Challenges in Evaluating Special Education Teachers and English Language Learner Specialists

Challenges in Evaluating Special Education Teachers and English Language Learner Specialists

This **Research & Policy Brief** addresses the challenges associated with evaluating special education teachers and English language learner specialists, particularly in accurately measuring student achievement growth and connecting that growth to teacher effects.

July 2010

Lynn R. Holdheide, *Vanderbilt University*

Laura Goe, Ph.D., *ETS*

Andrew Croft, *ETS*

Daniel J. Reschly, Ph.D., *Vanderbilt University*

NATIONAL COMPREHENSIVE CENTER
FOR **TEACHER QUALITY**

CONTENTS

OVERVIEW

Current emphasis on teacher effectiveness in educational policy poses a challenge for the evaluation of special education teachers and English language learner (ELL) specialists. Most evaluation systems focus on student achievement and teacher practice; however, few systems have the capacity to differentiate among specialty area educators, address the challenges in accurately measuring achievement growth for their students, and connect that growth to teacher effects. Questions arise as to how these interdependent foci may vary for at-risk populations and how evaluation systems should best reflect this variation. The purpose of this TQ Research & Policy Brief is to inform discussion of these issues.

The National Comprehensive Center for Teacher Quality (TQ Center), with support from the Council for Exceptional Children (CEC) and several national experts, surveyed more than 1,100 state and district directors of special education and interviewed numerous administrators throughout the nation to define the specific challenges in evaluating this population of teachers, determine the current status of state and district policy and practice, and identify promising evaluation practices and instruments. In addition, an analysis of state and district policy for evaluating ELL specialists was conducted primarily through an examination of relevant literature and current practice.

Survey results showed that slightly more than one half of districts (54.7 percent) develop their own evaluation system while approximately one third (36.3 percent) use or adapt the state's recommended evaluation system. At the teacher level, however, the majority of district respondents indicated that contractual agreements do not allow for modification in the evaluation process

for special education teachers. Conversely, nearly one half (49.9 percent) of respondents expressed the opinion that special education teachers should not be evaluated using the same process as that of general education teachers (National Comprehensive Center for Teacher Quality, 2010).

Few state and district respondents cited the use of student achievement data measured by standardized tests or curriculum-based measures in teacher evaluation. As an alternative, other student achievement measures, such as student learning objectives or individualized education program (IEP) goals, have been factored into teacher evaluation. Current teacher evaluation policy and practice is rapidly changing, potentially leading to increased use of these measures. This brief presents special considerations, particularly in the case of coteaching, for reliably using student achievement data to evaluate special education and ELL teachers.

Although many teacher evaluation instruments explicitly address teachers' contributions to meeting the needs of "diverse" learners, they may not consider the special skills and evidence-based instructional methods for students with disabilities and ELLs. A significant body of empirical evidence has identified instructional practices that are linked to improved academic achievement for students with disabilities and ELLs. Without a clear understanding of these special skills and instructional methods, evaluators' capacity to distinguish between effective and ineffective special educators and ELL specialists is limited. Reflecting on the research literature, survey results, and practitioner feedback, this brief offers policy recommendations to support the design and implementation of valid, comprehensive teacher evaluation systems based on teachers' specialties and areas of expertise.

INTRODUCTION

As evidenced in the requirements within the Race to the Top application and *A Blueprint for Reform: The Reauthorization of the Elementary and Secondary Education Act* (U.S. Department of Education, 2010), President Obama and his administration place teacher effectiveness at the heart of education reform, leading to a renewed emphasis on teacher evaluation. Consequently, states and districts are challenged to establish valid evaluation systems that focus on effective teaching and improved student achievement. Although current evaluation systems focus on these factors, questions arise as to how these interdependent foci may vary for teachers of at-risk populations, specifically students with disabilities and ELLs, and how evaluation systems should best reflect this variation.

In response to these questions, the TQ Center, with support from CEC and several national experts, recently surveyed more than 1,100 state and district directors of special education and interviewed numerous administrators throughout the nation to define the specific challenges in evaluating special educators, determine the current status of state and district policy and practice, and identify promising evaluation practices and instruments. In addition, a review of state and district policy for evaluating ELL specialists was conducted primarily through an examination of relevant literature and current practice. This TQ Research & Policy Brief provides the results of this inquiry in the context of current research and practice in teacher evaluation. The brief concludes with policy and practice recommendations for regions, states, and districts to assist in their efforts to create valid, reliable, and comprehensive evaluation systems for all teachers as they work to improve the achievement of all students. The recommendations provided within this brief hold value for practitioners at all levels and, in some respects, teacher educators charged with preparing educators.

THE POLICY AND PRACTICE CONTEXT

Recognizing and accounting for the unique contributions and roles of special education teachers and ELL specialists are critical whether states have designed or are in the process of designing evaluation systems using value-added measures based on student achievement scores, classroom observations, portfolios, and/or self-assessments. The American Recovery and Reinvestment Act (ARRA) of 2009 created an immediate need for guidance in evaluating specialty area teachers so that states and districts can develop and implement strong, valid, and reliable teacher evaluation policies that recognize and promote the unique contributions of these specialty area teachers.

Although designing and implementing teacher evaluation has been mostly left to districts, many states play a role through the creation of teacher evaluation rules and policies, though the degree and scope of involvement varies widely across states (National Council on Teacher Quality, 2008). However, current federal priorities and guidelines for various funding programs indicate an elevation of state-level involvement and accountability. Efforts within states and districts to develop systems to evaluate, compensate, and remediate education professionals are wide ranging—including both summative and formative measures and tangible and intangible aspects of teaching (National Council on Teacher Quality, 2008). Research suggests that evaluation systems should link to student growth, utilize multiple measures to gauge teacher performance, and align with teaching standards as well as professional development (Bill & Melinda Gates Foundation, 2010; Blanton, Sindelar, & Correa, 2006; Little, 2009).

Determining what constitutes effective teaching requires a thorough understanding of what an effective teacher does, the competencies needed to carry out those tasks, and how these competencies and practices result in a positive influence on academic and behavioral gains. In addition, teachers serve in various capacities (e.g., elementary, middle, high school, general, special, ELL, reading, behavioral, coteaching), all of which bring teachers' specialized training and knowledge to bear on student learning and development. The evaluation process ideally should take into account the specific roles of teachers as well as the contexts in which they teach. For example, does the performance of an effective teacher look different at the elementary and high school levels? Do special education teachers or those who teach ELLs require a special or additional set of skills to be considered effective?

The Elementary and Secondary Education Act (ESEA) as reauthorized by the No Child Left Behind (NCLB) Act addressed the long-standing problem of achievement gaps between groups of students through the provision of public accountability at the school, district, and state levels. Reporting under the current provisions of ESEA—including subgroups such as students with disabilities and students for whom English is a second language—has brought the persistent achievement gaps for this population of students to the forefront. Clearly, there are individual differences in overall levels of competence among all students in progress toward achievement goals; however, wide achievement gaps between groups are cause for concern.

Research-based evidence confirms that teacher effectiveness is the most influential school-based factor on student achievement (Rivkin, Hanushek, & Kain, 2002; Sanders & Horn, 1998; Sanders & Rivers, 1996). Therefore, improving teacher effectiveness is at the center

of education reforms designed to address the achievement gaps between groups. Moreover, as states, districts, and schools work to identify and recruit highly qualified, effective special educators and ELL specialists, valid teacher evaluation systems that account for diverse teacher roles and responsibilities are critical to improving teaching practices and student outcomes.

Research has identified a number of challenges related to assessing and supporting effectiveness in special education and ELL teachers:

- Many special education positions are left empty or are filled with uncertified personnel (Billingsley, Fall, & Williams, 2006; Boe & Cook, 2006; McLeskey, Tyler, & Flippin, 2004).

- Of secondary special education teachers, 82 percent to 99 percent are not highly qualified in the core content areas (e.g., mathematics, science) they teach (McLeskey & Billingsley, 2008).

- Although teacher attrition was recently found to be comparable between general and special educators, school migration was higher in special education, and the supply of special education teachers is a concern (Boe, Cook, & Sunderland, 2008).

- Limited preparation of teachers in the use of evidence-based practices has been noted (Reschly, Holdheide, Smart, & Oliver, 2007; Walsh, Glaser, & Wilcox, 2006).

- ELLs are more likely than any other group of students to be taught by a teacher who lacks appropriate teaching credentials (Gandara, Maxwell-Jolly, & Rumberger, 2008).

- Limited opportunities exist for rigorous training in teacher preparation programs for teachers of English to speakers of other languages (Menken & Antunez, 2001).

- Lack of organizational support as well as low self-efficacy present difficulties for ELL specialists (Eun & Heining-Boynton, 2007).

For these reasons, and given that research confirms the relationship between teacher evaluation systems and student achievement (e.g., Kimball, White, Milanowski, & Borman, 2004; Milanowski, 2004; Odden, Borman, & Fermanich, 2004), teacher evaluation can be a powerful lever for addressing achievement gaps for students with disabilities and ELLs. Comprehensive teacher evaluation frameworks that are fair, objective, reliable, transparent, focused on instruction, and linked to professional development hold promise in improving the effectiveness of special education and ELL teachers.

