

**A REVIEW OF SPECIAL EDUCATION TEACHER
(SET) EVALUATION PRACTICES**

*Jacob A. Tandy
Denise K. Whitford
Marilyn A. Hirth
Purdue University*

Abstract

Federal accountability for teacher evaluation is left up to interpretation by the states, who in turn often leave these decisions to individual districts. While there are no standard practices for how to evaluate special education teachers (SETs), there are multiple influences on how it should be done. With no commonly accepted evaluation used in practice, and relatively little research in the area, the attention of this article turns to the research literature in order to 1) shed light on best practices in SET evaluation and 2) provide a comprehensive conceptual model for special education teacher evaluation.

INTRODUCTION

There has been increased focus on teacher quality as multiple studies have found that teacher quality is among one of the most important school-based factors on student achievement (Goldhaber, 2010). In order to improve the quality of all teachers and administrators, policy makers took steps to require that teachers and administrators are “highly qualified” through the No Child Left Behind Act (NCLB, 2001), as well as placed emphasis on assessments to determine Adequate Yearly Progress (AYP) and the goal for all students to be 100% proficient in reading by 2013-14. Along with NCLB, the Individuals with Disabilities Education Act (IDEA, 2004) was reauthorized to align with the provisions in NCLB to include high academic standards for students with disabilities, along with the expectation and accountability of including students with disabilities in district and state-wide assessments, when appropriate and per students’ Individualized Education Plans (IEP), to be counted toward a school’s AYP. These legislations, as well as those introduced in more recent years (e.g., Race to the Top, the Elementary and Secondary Education Act, and the Every Student Succeeds Act), have put an emphasis on these assessments as evidence of student learning, and by extension, teacher quality. The issue, particularly in special education, is that there are a number of other responsibilities (i.e., writing and implementing IEPs, as well as collaborating with families, service providers and general educators) for special education teachers (SETs) beyond typical academic instruction (Benedict, Thomas, Kimerling, & Leko, 2013).

Special Education Teacher Evaluation (SET)

An issue to consider in evaluating SETs is that there are no set standards in practice (Brownell & Jones, 2015; Holdheide, 2015). Holdheide, Goe, Croft, and Reschly (2010) note that most states leave the evaluation of SETs up to individual districts, but recognized that Race to the Top has caused many states to be proactive in implementing rigorous evaluations that include student achievement. Accountability, as federal law dictates, is left to interpretation by individual states to make these decisions on how to create and implement evaluations for all teachers, including SETs. States often leave these decisions to individual districts themselves, and the ways in which evaluations are conducted are wide-ranging, including summative and formative measures, as from which data are both tangible and intangible aspects of teaching (National Council on Teacher Quality, 2008). Thus, there are no standard practices for how to evaluate SETs, but there are multiple influences on how this should be done, with a particular focus on student data from academic and functional assessments per federal law. The way in which this information is gathered however, is done from a variety of methods, from classroom observations, peer review, portfolios, value-added modeling, a combination of any of these or even other methods (Benedict et al., 2013).

Holdheide (2015) reviewed a number of different state and district models (Colorado, Massachusetts, Ohio, Pennsylvania, Bartholomew Consolidated School Corporation, Columbus, IN, and the Danielson Framework for Teaching) of SET evaluation, and identified five themes of SET evaluation that positively affect student learning and recognize the unique context and roles of SETs. The themes: recognition that all teachers support the learning of students with disabilities, promotion and reinforcement of the use of evidence-based practices, consideration

for differing measurement of student growth, attention to specific roles and responsibilities of SETs, and a focus on the role of leaders are all supported by evidence of either state policy and/or district practice. These exemplar practices are positive, but only represent a small portion of the policy and practice in this country, and as Brownell & Jones (2015) pointed out, there is little research base in SET evaluation on which to base policy and practice. Unlike the more robust research on general education teacher evaluation and teacher effectiveness literature, districts and states have little data on which to base decisions when developing their SET evaluations. There is, therefore, great diversity in policy and practice in regards to SET evaluations, with many states opting to leave these decisions to individual districts (Brownell & Jones, 2015). Without a firm research base to make these decisions and no standard practices in place, there needs to be guidance from another source.

