011). Effective Classroom Instruction: Implications of Child Characteristics by Reading Instruction Interactions on First Graders' Word Reading Achievement. *Journal of Research on Educational Effectiveness, 4*(3), 173–207.
<http://doi.org/10.1080/19345747.2010.510179>
- Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., McLean, L., & Morrison, F. J. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. *Journal of Educational Psychology, 106*(3), 762-778.
- Cook, B. G., & Schirmer, B. R. (2003). What is special about special education? Overview and analysis. *Journal of Special Education, 37*, 200-205.
- Cook, B. G., Tankersley, M., & Landrum, T. J. (2009). Determining evidence-based practices in special education. *Exceptional Children, 75*(3), 365–383.
- Council for Exceptional Children (2014). *Standards for evidence-based practices in special education*. Council for Exceptional Children: Arlington, VA.
- Cortiella, C. (2015). *The state of learning disabilities*. New York, NY: National Center for Learning Disabilities.

- Crawford, A. D., Zucker, T. A., Williams, J. M, Bhavsar, V., & Landry, S. H. (2013). Initial validation of the prekindergarten classroom observation tool and goal setting system for data-based coaching. *School Psychology Quarterly*, doi: 10.1037/spq0000033.
- Delquadri, J., Greenwood, C. R., Whorton, D., Carta, J. J., & Hall, R. V. (1986). Classwide peer tutoring. *Exceptional Children*, 52, 535-542.
- Derrington, M. L. & Campbell, J. W. (2015). Implementing new teacher evaluation systems: Principals' concerns and supervisor support. *Journal of Educational Change*, 16 (3), 305-26.
- Frost, L. A., & Kersten, T. (2011). The role of the elementary principal in the instructional leadership of special education. *International Journal of Educational Leadership Preparation*, 6 (2), 1-21.
- Fuchs, D. & Fuchs, L. S. (1994). What's special about special education? A field under siege. ED 379 817
- Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34, 174-206.
- Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., et al. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, 106 (2), 499-514.
- Gersten, R., Schiller, E. P., & Vaughn, S. (2000). *Contemporary special education research*. Mahwah, NJ: Erlbaum.

- Gersten, R., Vaughn, S., Deshler, D., & Schiller, E. (1997). What we know about using research findings: Implications for improving special education practice. *Journal of Learning Disabilities, 30*(5), 466–76. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9293227>
- Goldhaber, D. (2016). In schools, teacher quality matters most. *Education Next, 16*(2), 56-62.
- Graham, S., & Harris, K. R. (1989). Components analysis of cognitive strategy instruction: Effects on learning disabled students' compositions and self-efficacy. *Journal of Educational Psychology, 81*, 353-361.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record, 111*(9), 2055-2100.
- Harn, B. A., Parisi, D., & Stoolmiller, M. (2013). Balancing fit and flexibility with fidelity: What do we really know about “fidelity of implementation” in schools? *Exceptional Children, 79*, 181-194.
- Hill, H., Charalambous, C., & Kraft, M. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*, 56-64.
- Hill, H., & Grossman P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371-384.
- Ho, A. D., & Kane, T. J. (2012). *The reliability of classroom observations by school personnel.*

- Seattle, WA: Bill & Melinda Gates Foundation, Retrieved from
<http://k12education.gatesfoundation.org/teacher-supports/teacher-development/measuring-effective-teaching/why-met-additional-resources/>
- Holdheide, L. (2013). *Inclusive design: Building educator evaluation systems that support students with disabilities*. Available from www.gtlcenter.org
- Johnson E. S., & Boyd, L. (2013). Designing effective Tier 2 reading instruction in early elementary grades with limited resources. *Intervention in School and Clinic*, 48 (4), 203-209.
- Johnson, E. S., Ford, J. W., Crawford, A., & Moylan, L. A. (2016). *Recognizing Effective Special Education Teachers (RESET) Technical Manual*. Boise State University, Boise, ID.
- Johnson, E. S. & Semmelroth, C. L. (2015). Validating an Observation Protocol to Measure Special Education Teacher Effectiveness, *Journal of the American Academy of Special Education Professionals*
- Johnson, E. S., & Semmelroth, C. L. (2014). Special education teacher evaluation: Why it matters and what makes it challenging. *Assessment for Effective Intervention*. 39 (2), 71-82.
- Jones, N. D. & Brownell, M. T. (2014). Examining the use of classroom observations in the evaluation of special education teachers. *Assessment for Effective Intervention*, 39 (2), 112-124.
- Joyce, B., & Showers, B. (2002). *Student achievement through staff development (3rd ed.)*. Alexandria, VG: ASCD.
- Kane, T. J., & Staiger, D. O. (2012). *Gathering feedback for teaching: Combining high-quality*