CHALLENGES IN EVALUATING SPECIAL EDUCATION TEACHERS AND ELL SPECIALISTS

New models for teacher evaluation are emerging throughout the United States; however, few address the unique challenges associated with evaluating special educators and ELL specialists, particularly the challenges in accurately measuring achievement growth for their students and connecting that growth to teacher effects. Moreover, literature related to the evaluation of special educators and ELL personnel is limited. Much of the research and literature focuses on factors of special education teacher quality, such as experience, credentials, and self-efficacy (e.g., Blanton et al., 2006; Carlson, Lee, & Schroll, 2004; Goldhaber & Brewer, 2000; Greenwald, Hedges, & Lane, 1996); however, not all factors have been conclusively linked to improved teaching practices and student achievement (Goe, 2007; Harris, 2009; Kennedy, 1992).

In their study of special educators, Carlson et al. (2004) found that like general educators, the experience, attitudes, beliefs, and classroom practices of special education teachers correlated with improved student achievement; however, they noted the need for additional research to detect which measures hold promise in affecting student performance directly. Likewise, using value-added models of student achievement in Florida, Feng and Sass (2009) found that students with disabilities in general education courses "experience greater achievement gains in both math and reading when their teacher is certified in special education" (p. 1). However, causative factors documenting the significance of each of these dimensions are not fully understood, and the need to identify effective teaching and improve teaching practices remains.

Most evaluation systems lack the capacity to differentiate among teachers based on specialized roles and to consider the challenges of working with at-risk students and specific contexts (Chait, 2009; Toch & Rothman, 2008). Yet, most states require specialized training and certification, and many special education teachers and ELL specialists have distinct roles and responsibilities within the school building and classroom (Garcia & Potemski, 2009; McGraner, 2009). Several studies have demonstrated a relationship between special education teacher preparation and teaching practices (e.g., Algozinne, Morsink, & Algozinne, 1988; Nougaret, Scruggs, & Mastopieri, 2005; Sindelar, Daunic, & Rennells, 2004), which indicates that preparation in special education leads to improved teaching practices, although no studies directly link education and training of special education teachers to student outcomes (Feng & Sass, 2009).

States and districts should recognize and account for the unique contributions and roles of special education teachers and ELL specialists as they work to improve their teacher evaluation systems, particularly where evaluation results are used to make personnel and compensation decisions. The TQ Center's inquiry revealed mixed viewpoints and practices regarding the evaluation of special educators and ELL specialists—some advocating use of a common instrument and others promoting approaches that have unique components for specialized teachers.

Most states appear to leave the evaluation of ELL specialists and special educators up to the districts, which in turn may be constrained by local bargaining agreements in terms of differentiating evaluation instruments. However, many states are taking a more active role in evaluation given the Race to the Top requirements for rigorous evaluations that are comparable across schools and include student achievement. For example, Illinois, Tennessee, and Rhode Island recently changed state policy to include language about linking teacher evaluation to student achievement. In any case, support and guidance for states in developing appropriate evaluation systems is essential, especially considering the anticipation of increased state-level involvement. States and districts would benefit from guidance and policy recommendations on the following questions:

- When crediting teachers for student learning (employing a value-added evaluation system or other growth measure), how should the individual contributions of teachers acting in a coteaching or consultant role be determined?

- How can the contributions to student achievement be accurately measured for teachers instructing special populations for which alternative standards and/or assessments are used?

- Are the key features of teacher effectiveness different for specialized personnel, such as special education teachers and ELL specialists, and should those unique features lead to additional or different content on observation protocols, student growth assessments, or alternative instruments?

- When rating special education teachers and ELL specialists using an observation protocol or alternative instrument, what special training, if any, do evaluators need?

Because of the preliminary nature of this work, these questions cannot be adequately answered at this time. Nevertheless, the brief offers recommendations for addressing the challenges associated with evaluating specialty area educators to provide guidance for states and districts charged with improving teacher evaluation and suggests research to thoroughly explore uncertainties.

AN INQUIRY INTO EVALUATION PRACTICES FOR SPECIAL EDUCATORS AND ELL SPECIALISTS

Several components of the TQ Center review were conducted in the fall and winter of 2009–10: a survey inquiry, a series of interviews with state and district administrators and researchers from across the country, and an examination of relevant literature and current practices.

To collect information on state and district evaluation practices for special education teachers, the TQ Center staff, with support from CEC and several national specialists, developed a survey designed to determine the current status of state and district policy and practice and to identify promising evaluation practices and instruments. Two separate surveys were developed with slight modifications according to state and district contexts. The TQ Center produced the Web-based surveys using Microsoft.Net (an online survey program). Several state and district directors of special education pilot-tested the surveys to gauge the validity, ease of use, utility, and readability of the survey questions prior to dissemination.

TQ Center staff also contacted state and district special education administrators and researchers to conduct a more in-depth inquiry regarding evaluation practices. State and district personnel were identified through recommendations from national experts and analysis of survey results. State and district policy and practice for ELL specialists also were examined. Because a comparable electronic mailing list for ELL specialists was not available, however, this analysis consisted primarily of an examination of relevant literature and current practice.

Population and Sample

The state survey was e-mailed to all 51 state directors of special education, and the district survey was sent to more than 3,000 special education administrators using the membership list of the CEC Council of Administrators of Special Education. A total of 36 state special education directors participated in the state survey, and 1,107 respondents from the CEC membership list participated in the state and district surveys, resulting in 1,143 total respondents.

Data are reported in table and graph format based on responses that were available in the state and local data set. However, some respondents did not complete the entire survey. In addition, some survey questions were not posed when respondents indicated a lack of knowledge about the subject at hand in previous responses. As a result, some analyses have missing data, and the data presented in the tables may not equal the total number of respondents. In these cases, results are reported based on those respondents who completed the related survey items. In addition, the findings presented in this brief cannot be generalized to the entire population, as the study of current practice was not designed to collect a representative sample of practitioners.

The inquiry into the evaluation of special educators purposefully explored the knowledge base and opinions of special education administrators as opposed to personnel charged with state and district evaluation, compensation, and professional development. This deliberate design allowed for an analysis of current practice and provided an indication of the level of involvement of special education administrators in the evaluation framework design, process, and assessment. More importantly, this model gathered the collective opinions of special education administrators with considerable knowledge and experience in supervising, mentoring, and leading this significant population of teachers.

At times, however, this approach caused respondents to have difficulty responding to survey items. Two measures were taken to counter this challenge. First, an initial question was asked to determine whether the respondent was familiar with the teacher evaluation practices within the district or, for the state respondent, whether the state played a role in teacher evaluation. If the respondent was either not familiar or the state did not play a role, the specific questions regarding state/district practice were skipped, and only questions pertaining to their thoughts and opinions were posed. In addition, each question had a "don't know" option as a default in the event that the respondent was not familiar with current practice. These scenarios were both expected and noteworthy. State personnel appeared to be less familiar with current practice often resulting in the forwarding of the survey request to other units in the state department of education where it was thought that better information existed. In some cases, the other units did not respond, leading to either limited or no useful survey information for that state. Many state respondents had incomplete information because teacher evaluation is considered a local responsibility.

Survey and Interview Results

This subsection provides an overview of the wide disparity in approaches to teacher evaluation at the state and district levels with a particular emphasis on the evaluation of special education teachers and ELL specialists.

State and District Role in Evaluation

Historically, teacher evaluation has been a local responsibility. Yet, recent accountability mandates in student performance and teacher quality have elevated the demand for state-level involvement. As shown in Table 1, the majority of state-level respondents (46.2 percent) indicated state involvement in teacher evaluation, with an additional 28.8 percent

of the respondents predicting state involvement because of the Race to the Top requirements. A smaller number of respondents indicated that their state played no role in evaluation or that they were uncertain about the role the state played (11.5 percent and 13.5 percent, respectively).

Table 1. State Role in Teacher Evaluation

Does your state play a role in teacher evaluation?

	n	%
Yes	24	46.2
No, not yet, but Race to the Top will impact our involvement.	15	28.8
No	6	11.5
Don't know	7	13.5

Note: Both state special education director respondents and respondents from the CEC listserve who indicated state-level work are included in the state responses.

As shown in Table 2, a slim majority of state respondents indicated that the state mandates a state-developed evaluation system (28.6 percent); other respondents indicated that the state provides guidance through the provision of criteria and offers a state-developed system while allowing the use of a locally developed system (23.8 percent and 14.3 percent, respectively).

Table 2. State's Current Level of Involvement in Teacher Evaluation

Which statement most accurately reflects your state's current level of involvement in teacher evaluation used for all teachers?

	n	%
Our state mandates the use of a state-developed evaluation system.	6	28.6
Our state offers a state-mandated evaluation system but also allows the use of locally developed instruments.	3	14.3
Our state provides guidance through the provision of criteria districts should use in their evaluation system.	5	23.8
None of these efforts describe our state efforts. If so, please describe:	4	19.0
Don't know	3	14.3

Note: Both state special education director respondents and respondents from the CEC listserve who indicated state-level work are included in the state responses.

Table 3 indicates that the bulk of district respondents reported using a locally developed evaluation instrument (54.7 percent), 22.1 percent use the state-developed system, and 14.2 percent use a modified version of the state's system.

Table 3. District's Current Evaluation System

Which statement most accurately reflects your district's current evaluation system?

	n	%
Our district uses the state's recommended evaluation system.	200	22.1
Our district uses a slightly modified version of the state's evaluation system.	129	14.2
Our district has developed our own teacher evaluation system.	496	54.7
None of these efforts describe our state efforts. If so, please describe:	46	5.1
Don't know	35	3.9

Note: Responses were included for only those local special education administrators who indicated familiarity with the district teacher evaluation system.

Alignment to Standards

The development of a clear and coherent definition of effective teaching is a complex task. Evaluation systems demonstrating transparent alignment to research-based practices offer clear standards of practice to guide evaluation (Dwyer, 1998). Credible evaluation systems assess teachers' performance on tasks that can yield improvements in student learning.