SET guidelines. In 2013, the Council for Exceptional Children (CEC) published a position paper on special education teacher evaluation. The purpose of this paper was to set guidelines and recommend practices for quality special education teacher evaluation. As the leading voice for special education professionals, CEC aimed to use this platform to influence practices and policy at the local, state and federal level. In short, the position paper recommends that special education teachers be evaluated on one system that promotes effective teaching, encourages collaboration and professionalism, is research-based, and is differentiated for each particular educator's role and performance standards (CEC, 2013).

Rationale

As discussed above, there is no mandatory method in which SETs are evaluated nationally, with many states allowing districts to implement their own evaluation systems. With no commonly accepted evaluation used in practice, and relatively little research in the area, the attention of this study turns to the research literature in order to 1) shed light on best practices in SET evaluation and 2) provide a comprehensive conceptual model for special education teacher evaluation, with the overall goal to promote high teacher quality in special education, improve student achievement, foster professional development, and improve the role of special education teachers within the teaching profession.

METHODOLOGY

Inclusion and Exclusion Criteria

Given that the focus of this literature review was to examine research-based practices in SET evaluation, specific inclusion and exclusion criteria were necessary. First, only publications in peer reviewed, English academic journals were considered. Second, the publication had to address the issue of SET evaluation, either explicitly or by examining a practice for professional development (PD) that is often used in the evaluation process. Additionally, the participants or population in focus must be SETs, meaning the teachers discussed in the publication are licensed and teach students who qualify for special education services, as defined by IDEA. Thus, publications specifically providing interventions or only discussing pre-service teacher preparation and/or professional development did not meet inclusion criteria because these individuals are not yet licensed SETs. Naturally, this allowed for the inclusion of a great variety of students with disabilities and eligibility categories, but this was necessary because it allowed

us to illustrate the broad range of responsibilities and skills necessary for SETs, as well as the unique factors to necessary to evaluate these educators.

Third, only publications from 2010 through 2016 were considered, in order to provide recent research on practices of SET evaluation. Moreover, the 2010 date corresponds to research published after Holdheide et al. (2010), who addressed the challenges of evaluating SETs and teachers of English Language Learners (ELLs). Lastly, publications had to conduct research, or describe research, related to SET evaluation. Explicit editorials were not included in this literature review.

LITERATURE REVIEW

The first author searched the following databases to locate articles for this review: a) *PsychInfo*, b) *Education Resources Information Center*, c) *H.W. Wilson*, and d) *Education Source*. Searches included a combination of the terms *teacher evaluation*, *evaluation*, *special education*, and *accountability*. The four databases searches yielded, 118, 82, 56, and 82 results, respectively, with a total of 9 articles that met inclusion criteria. Lastly, a hand search of the *Journal of Special Education Leadership* was conducted and resulted in one publication that met criteria. After conducting these searches, 10 studies met inclusion criteria for review. It is important to note, that this was not a systematic search of the literature, and as such, this review is not systematic.

Coding Procedures

All 10 publications were coded for the grade level taught by the included SETs, the eligibility categories of students taught by the SETs, the type of evaluation(s) discussed, and the result of the evaluation method, or the recommendation of the author(s) on the method(s) of evaluation to use. Ideally, these procedures demonstrate the best evaluation practices for SETs, with consideration to student grade level and disabilities.

FINDINGS

In order to better understand the literature in the area of special education teacher evaluation, general demographic information related to the publications, as well as descriptions of the type of publication are necessary. Furthermore, the commonalities of these publications are addressed, namely the need for multiple sources of data in evaluation, issues with standardized tests to determine teacher effectiveness, the positive effects of a mentorship program (especially for beginning teachers), and the use of ongoing formative evaluation to facilitate growth in SETs.

Table 1

Characteristics of Studies Included in Review

Study	K-12	Disability/ Eligibility Categories of Students	Type of Evaluation	Results/Recommendations
Dingle, Brownell, Leko, Boardman, & Haager (2011)	3-5	ADHD, ASD, EBD, LD, and mild intellectual disabilities.	Monthly self- reflection and feedback from an instructional coach based on taped lessons	There needs to be customization to the specific needs of each teacher to facilitate growth. Self-reflection was helpful for teachers in identifying areas for improvement and then implementing changes.
Fall (2010)	K-12	All	Peer- evaluation, mentoring, and self- reflection	With a focus on high poverty schools, the author suggests more support, particularly for new SETs through collaborative mentor teachers to help navigate challenges. Self-evaluation can help prepare a SET for collaborative sessions.
Hartmann (2012)	K-12	Visual impairments, deaf or hard of hearing, and students with moderate to severe disabilities	Self- reflection through the Teacher Efficacy in Deaf- Blindness Education Scale (TEDE)	Self-efficacy, as measured by the TEDE, is more important than content knowledge and skills in order to achieve desired outcomes.
Israel, Kamman, McCray, & Sindelar	K-12	ADHD, ASD, EBD, LD, and moderate and severe disabilities	Mentor evaluation through classroom observation (two informal, two formal)	The mentors provided both emotional and professional support during the process, and this was not hindered by the evaluator/teacher relationship.