observations with student surveys and achievement gains. Retrieved from

http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf

LaSalle, T. P., Roach, A. T., & McGrath, D. (2013). The relationship of IEP quality to curricular access and academic achievement of students with disabilities. *International Journal of Special Education*, 28(1), 135-144.

Mathes, P.G., Howard, J.K., Allen, S.H., & Fuchs, D. (1998). Peer-assisted learning strategies for first-grade readers: Responding to the needs of diverse learners. *Reading Research Quarterly*, 33, 62-94.

McMaster, K.L., Fuchs, D., & Fuchs, L.S. (2007). Promises and limitations of Peer-Assisted Learning Strategies in reading. *Learning Disabilities: A Contemporary Journal*, 5, (2), 97-112.

Moats, L. C. & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia*, 53, 23-41.

Montague, M. (1992). The effects of cognitive and metacognitive strategy instruction on the mathematical problem solving of middle school students with learning disabilities. *Journal of Learning Disabilities*, 25, 230-248.

Montague, M., & Dietz, S. (2009). Evaluating the evidence base for cognitive strategy instruction and mathematical problem solving. *Exceptional Children*, 75,(3), 285-302.

National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel: Teaching children to read – An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.*

Bethesda, MD: NICHD, National Institutes of Health.

Odom, S. L. (2009). The tie that binds: Evidence-based practice, implementation science, and outcomes for children. *Topics in Early Childhood Special Education, 29*(1), 53–61.

Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*(2), 137–148.

Odom, S. L., Fleming, S., Diamond, K., Lieber, J., Hanson, M., Butera, G., Horn, E., Palmer, S., Marquis, J., & Children's School Success Project, (2010). Examining different forms of implementation and in early childhood curriculum research. *Early Childhood Research Quarterly, 25*(3), 314-328. Doi:10.1016/j.ecresq.2010.03.001.

Rosenshine, B., & Meister, C. E. (1993). *Reciprocal Teaching: A review of 19 experimental studies. Technical Report Number 574*. Center for the Study of Reading, University of Illinois: Champaign, Illinois

Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M.C. Wittrock (Ed), *Handbook of research on teaching, 3rd ed* (pp. 376-391). New York: Macmillan.

Steinberg, M. & Sartain, L. (2015). Does better observation make better teachers? *Education Next, 15*(1), 70-76.

Swanson, H. L., (1999). Instructional components that predict treatment outcomes for students with learning disabilities: Support for a combined strategy and direct instruction model. *Learning Disabilities Research & Practice, 14*, 129-140.

Swanson, H. L., & Deshler, D. (2003). Instructing adolescents with learning disabilities:

Converting a meta-analysis to practice. *Journal of Learning Disabilities, 26*(2) 124-135.

Swanson, H. L., & Hoskyn, M. (2001). A meta-analysis of intervention research for adolescent students with learning disabilities. *Learning Disabilities Research & Practice, 16*, 109-119.

Swanson, H. L., Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities, 33*(2), 114–136.

Taylor, E. S. & Tyler, J. H. (2012). Can teacher evaluation improve teaching? Evidence of systematic growth in the effectiveness of midcareer teachers. *Education Next, 12*(4), 78-84.

U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, March). *What Works Clearinghouse: Procedures and Standards Handbook (Version 3.0)*. Retrieved from <http://whatworks.ed.gov>

U.S. Department of Education (2015). *Teacher shortage area nationwide list*. Retrieved from <http://www2.ed.gov/about/offices/list/ope/pol/tsa.doc>

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in Schools, 47* (5), 432-444.

Vaughn, S. & Swanson, E. A. (2015). Special education research advances knowledge in education. *Exceptional Children, 82*(1), 11-24. doi: 10.1177/0014402915598781