According to the survey results, the majority of state and district evaluation systems align to the states' professional teaching standards (62.2 percent and 50.4 percent, respectively), with a smaller number of respondents showing alignment to the Interstate New Teacher Assessment and Support Consortium (INTASC) standards and Charlotte Danielson's (1996) *Enhancing Professional Practice: A Framework for Teaching* (3.8 percent and 18.3 percent, respectively) (National Comprehensive Center for Teacher Quality, 2010). Notably, 29 of the

102 respondents who selected "other" indicated that the evaluation system aligned to no particular standards (28.4 percent) (National Comprehensive Center for Teacher Quality, 2010). In addition, some of the open-ended responses provided insight to the general perceptions of the evaluation process. One respondent indicated, "Our evaluation tool was developed in the district over 40 years ago." Another respondent reported, "Our current evaluation system is outdated and applied to nothing!"

Preparation and professional development standards of practice for special education and ELL teachers are potential resources for teacher assessments and could serve as the basis for assessment (Blanton et al., 2006). Both CEC and INTASC offer teaching standards for general and special education teachers. CEC published the latest revisions to the preparation standards in 2009. The 10 revised content standards include knowledge and skill competencies for specialty areas (e.g., early childhood, secondary transition) (Council for Exceptional Children, 2009). Similarly, INTASC elaborates on its standards for general and special education teachers, with supplementary standards for special education teachers (Blanton et al., 2006). Both groups of standards were developed with considerable input from stakeholders (Blanton et al., 2006).

Likewise, the Center for Applied Linguistics Working Group on ELL Policy (2010) recommendations regarding ESEA reauthorization assert that insufficient attention is being paid to training and professional development for educators of ELLs, which limits educator effectiveness. The working group makes the following recommendation:

> Require states to demonstrate—as a precondition for receiving funds under Title II and Title III—that their credential

requirements and alternative routes to certification of teachers of core content include components that are effective in preparing these teachers to address both the content and academic language needs of English language learners. (p. 9)

The working group also notes that 15 states have no requirements for teacher preparation programs to incorporate components into their curriculum to prepare teachers for working with ELLs. Furthermore, the working group recommends that English as a second language be defined as a core academic subject under ESEA so that the same requirements would apply to teachers of ELLs as to teachers of other core academic subjects. It also states, "The capacity of teachers to meet the needs of English language learners must be a central component in all mechanisms for determining teacher effectiveness" (p. 10).

Given the increased likelihood that all teachers, including general education teachers, have or have had a student with a disability or student with limited English proficiency in their classroom (Ballantyne, Sanderman, & Levy, 2008; Office of Special Education Programs, 2001), the strengthening and alignment of standards that promote practices to support the education of students with disabilities and ELLs is essential. The inclusion of such standards could not only serve as a foundation for high-quality preparation but also as the basis for teacher assessment models.

It also should be noted that when queried about essential evaluation components, the majority of the survey respondents (91.7) indicated that special education teachers' use of evidence-based strategies should be a component of the evaluation process (see Figure 1). This assertion, combined with the fact that recent research has revealed inadequate teacher preparation in evidence-

based practices (Menken & Atunez, 2001; Reschly et al., 2007) advances the importance of this type of alignment within the definition and evaluation instrument(s).

Figure 1. Level of Agreement With Statements Regarding Teacher Evaluation
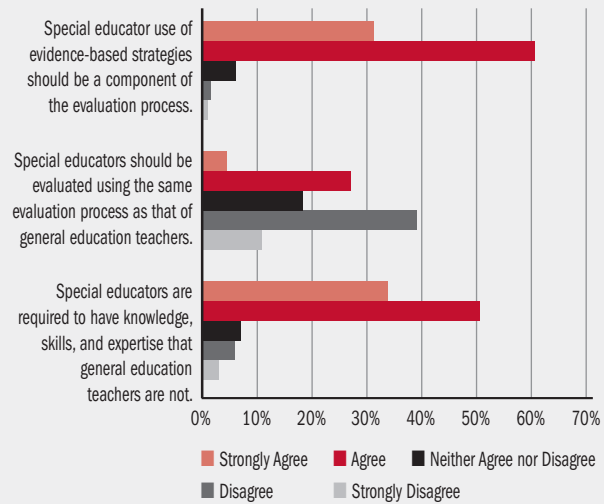
Modification of Evaluation Processes

As shown in Table 4, most state and district respondents indicated that their current evaluation system does not allow for a different or slightly modified evaluation process for special education teachers. Further, the majority of district respondents (81.4 percent) indicated that contractual agreements do not allow for modification in the evaluation process for special education teachers (National Comprehensive Center for Teacher Quality, 2010). Conversely, 49.9 percent of the respondents "disagreed" or "strongly disagreed" with the statement that special education teachers should be evaluated using the same process as that of general education teachers. In addition, 84.1 percent indicated that special education teachers are required to have knowledge, skills, and expertise that general education

teachers are not (see Figure 1). Many district-level interviewees expressed that using the same system to evaluate both general and special education teachers was not a "good fit" for special educators. It should be noted that, in systems that allow for a different or modified process, most (86.6 percent) include all special education teachers, with 6.3 percent including educators for low-incidence disabilities and 5.0 percent including educators serving as consultants/coteachers (National Comprehensive Center for Teacher Quality, 2010).

Table 4. State/District Allowance for a Different or Slightly Modified Process for Special Educators

Does your state's/district's current evaluation system allow for a different or slightly modified evaluation process for special education teachers?

	State		District		Total	
	n	*%**	*n*	*%*	*n*	*%*
Yes	3	18.8	220	26.1	223	26.0
No	11	68.8	606	72.0	617	71.9
Don't know	2	12.5	16	1.9	18	2.1

* The figures in this column do not add up to 100 percent due to rounding.

Note: Both state special education director respondents and respondents from the CEC listserve who indicated state-level work are included in the state responses. State and district responses were included for only those special education administrators who indicated familiarity or involvement with the teacher evaluation system.

MEASUREMENT INSTRUMENTS

For an evaluation system to be effective, it must be understood, credible, and valued. Much of the resistance to using evaluation results to make personnel and compensation decisions surfaces when judgments are based on a single score or data source (Blanton et al., 2006). In addition, the American Educational Research Association, ETS, and others have indicated that making high-stakes decisions based on a single measure is not sound. Research suggests that multiple sources are required to gain a full, fair, and accurate picture of a teacher's performance (Bill & Melinda Gates Foundation, 2010; Blanton et al., 2006; Little, 2009).

Table 5 provides respondent feedback about the type of measurement instruments used for all teachers. Observation protocols were cited most frequently as the data source at both the state and district levels (68.8 percent and 93.8 percent, respectively). Goal-driven professional development (i.e., teachers' progress on their professional goals factored into the evaluation results) was the second most cited (56.3 percent state and 62.1 percent district) followed by the evaluation of classroom artifacts (18.8 percent state and 43.9 percent district), teacher portfolios (37.5 percent state and 26.8 percent district), and self-report measures (18.8 percent state and 36.8 district).

Table 5. Measurement Instruments

Which measures are used/recommended within your state's/district's current evaluation system for ALL teachers?

	State		District	
	n	%	*n*	%
Standardized achievement test (e.g., value added)	2	12.5	125	4.9
Criterion-referenced or curriculum-based measures	3	18.8	170	20.2
Observation protocols	11	68.8	790	93.8
Classroom artifacts	3	18.8	370	43.9
Teacher portfolio	6	37.5	226	26.8
Self-report measures (teacher survey/checklists)	3	18.8	310	36.8
Student teacher evaluations	3	18.8	56	6.7
Parent/family teacher evaluations	1	6.3	46	5.5
Goal-driven professional development	9	56.3	523	62.1
Other:	3	18.8	54	6.4

Relatively few state and district respondents cited growth measures using standardized achievement tests or curriculum-based measures as a data source in teacher evaluation (12.5 percent and 18.8 percent state and 14.9 percent and 20.2 percent district, respectively). It should be noted, however, that the survey inquired about current practice, and many responses reflected student achievement as a main component in future teacher evaluation practice. The District of Columbia Public School's IMPACT System is one example of an evaluation system that uses more than one measure to determine teacher effectiveness (see "Practical Example: District of Columbia Public Schools IMPACT System").

PRACTICAL EXAMPLE: DISTRICT OF COLUMBIA PUBLIC SCHOOLS IMPACT SYSTEM

The District of Columbia Public Schools (DCPS) has recently developed IMPACT, a new system for assessing the performance of teachers and other school-based staff. Designed with considerable feedback from all personnel, IMPACT uses the following core measurement components as part of the evaluation of teachers:

- *Individual Teacher Value-Added Scores*
  Statistical measure of how much the teacher has affected students' learning during the year based on standardized test scores (DC-CAS)

- *Non-Value-Added Achievement*
  Measure of student growth demonstrated through standards-based measures (e.g., DRA, DIBELS)

- *Teaching and Learning Framework*
  Rubric used during observation of teacher practice

- *Commitment to the Community*
  Rubric to evaluate efforts related to school initiatives, support for special education and ELL teachers, and collaboration with colleagues

- *School Value-Added Scores*
  Measure of how much the school has affected students' learning during the year based on standardized test scores (DC-CAS)

- *Core Professionalism*
  Rubric to assess personnel professionalism

Dependent on discipline and available data, each measurement component is weighted differently with slight distinctions for specific groups. For example, all special education teachers are rated as follows:

- 50%—Teaching & Learning Framework
- 10%—Non-Value-Added Achievement
- 5%—School Value-Added Scores
- 5%—Commitment to the School
- 15%—IEP Quality
- 15%—IEP Timelines
- Core Professionalism

To access a listing of all 20 groups and their corresponding guidebooks, visit **http://dcps.dc.gov/DCPS/In+the+Classroom/Ensuring+Teacher+Success/IMPACT+(Performance+Assessment)**.