Johnson (2015)	K-12	All	Instructional observation	Observations should focus on the use of evidence-based practices in the classroom, and teachers need ongoing support to improve in their implementation of evidence-based practices.
Kearns, Kleinert, Thurlow, Gong, & Quenemoen (2015)	K-12	Significant intellectual disabilities	Use of Alternate Assessments to evaluate effectiveness	Student achievement data must be paired with ongoing professional development for teachers to help them improve as teachers and to provide curricular opportunities for their students.
Kersting, Chen, & Stigler (2013)	3-5	ADHD, ASD, EBD, and LD	Value-Added Modeling (VAM)	VAM is a reliable way to measure teacher impact across all groups of students, including students with disabilities
Martin (2012)	K-12	All	High stakes student test scores	Using high stakes testing alone is problematic because it does not account for individual differences of students. These data should be used formatively, not summative.
Pogrud & Cowan (2013)	K-12	Visual Impairments	Mentor evaluation through state mandated programing	Teachers reported that the mentor program helped them to improve their practices and adjust well to their positions. Accountability and frequency of mentor/protégé contact was a strength.
Steinbrecher, Selig, Cosby, & Thorstensen (2014)	K-12	All	Value-Added Modeling (VAM)	Rejects use of VAM based on the assumption that standardized tests are valid indicators of learning, and no consideration of individualization of instruction.

Descriptive Information

Many of the studies included in this review are descriptive in nature, examining a specific issue related to SET evaluation and providing some recommendations based on the research literature. These publications generally discuss K-12 SETs and how different evaluation practices will generally impact the field. Other publications, some of them experimental or quasi-experimental, discuss specific SETs and the grade levels and disabilities of their students. Dingle, Brownell, Leko, Boardman, & Haager (2011) and Kersting, Chen, & Stigler (2013) focused on teachers of 3rd through 5th graders who had a wide range of high incidence disabilities (i.e., Attention Deficit-Hyperactivity Disorder (ADHD), Autism Spectrum Disorder (ASD), Emotional and Behavioral Disorders (EBD), Learning Disabilities (LD), and mild intellectual disabilities). In Israel, Kamman, McCray, & Sindelar (2014), a new teacher mentorship program described SETs in the K-12 education of students with a wide range of disabilities (ADHD, ASD, EBD, LD, and moderate and severe disabilities). Lastly, Hartmann (2012) and Pogrud and Cowan (2013) specifically addressed K-12 teachers who taught students with visual impairments along with students who are deaf or hard of hearing and some students with moderate to severe disabilities, in addition to their visual impairment.

Multiple Sources of Evidence for Teacher Quality

An ongoing commonality among the literature is that there should be multiple sources of evidence for teacher effectiveness that factor into an evaluation. An interesting finding related to the sentiment for multiple sources involves the controversial use of Value Added Modeling (VAM; Kersting et al., 2013; Steinbrecher, Selig, Cosbey, & Thorstensen, 2014). This method involves the use of high stakes student test scores compounded over time to determine trend data, as well as the effect individual teachers have on the growth of a student. One could argue that a VAM approach uses two sources, the high stakes test score, and then the VAM data, but since this approach is highly dependent on high stakes testing, it presents some issues. Kersting et al. (2013) argued that VAM is a reliable way to measure teacher effectiveness based on the statistical tests, but Steinbrecher et al. (2014) strongly disagree, based on the reliance upon the state-based high stakes testing and the lack of consideration on individual differences that are inherent within special education students and settings.

Also, Martin (2012) argued against the use of high stakes test scores as the sole, or even primary source, of SET evaluation. The main argument Martin (2012) presented, just as in Steinbrecher et al. (2014), is that individual differences of students are not reflected in the testing data. Martin (2012) contended that SETs should not have summative evaluations solely based on high stakes testing scores, but that these data should be used formatively to inform practice and to be responsive to individual students' needs.