Value-Added Models

Recent advances in technology have created the capability to use student standardized test scores and curriculum-based/criterion-referenced measurements to compare students' actual growth with predicted growth and assign scores to teachers (or a team of teachers) based on prior student performance, representing their contribution to student learning growth. Value-added models are statistical techniques that use multiple years of student standardized test data to estimate the effects of schools and/or teachers (McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004). Triggered by Race to the Top requirements, value-added models have attracted the attention of policymakers, researchers, and the community at large. Virtually all research on value-added models, however, has focused on general education students, excluding special populations.

Still relatively new as evaluation measures, value-added models can be powerful tools to identify effective teachers; however, there are many challenges in implementing value-added assessments and several notable obstacles in using value-added measures for special education and ELL teachers.

General Challenges in Using Value-Added Models

- It is uncertain how much of a value-added score represents classroom effects versus teacher effects—that is, how much of the value-added score truly represents the teachers' contributions and how much reflects differences in classrooms, such as the effects of curriculum, resources, school culture, and peer effects (Braun, Chudowsky, & Koenig, 2010).

- Value-added scores can be generated only for teachers whose students have prior standardized achievement test scores, which excludes about 70 percent of teachers (Prince et al., 2006).

- In schools with highly mobile populations, value-added models that adjust for missing data must be used, and some models do this better than others (Wright & Sanders, 2008).

Challenges Specific to Special Education and ELL Classrooms

- Because the science of value-added modeling has not included and specifically addressed special educators and ELL specialists, a research-derived model for these teachers does not exist. Thus far, the research literature on challenges in using standardized test scores of students with disabilities and ELLs in the same manner as those of general education students for the purposes of measuring growth is limited but suggests that standardized tests are tenuous measures of ELL achievement unless ELL characteristics are considered throughout all stages of test development (Solórzano, 2008).

- Learning trajectories for ELLs may be different than those of native English speakers as well as for students with disabilities and students without disability labels. Adjustments may need to be made for what constitutes achievement for ELLs or students with disabilities.

- The mobility of some immigrant populations and students with disabilities may make it difficult to link student learning to individual teachers in order to determine teachers' contributions to student learning growth.

- Small student samples commonly associated with special education and ELL caseloads produce results that are statistically less reliable (Amrein-Beardsley, 2008; Blanton et al., 2006; Feng & Sass, 2009; Koretz, 2008).

- Value-added literature shows researchers' concerns over whether "all gains are created equal," stressing that little is known as to whether a 10-point gain at different points on an achievement scale for a higher and lower achiever is comparable (e.g., Braun, 2005; Rothman, 2009).

Challenges Specific to ELL Teachers

- In order to make accurate predictions about what a student's achievement level should be in a given year, test scores from previous years must be included in the model. For

many ELLs, such test scores do not exist, particularly if they have only recently immigrated to the United States.

- When students who are learning English take a test, they are essentially tested on two factors: English comprehension and knowledge of the subject being tested. Standardized tests that have been validated for testing subject matter knowledge assume a certain level of English proficiency. When students are not proficient in English, it may appear that they do not know the content when in fact, they may not understand the question, reading passage, or answer choices. Abedi (2003) analyzed data from four schools nationwide to examine ELLs' performance on tests with varying levels of "language demand" (how advanced or complex the English is) and concluded that language demand was confounded with students' content knowledge. Although a primary responsibility of the ELL teacher is to know the language demands across content areas (Harper & de Jong, 2004), teachers' contributions to ELLs' learning in content areas is entangled with their contributions to progress in learning English.

- ELL progress rates in acquiring English proficiency vary depending on a number of factors, including how much education they had prior to immigrating to the United States; how proficient they are in their home language; whether they have support in their homes and communities for developing further proficiency; and whether they have teachers who provide the support and instruction needed to learn English as well as content in mathematics, science, and other subjects. Researchers are still trying to understand the development of ELLs as they gain proficiency in speaking, writing, reading, and academic skills (Genesee, Lindholm-Leary, Saunders, & Christian, 2009). As Reeves (2009) states, "Language…is taught in service to the learner rather than the learner in service to the language-learning task" (p. 112).

Challenges Specific to Teachers of Students With Disabilities

- There is little research on the use of accommodations on standardized tests and their specific impact on value-added scores. Both IDEA and the current provisions of ESEA entitle students with disabilities to accommodations. As such, the interpretation and treatment of standardized test scores becomes a challenge (Koening & Bachman, 2004). For example, evaluators need to determine how the number and type of accommodations affect levels of performance and whether they should be factored into value-added scores (McLaughlin, Embler, Hernandez, & Caron, 2005).

- Students with intellectual and developmental disabilities who are assessed using alternative assessments are presently excluded from value-added models. Establishing reliable and valid methods to determine teacher effectiveness for this population of students is necessary.

Given the increasingly high stakes associated with the use of value-added models, these uncertainties need to be thoroughly explored through research. Survey results indicate that state and district systems have attempted to address some of these issues. A large portion (43.2 percent) of the survey respondents indicated that they use standardized assessment scores that exclude special education teachers, 40.8 percent include special education teachers, and another 4.8 percent have established specific inclusion criteria (National Comprehensive Center for Teacher Quality, 2010).

For example, Battelle for Kids, a privately funded value-added research center based in Ohio, works with districts where value-added reports are generated only for teachers who have 10 or more students per tested area/per grade level. Other districts have established criteria for teachers of only five or more students per tested area/per grade. Respondents who used these criteria indicated that test scores for students with disabilities are generally used

in the same manner as general education students (83.9 percent), whereas 10.7 percent of the respondents indicated that scores were used differently (National Comprehensive Center for Teacher Quality, 2010).

Other Indicators of Achievement

Some students' gains are measured by their progress on alternative standards. IDEA requires student participation in standardized testing and reporting but also allows the use of alternative assessments for students for whom the standardized assessment is inappropriate even with reasonable accommodations (IDEA, 2004). Likewise, many areas of academic and behavioral growth are not measured within standardized assessments. For example, teacher contributions to social and behavioral growth cannot be measured using value-added models. In some cases, districts have opted to institute student learning objectives that are measured using criterion-referenced or curriculum-based measurements (see "Practical Example: Austin Independent School District Reach Compensation and Retention System, Texas").

PRACTICAL EXAMPLE: AUSTIN INDEPENDENT SCHOOL DISTRICT REACH COMPENSATION AND RETENTION SYSTEM, TEXAS

This four-year pilot incentive pay program for teachers and principals was initiated in 2007–08. The program goals are as follows:
- Ensure quality teachers in every class.
- Provide professional growth opportunities.
- Increase retention.

Student growth, professional growth, and highest needs schools are program elements. Student growth is measured by student learning objectives. Each teacher develops two student learning objectives—one that targets classroom performance and the other focused on particular skills or a subgroup of students (e.g., students with special needs). Each student learning objective should be measurable and approved by the principal. Teachers and principals undergo a series of trainings in how to establish and measure learning objectives.*

For more information regarding this system, visit **http://www.austinisd.org/inside/initiatives/ compensation/docs/SCI_SLO_Campus_Quick_ Reference_Chart_2009-10.pdf.**

*Student learning objectives are used to determine incentives and are not an integral part of the evaluation of teachers at this time.

Student progress toward these objectives is sometimes included as an evaluation measure to determine teacher effectiveness.

Given the increasing use of growth models as measures of teacher effectiveness, the Q Center inquiry explored the perceptions of special education administrators. As illustrated in Figure 2, special education administrators at both the state and district levels were largely in favor of using student achievement data as a source to determine special educator effectiveness, with 60.4 percent indicating that they "agreed" or "strongly agreed." Respondents were less likely to indicate support in using standardized tests as a component of teacher evaluation (20.8 percent); however, they were in support of using student progress on their IEP goals (73.3).

Figure 2. Level of Agreement With Statements Regarding the Use of Student Achievement

Legend: Strongly Agree, Agree, Neither Agree nor Disagree, Disagree, Strongly Disagree
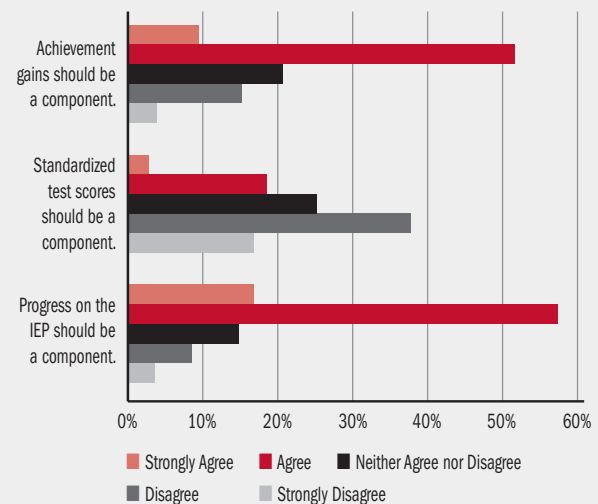
Measuring the progress of students with disabilities on their IEP goals appears to be one potential approach to assessing academic and social/behavioral growth that would address the issue of students working toward alternative standards and the challenge of determining individualized (and perhaps realistic) growth patterns. However, this approach would be highly dependent on the competencies of teachers in writing *high-quality* IEPs and using reliable and valid measurements and data to evaluate effectiveness. This approach becomes much

more contentious if used to make high-stakes decisions (e.g., personnel and compensation decisions). In at least one teacher evaluation system—Washington, D.C.'s IMPACT system— a significant portion of teachers' evaluation scores (15 percent) is based on producing quality IEPs. However, evaluating teachers on the quality of their IEPs requires that teachers be given appropriate guidance. States and districts that make a concerted effort to train and provide sustained support for teachers who write and facilitate IEPs would improve the likelihood of success (see "Practical Example: The Role of Student Progress on IEP Goals in Norwell Public Schools, Massachusetts").

PRACTICAL EXAMPLE: THE ROLE OF STUDENT PROGRESS ON IEP GOALS IN TEACHER EVALUATION IN NORWELL PUBLIC SCHOOLS, MASSACHUSETTS

In Norwell, Massachusetts, student progress on the IEP is factored into the evaluation of special education teachers. This is a full-inclusion school district in which all students with disabilities, by Grade 10, meet adequate yearly progress. The special education administrator dedicates a full year of training and support to special educators to develop their goal-writing skills—ensuring they are measurable and aligned to the state standards. Teachers also receive training and job-embedded supports to address the effective collection and use of data to make instructional decisions and evaluate student performance. Each student has a data binder that accounts for student progress.

Classroom Observations

As indicated, the majority of states and districts utilize an observation instrument as their primary measure of teacher effectiveness. Literature has noted that the vast majority of observations are informal and lack a standardized process, often resulting in inconsistent feedback and support (Blanton et al., 2006; Stuhlman, Hamre, Downer, & Pianta, n.d.). In addition, some instruments lack alignment to practices backed by empirical evidence (Stuhlman et al., n.d.). For instance, more than one fourth of the state and district respondents indicated that they "didn't know" the standards to which the evaluation system is aligned, with several more indicating that

the evaluation system was not aligned to any particular standards (National Comprehensive Center for Teacher Quality, 2010). Most respondents indicated the use of the state-developed protocol (27.1 percent). Multiple published and unpublished classroom observation systems are used at the district level. For example, Charlotte Danielson's *Enhancing Professional Practice: A Framework for Teaching*/PRAXIS III (or an adapted version) is a well-documented and researched instrument that is receiving increased attention as an evaluation tool. Of the survey respondents, 7.8 percent indicated alignment to Charlotte Danielson's (1996) *Framework* or PRAXIS III, and 11.8 percent indicated use of an adapted version of Charlotte Danielson's *Framework* (National Comprehensive Center for Teacher Quality, 2010).

It should be noted that no adaptations for special educators are made within PRAXIS III or Charlotte Danielson's *Framework* (Blanton et al., 2006). Sindelar et al. (2004) used the PRAXIS III with a sample of special educators and determined that it effectively measured the competencies of special educators "despite the fact that PRAXIS III was designed to access general education teachers" (p. 222). Survey and anecdotal information are both in support and wary of this approach, but some interviewees acknowledged that special educators have multiple responsibilities outside of teaching— some of which have legal ramifications if compliance is not met.

Only 12.2 percent of the respondents indicated that different observation protocols are used for special education teachers, some of which are designed to address the distinct roles of special educators (e.g., low-incidence teachers, consultants). Most (55.7 percent) simply modify the observation protocol to reflect the specialized roles and responsibilities; others (38.1 percent) use a separate instrument (see Table 6). For example, the Alabama Department of Education has modified the observation instrument for teachers of low-incidence disabilities to reflect instruction toward

alternative standards (see "Practical Example: Alabama Department of Education's Professional Education Personnel Evaluation Program").

Table 6. Different Observation Tools for Special Educators

Is your observation protocol different for special educators?

	n	%
Yes	97	12.2
Modified	54	55.7
A separate instrument	37	38.1
Don't know	6	6.2
No	680	85.6
Don't know	680	85.6

PRACTICAL EXAMPLE: ALABAMA DEPARTMENT OF EDUCATION'S PROFESSIONAL EDUCATION PERSONNEL EVALUATION PROGRAM

Alabama adopted a resolution in 1988 requiring the evaluation of all professional public education personnel either by a state-developed evaluation system or by one that each school system may opt to formulate pursuant to the Board-established criteria.

The Professional Education Personnel Evaluation (PEPE) Program is designed as a formative and summative evaluation system using multiple measures—one of which includes three instructional observations from trained evaluators. The tool is slightly modified for specialty area systems (e.g., special education teachers of students with cognitive disabilities, speech pathologists, library specialists). For instance, teachers of students with significant cognitive disabilities are evaluated using a slightly modified process (Special Education II). Modifications were made with direct input from the teachers. The competencies are essentially the same, with added indicators in certain areas. For example, the classroom is expanded to include community settings, and academic content is expanded to include functional life skills.

For more information about PEPE, visit **http://www. alabamapepe.com/alpepetext.htm.**

Note: The Alabama Department of Education is in the process of designing and adopting new evaluation systems. Updates will be provided on the state website.

The majority (85.6 percent) of the respondents indicated that the same observation protocol was used for all teachers (National Comprehensive Center for Teacher Quality, 2010); however, anecdotal information and survey data indicate that standard protocols are not always a good fit. Several directors indicated that they "modified" the narrative portion of the observation feedback form to differentiate the process for special educators and ELL specialists; however, this approach seemed to be unsystematic and, therefore, highly dependent on the knowledge base and skill set of the evaluator. This rather subjective method could be replaced with a skillfully and explicitly designed rubric delineating clear expectations and criteria for performance.

Observation Instruments for ELL Specialists

Teachers of ELLs are required to have certification and training to instruct students with limited English proficiency. This certification recognizes that these teachers must have specific knowledge and training to ensure that they can effectively teach students with special needs. For example, researchers have found that certain ways of communicating language to ELLs are more effective. Long (1997) found that teaching grammar was most effective as part of meaningful communication accompanied by brief interventions to point out grammatical structures that may be causing trouble for ELLs. Norris and Ortega (2000) found that explicit types of instruction were more effective than implicit types for ELLs. Neuman and Koskinen (1992) found that context was imperative in helping ELLs to acquire and use new vocabulary. Unless the observer in a classroom of ELLs understands appropriate instructional methods for teaching language to them, it is likely that the observer will not be able to distinguish between effective and ineffective teaching. Researchers have developed observational instruments to analyze the ELLs' learning experience at the student, teacher, and classroom levels (Waxman, Padrón, Franco-Fuenmayor, & Huang, 2009). In addition, some states have developed standards and require preservice teachers to complete some coursework related to teaching ELLs in order to ensure that teachers of ELLs have the knowledge and skills to work with this population of students (see "Practical Examples: State ELL Credentials and Standards").

PRACTICAL EXAMPLES: STATE ELL CREDENTIALS AND STANDARDS

California's Bilingual, Culture, Language, and Development

The Bilingual, Culture, Language, and Development (BCLAD) certificate is a rigorous credential used in California to ensure that ELL teachers have the appropriate skills and knowledge to meet the challenges of teaching diverse populations of ELLs. For example, Specially Designed Academic Instruction Delivered in English describes content instruction that is delivered in English but specially designed to provide meaningful access to the curriculum for ELLs. For more information about BCLAD and other credentials, visit **http://www.ctc.ca.gov/credentials/CREDS/english-learners.html.**

Florida's Performance Standards for Teachers of English for Speakers of Other Languages

Florida lists specific performance standards for teachers of ELLs that include attention to cultural characteristics of ELLs as well as references to specific strategies, knowledge, and skills required by these teachers. For more information about Florida's Performance Standards for Teachers of English for Speakers of Other Languages, visit **http://www.fldoe.org/aala/perstand.asp.**

Texas English as a Second Language Standards

Texas has specific standards for teachers of ELLs that go far beyond knowing how to deliver instruction. The following standards focus on teachers' abilities to make connections to the culture and communities of ELLs (Texas Education Agency, 2001):

Standard I. The ESL teacher understands fundamental language concepts and knows the structure and conventions of the English language.

Standard II. The ESL teacher has knowledge of the foundations of ESL education and factors that contribute to an effective multicultural and multilingual learning environment.

Standard III. The ESL teacher understands the processes of first- and second-language acquisition and uses this knowledge to promote students' language development in English.

Standard IV. The ESL teacher understands ESL teaching methods and uses this knowledge to plan and implement effective, developmentally appropriate ESL instruction.

Standard V. The ESL teacher has knowledge of the factors that affect ESL students' learning of academic content, language, and culture.

Standard VI. The ESL teacher understands formal and informal assessment procedures and instruments (language proficiency and academic achievement) used in ESL programs and uses assessment results to plan and adapt instruction.

Standard VII. The ESL teacher knows how to serve as an advocate for ESL students and facilitate family and community involvement in their education.

For more information about the Texas English as a Second Language Standards, visit **http://www.sbec.state.tx.us/SBECOnline/standtest/standards/allesl.pdf.**

These three states have large populations of ELLs, and they require that teachers of ELLs have specialized knowledge beyond the subjects they teach. Thus, accurately assessing these teachers may require that the evaluator have some knowledge of the appropriate practices and strategies used by teachers of ELLs or that the directions for the evaluation instruments themselves provide some guidance about the effective instructional practices.