High Stakes Testing to Determine Effectiveness

In addition to the above discussion about the issues with utilizing only one primary source of evaluation, and how high stakes test scores are often that one source, there are some other issues with high stakes testing in the evaluation process. As students with disabilities are expected to participate in these tests, per IDEA, it is expected that many students with disabilities will not perform as well as their peers without disabilities (Martin, 2012). Some students with disabilities can receive access to accommodations on these tests as part of their IEP, but even then, many of these students will demonstrate that they struggle with high stakes tests and be the “floor” of the data (Steinbrecher et al., 2014). Also, each state uses its own form of high stakes testing and Steinbrecher et al. (2014) asserted that these tests are not necessarily valid indicators of learning, and there is the issue that they typically only occur one time per school year. Additionally, being timed tests, they are not necessarily assessing students’ abilities to simply complete the required tasks, but to complete these tasks quickly.

Some students with disabilities are not required to take the high stakes test required in their state because the student’s educational needs, as determined by the his or her IEP team, are severe enough that an alternate assessment is necessary to measure the child’s academic and or functional skills (Kearns, Kleinert, Thurlow, Gong, & Quenemoen, 2015). Even for SETs who teach students with severe disabilities, there must be ongoing professional development and support to help facilitate the learning that is being measured in alternate assessments. As such, a high stakes test, whether the general or alternate test, should not be the sole evidence of teacher effectiveness for special education teachers. SETs have to provide individualized instruction that meets the needs of a child in academic and functional performance, per IDEA, which means that a high stakes test that may only measure reading, writing and math skills may not be reflective of the learning taking place for a student based on other needs outlined in the IEP (i.e., social skills, behavior challenges, transition goals, etc.).

Mentorship Programs

Dingle et al. (2011) made an argument that SET evaluations should be customized, using a broad range of sources, to the individual needs of the SETs to facilitate growth. Dingle et al. (2011) paired new SETs with experienced mentors who worked with the new teachers to evaluate their video recorded lessons over the course of a school year. In addition to these sessions, the new teachers engaged in self-reflection that also factored into the process. Fall (2010), Israel et al. (2014), and Pogrund and Cowan (2013) employed mentorship programs in which mentors worked with new teachers to help them improve practices, adjust to the myriad of responsibilities in teaching, and provide evaluative and ongoing feedback to SETs. These publications all suggest that having a strong mentor, with clear guidelines, regular meetings and accountability helps to provide more support to new teachers, which will help them adjust to the profession, improve as a teacher, and hopefully successfully remain in the field. Interestingly, Pogrund and Cowan (2013) assessed the impact of a mandatory mentorship program in Texas for teachers working with students with visual impairments. They demonstrated that mentorship experiences can work for SETs within a particular eligibility category, and that states can implement these programs to support teachers effectively.

Ongoing Formative Evaluation

Eight of the ten reviewed publications employ and/or recommend an ongoing formative evaluation process that focuses on improving teacher practices. All of the other studies discuss the use of ongoing support and professional development, whether through a mentoring/coaching program for teachers (Dingle et al., 2011; Fall, 2010; Israel et al., 2012; Johnson, 2015; Pogrud & Cowan, 2013), or through the use of high stakes testing to inform practice; not just as a method of summative evaluation (Kearns et al., 2015; Martin, 2012; Steinbrecher et al., 2014).

DISCUSSION

The findings of this review have broader connections to the research literature, as well as implications for research and practice. The major points for discussion are a consistent need for ongoing formative evaluations, in addition to using multiple sources of teacher quality for evaluation. These are the overarching findings from the review, and have the biggest implications for the field of special education. Additionally, these findings are further compared to the CEC (2013) position paper on SET evaluation to see how the research literature compares to CEC recommendations for special education evaluation standards. Furthermore, we outline a comprehensive conceptual model for special education teacher evaluation.

Ongoing Formative Evaluation

As explained in the findings, eight of the ten reviewed studies directly dealt with providing SETs with ongoing formative evaluation, either through an experiment, providing a recommendation for practice based on research, or in reaction to what authors think are ineffective current practices in SET evaluation. Fall (2010) provided a great deal of discussion on the need for recruiting and retaining highly qualified SETs in high poverty schools. The author pointed out that there has been a shortage of highly qualified SETs, particularly in high poverty schools, but it is a recognized issue nationally as well. There is a growing concern that there are inadequate numbers of highly qualified SETs to serve students with disabilities, and that this is especially troubling because quality teachers are one of the most important school based factors in student achievement (Fall, 2010; Israel et al., 2014).

Due to concerns regarding the recruitment and retention of highly qualified and high quality SETs, one approach revealed in this literature review is the practice of providing ongoing formative evaluation and support for SETs in order to improve their practices. One way in which SETs can be provided this support, particularly for new, inexperienced teachers, is through a mentorship program. The studies that discussed a mentorship program (Dingle et al., 2011; Fall, 2010; Israel et al., 2014; Pogrud & Cowan, 2013), all reported that it was an effective way to support SETs and provide ongoing formative evaluation in a positive, collaborative way. An interesting finding in Israel et al. (2014) was that the evaluator/teacher relationship between the mentor and the protégé was not a hindrance to their collaborative efforts, and that the protégé teachers were very comfortable working with their mentor to improve their practices, even though the mentor was evaluating them periodically. While there is no doubt in the existence of a power dynamic between an evaluator and a teacher, Israel et al. (2014) determined that in an ongoing, collaborative setting in which the improvement of the teacher was the priority, both the evaluator and the teacher were able to work together effectively to achieve positive results. In the position paper presented by CEC (2013), it is clear that a

structured, collaborative integration of professional development and evaluation is important in evaluating SETs. By providing regular feedback to teachers and providing them opportunities and resources to improve in instrumental areas and skills, the evaluation process is much more focused on growth, which can then enhance the role of special education teachers within the teaching profession, and may subsequently have positive impacts on teacher retention through such support.

Multiple Sources of Evidence for Teacher Quality

Along with the need for more ongoing formative evaluation and feedback, there should be multiple sources of teacher quality used in the SET evaluation process. As explained above, there are many ways in which SETs are evaluated; some of them are primarily from one piece of data (i.e., VAM, high stakes tests, and classroom observations). Both Martin (2012) and Steinbrecher et al. (2014) argued against using high stakes tests and the data associated with such tests as the primary, or only source, of evidence for teacher quality. Their primary argument is that all students, particularly students with disabilities, have individual differences and challenges that could be inaccurately or unfairly represented by only considering test scores for teacher evaluation. Not only do students with disabilities have individual differences, but per IDEA, SETs must meet the functional and academic needs of these students through goals, services, accommodations, and more within the students' IEPs (Brownell & Jones, 2015; Holdheide, 2015). The general argument is that SETs have an inherently different and multi-faceted job compared to general education teachers, and as such, evaluation for SETs should look different and consider these different factors and responsibilities.

This sentiment is strongly supported by Dingle et al. (2011) who posited that evaluation, and the professional development associated with the evaluation, must be tailored to the specific needs of individual teachers. Thus, a high stakes test as the tool for evaluation cannot be adequate in specifically addressing and measuring the quality of a SET based on his or her specific needs as a teacher. CEC (2013) parallels this contention, indicating that any one source of teacher effectiveness or quality by itself is not adequate to evaluate a SET. Also, there must be consideration of each teacher's specific role and responsibility when designing and implementing the evaluation for said teacher, because there are a broad range of responsibilities, student needs, and instructional models, among others that must be considered to accurately evaluate a SET.

IMPLICATIONS

Implications for Research

Based on these findings, there are a number of implications that researchers should consider in designing and implementing research with SETs, administrators, and districts in the future. There is a continual need for more research of evaluation practices in relation to special education teachers (Brownell & Jones, 2015). While this is not a systematic review, the limits of ten studies within a five-year range show that there is a need for additional research in the area. Furthermore, only five of these studies were experimental in nature, demonstrating the need for more extensive experimental research in the area to test the effectiveness of practices and accurately report teacher quality and drive professional development for teachers. Additionally, researchers should work on evaluation practices within individual states (i.e., Pogrud & Cowan,

2013) to see how effective these practices have been, and then utilize the information to inform policy and practice.

Implications for Practice

Holdheide et al. (2010; 2015) reported that most states allow individual districts to determine how to evaluate SETs, with some guidelines on what should be assessed. This allows districts, administrators and the teachers within those districts to shape what should be done in evaluating SETs, in consideration of the unique challenges of the profession and the students within those districts. Depending on how a state or district currently evaluates SETs, practitioners should consider implementing a more collaborative, ongoing formative evaluation model.