Expanding the scope of observations to include the special knowledge and training ELL specialists possess may improve the accuracy of these protocols in evaluating instruction. Additional sources of evidence for teacher effectiveness include student reports, survey and interview data, and extensive field notes, which can complement observation data. Some experts believe that teaching ELLs requires more than just good instructional practice; finding ways to teach English without slighting students' home language and culture is essential. According to Villegas and Lucas (2007), teachers need to do the following in order to effectively teach students from diverse cultures and linguistic backgrounds:

Teachers need to be familiar with constructivist views of learning, gather information about their students' lives, develop sociocultural consciousness, hold affirming views toward diversity, use

instructional strategies that help students build connections between their lives in and out of school, and advocate for all students. (p. 28)

Although many teacher evaluation instruments explicitly address teachers' attention to meeting the needs of "diverse" learners, they may not attend to the special skills and strategies that are required to balance the acquisition of English with the affirmation of students' home culture and language. Teacher evaluation protocols should include and affirm teachers' abilities to navigate this territory. The Sheltered Instruction Observation Protocol, which was developed to assist teachers of ELLs in designing lessons to meet the needs of ELLs, may be a potential starting point for observation protocols for teachers of ELLs (see "Practical Example: Observing ELL Instruction With the Sheltered Observation Protocol").

Evaluators of Special Educators and ELL Specialists

As progressively higher stakes are attached to evaluation results, teachers are beginning to question the validity and reliability of instruments used by evaluators who, in their opinion, lack the qualifications and/or experience to make legitimate judgments about their effectiveness (Tollefson, Lee, & Webber, 2001). Of the respondents who use observation instruments, 60.2 percent indicated that training was mandated for evaluators, with only 12.4 percent of the respondents citing explicit training designed for evaluators of special education teachers. Yet, 77.4 percent of the respondents believe that evaluators of special education teachers should receive specialized training, and 61.0 percent agreed that evaluators of special educators should have experience in special education (see Figure 3).

Historically, principals have had primary responsibility for teacher evaluation (Halverson, Kelley, & Kimball, 2003). Literature has cited several problems with this approach, most notably time and expertise (Blanton et al., 2006). To combat this, teacher-to-teacher observations, such as those used in Toledo's Peer Assistance and Review (PAR) model in Ohio in which school-based teams evaluate and

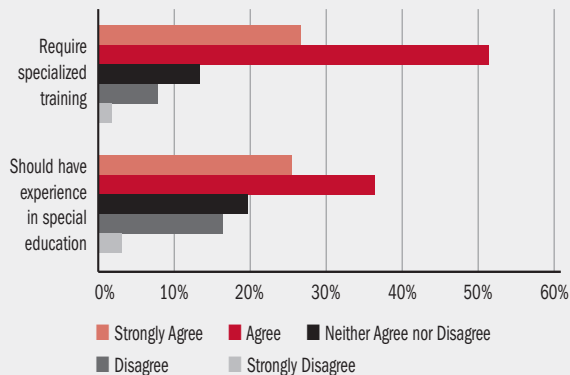### PRACTICAL EXAMPLE: OBSERVING ELL INSTRUCTION WITH THE SHELTERED INSTRUCTION OBSERVATION PROTOCOL

The Sheltered Instruction Observation Protocol (SIOP) was designed by researchers to support teachers of ELLs (Echevarria, Vogt, & Short, 2008). The SIOP contains a dual focus on content and language objectives in order to ensure that both state standards for content and language development are incorporated into the lessons. Teachers design lessons using instructional strategies for each of the SIOP components, including lesson preparation, building background, practice, and application. An alternative model is the Cognitive Academic Language Learning Approach (CALLA) Handbook, which includes subject-specific instructional strategies to promote ELL growth in academic language proficiency (Chamot, 2009).

The SIOP can be used for teacher evaluation purposes, though the creators recommend that it not be used for that purpose until teachers have mastered the components and related instructional strategies. Like Charlotte Danielson's (1996) *Framework for Teaching*, its main purpose is to create opportunities for teacher reflection and discussion with colleagues. The SIOP has been used in districts in many states, including Alaska, Arizona, Florida, Idaho, Minnesota, North Carolina, Ohio, and Tennessee.

Validated strategies for effective ELL instruction can be a useful starting point for observation protocols. However, more research is needed to empirically support the effectiveness of these strategies through studies that include comparison groups (What Works Clearinghouse, 2009). In general, there is a lack of research that identifies the most effective practices for literacy instruction of ELLs, particularly sheltered instruction and cultural awareness (August & Shanahan, 2006). Current research is in progress to establish the effectiveness of SIOP through comparison of teachers who use its strategies and those who do not.

For more information about SIOP, visit **http://www.siopinstitute.net/research.html** or **http://www.cal.org/siop/**.

Figure 3. Training and Requisite Competencies for Evaluators of Special Educators

Legend:
- Strongly Agree
- Agree
- Neither Agree nor Disagree
- Disagree
- Strongly Disagree

provide supports to novice and veteran teachers, are emerging (Harvard Graduate School of Education, n.d.). Norwell Public Schools in Massachusetts uses another promising practice (see "Practical Example: Classroom Observation Practices in Norwell Public Schools, Massachusetts"). In this district, both the principal and special education administrator observe the teacher separately; both analyze teaching practices in their area of expertise. Both then work collaboratively to develop a summative report. These models show promise with the pairing of identified, trained, and experienced master teachers and administrators to novice and/or underperforming teachers within their level of expertise.

PRACTICAL EXAMPLE: CLASSROOM OBSERVATION PRACTICES IN NORWELL PUBLIC SCHOOLS, MASSACHUSETTS

All teachers are evaluated using the same observation instrument. For teachers with fewer than three years of experience, two formative assessments are completed annually—one of which is conducted by the building principal and the other by the special education administrator. Evaluators focus on observing practices within their areas of expertise. For example, the principal might evaluate the teacher instructing a lesson within the general education classroom. Although the special education administrator may review lesson plans and instruction with a specific focus on differentiation and/or attend an IEP meeting in which the teacher's ability to effectively facilitate the meeting is evaluated, both the principal and special education administrator work collaboratively to develop a summative evaluation report.

Classroom Artifacts and Portfolios

Evaluating classroom artifacts and/or portfolios is another method of assessing teacher performance (Doolittle, 1994). Often, such documentation augments other measures (e.g., value-added scores and observations) as supporting evidence of teacher efforts in planning, instruction, assessment, reflection, and professional growth. Use of portfolios to make personnel decisions can be controversial due to the often subjective nature of the review (Doolittle, 1994). As such, the exclusive use of portfolios and artifacts as measures to make personnel and compensation decisions is not recommended (Doolittle, 1994). Of the states and districts participating in the study, 27.0 percent use the evaluation of portfolios, and 43.5 percent use classroom artifacts as part of the evaluation process. Most use portfolios to complement other measurement tools, with only one respondent indicating that the evaluation of classroom artifacts was used exclusively. Approximately 16 percent of the districts have established different criteria for the evaluation of special education teacher portfolio/artifacts (National Comprehensive Center for Teacher Quality, 2010).

Although many evaluation frameworks require the collection and analysis of lesson plans, classroom artifacts, and portfolios, the process to gather and analyze these items is often not systematic and, not surprisingly, teachers often view this process as overwhelming and time-consuming. In addition, when evaluators lack expertise and experience in special education or ELL instruction, the evaluation of these work products can be even more troublesome.

Not many exemplary practices were gleaned from the examination of current practice; however, additional valuable information is available from the Kansas Performance Teaching Portfolio (KPTP) (http://www.ksde.org) and the National Board for Professional Teaching Standards (NBPTS) (http://www.nbpts.org). Designed for preservice teachers, the KPTP

system requires teacher candidates to provide documentation about lesson design and assessment—including specific information on two identified students requiring instructional modifications. NBPTS offers an evidence-based performance assessment requiring four portfolio entries and assesses teachers across all developmental levels. Nevertheless, for these measurements to be successfully implemented, teachers and administrators must receive training to understand what is expected in the development, collection, and analysis of classroom artifacts/portfolios. This training could be developed with a rubric design that includes specialized criteria for special educators and ELL specialists.

Self-Report Measures, Student/Parent Teacher Evaluations

Self-report measures and student/parent teacher evaluations are another method used to determine teacher effectiveness. Although results are highly subjective and not directly linked to student achievement, self-assessment and survey information provide valuable information regarding teacher effectiveness. Self-assessments provide teachers with a highly reflective process to analyze success/failure and determine professional development needs. Student/parent teacher evaluations provide a window into the perceptions of students and parents. As with the evaluation of artifacts, both measures are insufficient in determining teacher effectiveness exclusively; however, both have value. Of those respondents who indicated use of these measures, only 26.6 percent include evaluation items/questions that are specifically designed for special education teachers (National Comprehensive Center for Teacher Quality, 2010).

Goal-Driven Professional Development Measures

Systematically using and incorporating evaluation results to form and target job-embedded professional development and support is likely the most important part of the evaluation cycle. Some states and districts have recognized the importance of the relationship between professional development and evaluation and have moved beyond the historically disconnected professional development plan to one that is deeply embedded into the evaluation cycle and used as a component to determine teacher effectiveness (see "Practical Example: Instructional Performance Evaluation and Growth System, Dade County, Florida").

PRACTICAL EXAMPLE: INSTRUCTIONAL PERFORMANCE EVALUATION AND GROWTH SYSTEM, DADE COUNTY, FLORIDA

The Instructional Performance Evaluation and Growth System (IPEGS) is an evaluation system designed to promote the growth and development of personnel. Measurable goals that align with school improvement plans are established and evaluated regularly throughout the evaluation cycle. IPEGS offers Support Dialogue, a school-based process designed to facilitate discussion between the administrator and personnel and to identify personnel needs and necessary supports.