SET Evaluation Model (SETEM). Based on the information gathered in this literature review, we propose a comprehensive model for special education teacher evaluation. Figure 1 provides a diagram of the Special Education Teacher Evaluation Model (SETEM). First, multiple sources of information need to be utilized for SET evaluation, including the use of biannual teaching observations, random sampling of IEP documents, IEP team participant feedback, student work samples and portfolios, and/or student academic and functional assessments. These multiple sources of information must include an evaluation based on individualization, considering SET students' educational placement and setting, present levels and progress toward IEP goals, as well as abilities and disabilities. Lastly, SETs must be supported through formative evaluations, facilitated through the use mentoring partnerships, collaborative programs, and interactive professional development participation.

Multiple sources of information. By multiple sources, we recommend three or more of the five identified blocks of information. Three or more blocks should be utilized to obtain a more complete picture of teacher performance, with the higher number of blocks corresponding with the most complete picture of performance. First, teaching observations are essential for the betterment of all teachers; however, the general annual review may not be sufficient and a biannual observation may be more conducive to identifying classroom environments and teaching situations across the academic year. Again, more observations, as opposed to less, should allow SETs to refine their skills. These biannual observations should be conducted by an administrator, but informal peer observations should also be encouraged to help new SETs improve their teaching.

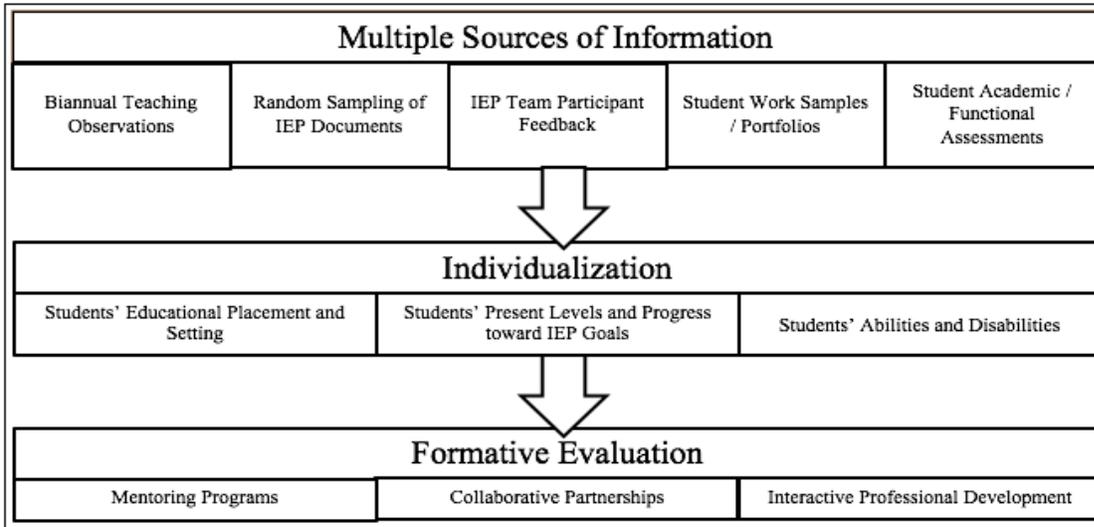
Second, a large portion of the SETs academic year will be focused on completing IEP documents. As a major part of the SET's job, its proper completion should be included in SET evaluations. Additionally, IEPs that are not properly completed are considered out of compliance with federal regulations and can incur legal and financial costs to districts. Third, another area of SET job responsibilities includes the preparation and facilitation of IEP meetings. We recommend that the random sample include a minimum of 10% of the SET's caseload and include at least one initial IEP, one annual IEP, and one triennial IEP. As IEP meetings are essential and can facilitate relationships between SETs and the remaining team members, including parents and guardians, general education teachers, specialized staff, and administration, as well as students, it's important that how they are planned and facilitated are important to SET overall job performance. Asking team members to complete a brief survey regarding their experiences leading up to, during, and after the IEP meeting may demonstrate SET ability to perform this important function within their identified list of responsibilities.

Fourth, student work samples, gathered throughout the time each student was with their SET, and placed in a portfolio, can lend itself to identifying productive work the SET has accomplished with his or her students. Student work samples are an especially useful alternative or supplement to high stakes test results, as they can demonstrate actual ability, instead of simple memorization ability or poor test taking skills. It is important to point out that work samples should reflect progress toward students' IEP goals, as IEP goal attainment is the utmost importance when it comes to teaching for SETs. Fifth, formal academic or functional assessments, including high stakes test results may be useful for some student progress evaluation.

Individualization. As special education is an individualized field, the use of any of the five identified blocks of information should be selected based on the characteristics of each SET's circumstances. First, individualization should consider the educational placement and classroom setting of individual SETs. Special education classrooms can vary on a wide continuum of minimally restrictive environments to considerably restrictive environments, with SET responsibilities drastically different from one end of the continuum to the other. Second, as previously discussed, a major part of SET responsibilities is to assist students in making progress toward their IEP goals. As such, each SET's students' previous present levels of performance, as well as comparisons with current progress toward IEP goals should be considered in SET evaluations. This recommendation should be considered with caution, because every student will make progress toward their IEP goals at a different rate. However, it is important that SETs work with IEP teams to establish goals for individual students that can be realistically met within one year. Third, special education is based on individualized plans because students in special education vary from one to the next. The abilities and disabilities of individual SETs should be considered during the evaluation process.

Formative evaluation. To enable proficient or excellent SET performance it is essential that SETs are fully supported throughout the process of information gathering and identification of individualization needs and rationale. This support should come in the form of formative evaluations, facilitated through mentoring partnerships, collaborative programs, and interactive professional development participation. First, new SETs who are mentored by highly productive veteran SETs, as well as administrators will presumably have many opportunities to receive informative feedback toward annual progress if they have someone they can share their SETEM information and individualization needs and rationale with. These mentoring partnerships will be most beneficial if SETs can be paired based on similar individualization needs and rationale. Second, similar to mentoring programs, collaborative partnerships can foster a support network for all SETs, whether new or veteran. Collaborative partnerships have the ability for SETs to bounce ideas off of each other as they complete work toward their annual evaluation. Third, interactive professional development has the potential to guide SETs in teaching and all of the additional responsibilities they hold. We emphasize the term "interactive", as these professional development opportunities will not be as beneficial if SETs cannot get assistance toward their own professional goals, working toward their students' goals, and working toward a productive evaluation of their job performance.

Figure 1. Special Education Teacher Evaluation Model (SETEM)



Limitations

A major limitation of this study is that it is not a systematic review, but rather a literature review. Therefore, there could be publications that have not been included that may have contributed to better or different findings. Also, there are a limited amount of experimental studies in this review, meaning that although the authors in the descriptive publications based their writings on research, these publications could be more of a reflection of the opinions of those corresponding authors. However, first and third author experiences in special education administration lead to agreement on the need for SET-specific evaluation procedures as well as the methods proposed in the comprehensive model.

CONCLUSION

Without a national standard or identified practice for how SETs are evaluated in K-12 public schools in the United States, CEC (2013) released a position statement to provide a list of best practices to consider and consult when designing and implementing a SET evaluation process. This position, as well as the research literature, underscore a need for a variety of sources of evidence for teacher quality (i.e., classroom observations, peer review, portfolio formation, IEP writing and implementation rubrics, VAM), not just one source such as high stakes testing (Kearns et al., 2015; Martin, 2012; Steinbrecher et al., 2014). Moreover, SETs must be provided with ongoing feedback through formative evaluations that identify areas for improvement and provide opportunities for professional development to address those areas. With a substantial need for highly qualified and high quality SETs, the challenges of teaching students with disabilities, and the high stakes of ensuring a high quality education for all students, there is a great need to review the current evaluation practices of all teachers, particularly SETs, and to collaborate with teachers, districts and administrators to develop and implement a professional and effective evaluation system for SETs, such as the model outlined in this literature review.

REFERENCES

*Studies denoted with an asterisk were included in the literature review.