For more information about IPEGS, visit **http://ipegs. dadeschools.net.**

Approximately 62 percent of the respondents indicated the use of professional development growth plans to determine teacher effectiveness as a piece of the evaluation process (see Table 5). These plans include self-driven goals for which progress is factored into the evaluation results. Viewing professional growth as a necessary and continuous cycle of developing highly effective teachers is essential. Evaluation frameworks that possess the capability to detect the professional development needs of special educators and ELL specialists will inevitably lead toward more focused, goal-driven professional development plans and thereby improved student achievement.
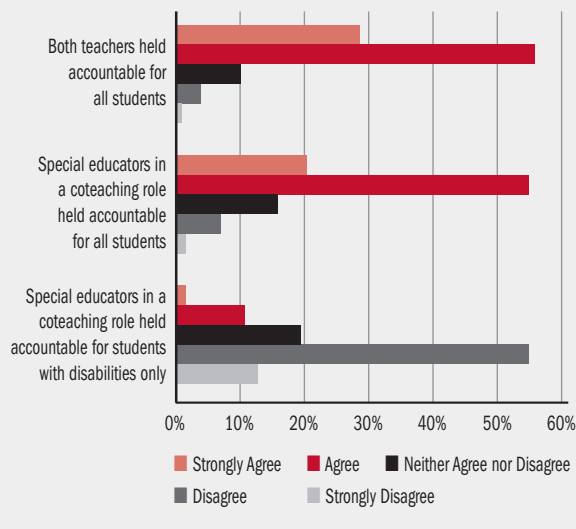
COTEACHING

Since IDEA was amended in 1997, schools have been charged with ensuring that all students have access to the general education classroom. The majority of students with disabilities (57.21 percent) are placed in the general education classroom for more than 80 percent of the school day (Office of Special Education Programs, 2007). Increasingly, various models of inclusive practices in which general and special educators share teaching responsibilities within a classroom are prevalent. The complexity of coteaching makes it difficult to determine how value-added or non-value-added scores are attributed to the general and special education teachers as "more than one teacher contributes to their academic growth" (Blanton et al., 2006, p. 124). Likewise, when observing, teacher evaluators need to determine whether they are monitoring teacher interactions with all students or only students with disabilities.

The role of ELL teachers also is not always clear in coteaching environments, making the use of observation instruments difficult. The knowledge, skills, and abilities general education teachers should hold in instructing ELLs are unclear (Janzen, 2008), further complicating the evaluation of ELL teachers. Given the importance of collaboration between ELL and general education teachers as equal professionals (DelliCarpini, 2009), more sophisticated observation protocols and additional evaluative measures are needed to capture the work that ELL teachers conduct cooperatively. Figure 4 shows that most state and district respondents are of the opinion that both teachers should be accountable for student learning for all students—regardless of disability status—with approximately 84.7 percent of the respondents selecting "agree" or "strongly agree."

One of the difficulties in coteaching is the variation in how the term is defined and how

the practice is implemented. This becomes particularly challenging when using observation instruments as one measurement to evaluate teachers. If true coteaching is absent, both teachers are not equally responsible for lesson design and implementation, complicating the evaluation of both teachers. For example, in teaching situations in which special educators serve as aides in the classroom or only assist students with disabilities, observing their teaching and interaction in a general education classroom may not be an accurate reflection of their contributions. Anecdotal information from respondents suggests that special arrangements are made to ensure that observations are scheduled at appropriate times. Others have indicated that using the same instrument during observation would be a challenge in coteaching if the shared responsibility among teachers was not evident.

Figure 4. Level of Agreement With Statements Regarding the Evaluation of Coteachers

Another complication with coteaching is determining how teacher attribution (e.g., value-added scores) is divided among students when both the general and special education teachers contribute to the academic and behavioral growth of students with disabilities—particularly when instruction is received in both a general and special education setting. As a solution to this issue, Battelle for Kids in Ohio

works with districts to accurately account for shared instruction by linking individual teachers to students and distributing student standardized test scores accordingly (see "Practical Example: Crediting Coteachers for Student Learning With Battelle for Kids BFK•Link™ Solution in Olentangy Local Schools"). Both teachers work collaboratively by evaluating each student and allocating the percentage of time each contributes to student learning in the assessed content area. Several district administrators remarked that this process is invaluable as it facilitates a deeper and necessary discussion regarding the roles each educator will assume and further promotes the implementation of inclusive services.

With students for whom value-added scores are not available, districts have opted to develop student learning objectives, which are assessed and measured using reliable and valid progress monitoring measures. Student learning objectives are determined for individual students and groups of students depicting the anticipated growth according to baseline, postmeasurements, and expected rate of growth. Some districts require general educators to develop student learning objectives for all students—regardless of disability status. Special educators are held accountable only for the progress of students with disabilities (see "Practical Example: Crediting Coteachers for Student Learning at Edison Elementary School, Denver, Colorado").

PRACTICAL EXAMPLE: CREDITING COTEACHERS FOR STUDENT LEARNING WITH BATTELLE FOR KIDS BFK·LINK™ SOLUTION IN OLENTANGY LOCAL SCHOOLS

In Olentangy Local Schools, the district's value-added scores for teachers serving in a coteaching capacity are determined by teachers (both general and special education) working collaboratively to identify their level of attribution for each student. Olentangy Local Schools completes this process as part of its involvement in Battelle for Kids' Teachers Connecting Achievement & Progress (T-CAP) initiative. Launched in 2006 and funded by The Joyce Foundation, nearly 40 Ohio school districts participated in the three-year initiative to learn the best ways to assure accurate and timely reporting of student growth as measured through value-added analysis; provide professional development and tools to help teachers use this information; develop and implement the support systems necessary to meet initiative objectives; and use the information collected to pilot research strategies addressing key policy questions pertaining to large-scale implementation.

Districts such as Olentangy and others across the country that are taking value-added analysis to the classroom level are using Battelle for Kids' innovative, Web-based BFK•Link™ solution to accurately "link" teachers to students. During the linkage process, teachers review and correct data used for teacher-level measures of effectiveness, including value-added analysis, by ensuring that all students taught are claimed for all subjects; accounting for student mobility; accounting for shared instruction/coteaching; and making the students considered in a teacher's analysis transparent. For example, for teachers working in a true coteaching situation, both teachers may each allot 50 percent. Or, if students are receiving some support services in a resource room, the general educator may allot 70 percent with the special education teacher at 30 percent. Student standardized test scores are then distributed accordingly. If teachers have allocated more than 100 percent to students, then they are identified through the system, and the teachers are requested to reevaluate. In the event this does not occur, the BFK•Link™ solution calculates it proportionally.

To use value-added analysis to inform instruction and for high-stakes decisions, accurately linking teachers to students is essential. For more information, visit **http://www.BattelleforKids.org/go/publications** and download Battelle for Kids' The Importance of Accurately Linking Instruction to Students to Determine Teacher Effectiveness white paper, commissioned by the Bill & Melinda Gates Foundation, which addresses these issues in more detail.

PRACTICAL EXAMPLE: CREDITING COTEACHERS FOR STUDENT LEARNING AT EDISON ELEMENTARY SCHOOL, DENVER, COLORADO

At Edison Elementary School in Denver, Colorado, teachers establish student growth objectives that hold teachers accountable for the individual growth of each student according to baseline, postmeasurements, and expected rate of growth. General education teachers submit two student growth objectives and are held accountable for all students in the classroom—including students with disabilities. Special education teachers submit two student growth objectives and are held accountable for measuring the individual growth for students on their IEP caseload only.

DISCUSSION OF ISSUES AND STRATEGIES

Relatively few respondents at the state and district levels (36.6 percent) "agreed" or "strongly agreed" that the present evaluation system adequately measures teacher effectiveness, with fewer indicating efficacy in the process for special educators (32.2 percent) (National Comprehensive Center for Teacher Quality, 2010). Most respondents indicated that the same evaluation process was used for all teachers; however, survey and interview data indicate that some experts in the field believe the teacher evaluation process should be differentiated to account for the additional and differentiated roles and responsibilities of special education and ELL teachers. The scarcity of evaluation programs that allow for differentiation for special educators and ELL specialists indicates that much work remains to be done.

This review revealed that although 26 percent of the respondents indicated allowance for modification, in actuality, few evaluation programs reflect modifications for special education or ELL teachers that are supported with empirical evidence and strategically and comprehensively designed. For instance, as TQ Center staff attempted to explore these nuances via follow-up phone calls to respondents (34 phone calls to states and districts), few revealed a systematic and coordinated development process. Conversely, anecdotal information reflected a much more unsystematic approach resulting from a common dissatisfaction in the current evaluation process. In fact, several administrators indicated that contractual agreements prevented specialization, so they adapted the process through the narrative portion of the evaluation to the best of their ability. Although this may be a viable option, it would likely be highly dependent on the knowledge base and skill set of the evaluator.

Although many teacher evaluation instruments explicitly address teachers' attention to meeting the needs of "diverse" learners, they may not attend to the special skills and evidence-based instructional methods for students with disabilities and ELLs. There is a fairly extensive amount of empirical evidence that identifies instructional practices that are linked to improved academic achievement for students with disabilities and ELLs. In addition, the following special skills and responsibilities for special education teachers are significant: increased collaboration in a coteaching setting, increased family engagement, IEP meeting coordination and facilitation, social and behavioral interventions, paperwork requirements, compliance with legal mandates, and the delivery of specialized interventions for students with severe cognitive disabilities or other complex impairments (e.g., autism, hearing impaired, visually impaired).