- 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.
- Brownell, M. M., & Jones, N. N. (2015). Teacher evaluation in special education: Approaches, supporting research, and challenges confronted. *Journal of Special Education Leadership*, 28(2), 63-73.
- Council for Exceptional Children [CEC]. (2013). The Council for Exceptional Children's position on special education teacher evaluation. *Teaching Exceptional Children*, 45(3), 73-76.
- *Dingle, M. P., Brownell, M. T., Leko, M. M., Boardman, A. G., & Haager, D. (2011). Developing effective special education reading teachers: The influence of professional development, context, and individual qualities. *Learning Disability Quarterly*, 34(1), 87-103.
- *Fall, A. M. (2010). Recruiting and retaining highly qualified special education teachers for high-poverty districts and schools: Recommendations for educational leaders. *Journal of Special Education Leadership*, 23(2).
- Goldhaber, D. (2010). *When the stakes are high, can we rely on value-added? Exploring the use of value added models to inform teacher workforce decisions*. Washington, DC: Center for American Progress.
- *Hartmann, E. (2012). A scale to measure teachers' self-efficacy in deaf-blindness education. *Journal of Visual Impairment & Blindness*, 106(11), 728.
- Holdheide, L. G. (2015). Same debate, new opportunity: Designing teacher evaluation systems that promote and support educators in practices that advance all students' learning. *Journal of Special Education Leadership*, 28(2), 74-81.
- Holdheide, L. R., Goe, L., Croft, A., & Reschly, D. J. (2010). Challenges in evaluating special education teachers and English language learner specialists. *Research & Policy Brief*. National Comprehensive Center for Teacher Quality.
- Individuals with Disabilities Education Act [IDEA]. (2004). Public Law 94-142. 20 U.S.C. § 1400 *et seq.* Retrieved from <http://www.copyright.gov/legislation/pl108-446.pdf>
- *Israel, M., Kamman, M. L., McCray, E. D., & Sindelar, P. T. (2014). Mentoring in action: the interplay among professional assistance, emotional support, and evaluation. *Exceptional Children*, 81(1), 45-63.
- *Johnson, E. E. (2015). Integrating special education teachers into a statewide teacher evaluation system. *Journal of Special Education Leadership*, 28(2), 82-88.
- *Kearns, J. F., Kleinert, H. L., Thurlow, M. L., Gong, B., & Quenemoen, R. (2015). Alternate assessments as one measure of teacher effectiveness: Implications for our field. *Research and Practice for Persons with Severe Disabilities*, 40(1), 20-35.
- *Kersting, N. B., Chen, M. K., & Stigler, J. W. (2013). Value-added teacher estimates as part of teacher evaluations: Exploring the effects of data and model specifications on the stability of teacher value-added scores. *Education Policy Analysis Archives*, 21(7), n7.
- *Martin, P. C. (2012). Misuse of high-stakes test scores for evaluative purposes: Neglecting the reality of schools and students. *Current Issues in Education*, 15(3).

- National Council on Teacher Quality [NCTQ]. (2008). *State teacher policy yearbook 2008, national summary*; Washington DC: Author.
- No Child Left Behind [NCLB] Act (2001). Public Law 107-110 § 115, Statute 1425.
- Pogrud, R. L., & Cowan, C. (2013). Perceptions of a statewide mentor program for new itinerant vision professionals. *Journal of Visual Impairment & Blindness*, 107(5), 351.
- *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.

AUTHOR BIOGRAPHIES

Jacob A. Tandy is a PhD student of Special Education at Purdue University. His research interests are the use of technology to improve reading comprehension for students with disabilities, special education teacher education, and the organization and administration of special education.

Denise K. Whitford is an Assistant Professor of Special Education at Purdue University. She specializes in the field of emotional and behavioral disorders (EBD) and her research interests include disproportionality in school discipline and teacher education in cultural awareness and culturally responsive teaching practices.

Marilyn A. Hirth is an Associate Professor in the Department of Educational Studies at Purdue University. She teaches courses in educational policy, school finance/business management, and special education administration. Her research includes studies of educational policy issues, school finance equity and adequacy, and the organization and administration of special education.

PREFERRED CITATION

Tandy, J.A., Whitford, D.K., & Hirth, M.A. (2016). A review of special education teacher (set) evaluation practices. *Journal of Ethical Educational Leadership*, 3(5), 1-16.
Retrieved from: <http://www.cojeel.org>.

JEEL

www.cojeel.org

The views expressed in this publication are not necessarily those of
JEEL's Editorial staff.

JEEL is a free, open-access online journal.

Copyright ©2016 (ISSN 2377-4975)

Permission is hereby granted to copy any article provided that the Journal of Ethical
Educational Leadership is credited and copies are not sold.