Likewise, the special skills and responsibilities of ELL specialists for increased collaboration in a coteaching setting, increased attention to home language and cultures, and the need to build connections between the students' school and home lives are central. Given these documented requirements and research-based interventions as well as persistent achievement gaps for this population of students, it seems imperative that these competencies be integrated into the teacher evaluation design. Unless evaluators have a clear understanding of these special skills and instructional methods, their capacity to distinguish between effective and ineffective special educators and ELL specialists is limited.

Of course, general education teachers also take on additional roles outside of teaching. Many teach students with disabilities and ELLs on a regular basis, but the level of intensity and responsibility differs across special and general education teachers. Although one universal evaluation system for all teachers has the virtues of simplicity and implementation ease, the different teacher roles and responsibilities with ELL populations and students with disabilities necessitates an evaluation system with the capacity to differentiate.

POLICY AND PRACTICE RECOMMENDATIONS

Many states and districts are at the starting line when it comes to developing or retooling teacher evaluation systems. Now is the time to consider the unique challenges of evaluating special educators and ELL specialists. False starts may be prevented if states, districts, and, in some cases, teacher educators consider the following recommendations:

- **Include special education and ELL administrators and teachers when revamping/designing evaluation frameworks.** Special educators and ELL specialists need to anticipate rather than react. Practitioners need to be at the forefront of these discussions to ensure that the unique skills and knowledge of special educators and ELL specialists are considered and addressed. As plans for teacher evaluation systems progress, states and districts need to account for the varying roles and responsibilities these teachers assume and incorporate these distinctions within the evaluation framework.

- **Identify a common framework that defines effective teaching for all teachers. Where appropriate, include differentiated criteria/ expectations for special education teachers and ELL specialists.** Involve all teachers in the development of the framework to foster better understanding and secure teacher empowerment and ownership. Include special educators and ELL specialists to account for the varying roles and responsibilities of special educators and ELLs. Make certain that there is prior research that supports the measurement of specific components in the evaluation system. Continue to collect evidence on instructional strategies that are known to improve outcomes for students with disabilities and ELLs.

- **Integrate evidence-based practices for students with disabilities and ELLs into evaluation models.** Incorporating and pairing evidence-based practices with consistent and explicit feedback and opportunities for job-embedded professional development will contribute to improved practice and student outcomes.

- **Establish a culture of collaboration, trust, and empowerment in which clear expectations of performance are explicitly stated and expected.** Whether designing a standardized or specially designed evaluation system, involving teachers in the process is essential to success. Establishing and communicating clear expectations reduces motivation to disagree with the content and results of teacher evaluation systems. In situations in which the evaluation measurement is potentially more subjective, offer rubrics with clear expectations and criteria for performance.

- **In addition to—or, in some situations, in the absence of—appropriate standardized assessment data, incorporate other concrete evidence of teachers' contributions to student learning into the teacher evaluation system** (e.g., progress toward accomplishing IEP objectives and student learning objectives across broad academic and behavioral domains). Criterion-referenced, curriculum-based measurements, IEP measurement data, and/or progress-monitoring data collected within the response to intervention process can be valuable to assessing teacher effectiveness. States can utilize Race to the Top funding to develop instructional improvement systems or technology-based tools and strategies that provide teachers with the data and support to make informed instructional decisions and can potentially be used to measure teacher effectiveness.

- **Improve data quality.** Whether utilizing standardized assessment data or other sources of information, ensuring teacher

evaluation system validity and reliability is essential, especially if used to make personnel and compensation decisions. When standardized assessment data are used, specific inclusion criteria that protect the validity of value-added scores should be established. Procedures should be adopted that correctly allocate teacher contribution where coteaching occurs. In situations in which IEP goals and/or criterion-referenced/curriculum-based measures are used, personnel training is imperative. Given the potential for subjectivity, staff need consistent training and support to establish accountability. Goals and performance objectives should be aligned to state standards, designed to reflect adequate student growth, and assessed appropriately to ensure progress.

- **Ensure that the evaluation framework can identify and provide the professional development needs of special educators and ELL specialists and detect improvements in practice resulting from sustained, job-embedded professional development.** Using data gathered through evaluation measurements to refine or modify professional development activities for all teachers is a powerful process to directly address identified needs. Honing in on the effective use of evidence-based practices for students with disabilities and ELLs will underscore the training needs of these teachers and provide validation of professional development efforts.

- **Establish evaluator training that includes explicit training on the specific teacher effectiveness measures used with special educators and ELL specialists and/or consider establishing a model in which peer-to-peer observations or evaluators are matched to a specific discipline.** The credibility of the evaluation measurement—whether it is an observation instrument or the evaluation of classroom artifacts/portfolios—gains integrity when the personnel being evaluated deems the evaluator as credible. Matching special educators and ELL specialists to credible evaluators will result in gains in validity.

- **Offer a checklist or rubric that offers selection criteria for evaluation models and includes specific standards for special educators and ELL specialists.** Incorporating multiple measures into the evaluation framework is a daunting task. Providing states and districts with tools to support the selection and development will advance efforts to differentiate for this significant population of teachers.

- **Support research in determining means to construct and validate value-added scores for teachers working with students on alternative standards.** Prominent researchers in the field indicated that this area is in need of further research. Allocating funds to support such efforts is imperative.

- **Consider modifying existing statute and/or policy to allow for modifications in the evaluation of special educators and ELL specialists.** Providing allowances for modifications within the teacher evaluation process for a group of students can become quite contentious. Garnering and sustaining stakeholder buy-in and support requires active and continual stakeholder involvement.

- **Collaborate with teacher preparation programs to ensure that evidence-based practices are incorporated into teacher preparation coursework and professional development activities.** Teachers and teacher educators need to be sufficiently prepared to identify and implement evidence-based practices with fidelity for academic and behavioral gains to be achieved. Identifying and incorporating these competencies into teacher preparation coursework and professional development activities will likely result in improved teacher evaluations.

CONCLUSION

The primary purpose of teacher evaluation should be to improve teaching and learning. How teacher effectiveness should be evaluated is the source of considerable discussion and debate. Under the assumption that teacher effectiveness represents, in part, a teacher's contribution to student achievement, teacher evaluation should consider evidence of student learning growth that can be reasonably attributed to the teacher. When direct evidence is difficult to evaluate or incomplete, as is often the case with these student populations, collecting evidence on specific teacher practices that are known to improve outcomes for students with disabilities and ELLs may be essential. Evaluation systems that recognize and account for the extensive training and education that ELL specialists and special educators bring to the classroom will be better able to identify practices that contribute to improved learning and allow administrators to make sound hiring and performance decisions. If evaluators lack an understanding of specific practices that contribute to improved student outcomes, then the assessment of the teacher's effectiveness may be less precise.

This brief serves as a starting point to additional discussion and research. As indicated, little is known—in terms of research and practice—as to whether evaluation systems designed for all teachers adequately measure, identify, and remediate the effectiveness of special educators and ELL specialists. This brief provides a window into the thoughts and perceptions of special education administrators and an overview of evaluation practices for this population of teachers. Research determining the validity of these methods and their ability to adequately measure teacher effectiveness and their correlation to student achievement is necessary—particularly if results are used to make personnel and compensation decisions.

◼ REFERENCES

Abedi, J. (2003). *Impact of student language background on content-based performance: Analyses of extant data (Center for the Study of Evaluation report).* Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Algozinne, B., Morsink, C. V., & Algozzine, K. M. (1988). What's happening in self-contained special education classrooms? *Exceptional Children, 55*, 259–265.

American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, 123 Stat. 115 (2009). Retrieved June 10, 2010, from http://frwebgate.access.gpo.gov/cgi-bin/getdoc. cgi?dbname=111_cong_bills&docid=f:h1enr.txt.pdf

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher, 37*(2), 65–75.

August, D., & Shanahan, T. (2006). Synthesis: Instruction and professional development. In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 351–363). Mahwah, NJ: Lawrence Erlbaum.

Ballantyne, K. G., Sanderman, A. R., & Levy, J. (2008). *Educating English language learners: Building teacher capacity roundtable report.* Washington, DC: National Clearinghouse for English Language Acquisition. Retrieved June 10, 2010, from http://www.ncela.gwu.edu/files/ uploads/3/EducatingELLsBuildingTeacherCapacityVol1.pdf

Bill & Melinda Gates Foundation. (2010). *Empowering effective teachers: Strategies for implementing reforms.* Seattle, WA: Author.

Billingsley, B., Fall, A., & Williams, T. O. (2006). Who is teaching students with emotional disorders? A profile and comparison to other special educators. *Emotional Disorders, 31*(1), 252–264.

Blanton, L. P., Sindelar, P. T., & Correa, V. I. (2006). Models and measures of beginning teacher quality. *The Journal of Special Education, 40*(2), 115–127.

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. Retrieved June 10, 2010, from http://www.gse.upenn.edu/cresp/pdfs/EC%20Shortage%20Article.pdf

Boe, E. E., Cook, L. H., & Sunderland, R. J. (2008). Teacher turnover: Examining exit attrition, teaching area transfer, and school migration *Exceptional Children, 75*(1), 7–31.

Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models.* Princeton, NJ: Policy Information Center, Educational Testing Service. Retrieved June 10, 2010, from http://www.ets.org/Media/Research/pdf/PICVAM.pdf

Braun, H., Chudowsky, N., & Koenig, J. (2010). *Getting value out of value-added: Report of a workshop.* Washington, DC: National Research Council Committee on Value-Added Methodology for Instructional Improvement, Program Evaluation, and Accountability.

Carlson, E., Lee, H., & Schroll, K. (2004). Identifying attributes of high quality special education teachers. *Teacher Education and Special Education, 27*(4), 350–359.

Center for Applied Linguistics Working Group on ELL Policy. (2010). *Recommendations for the reauthorization of the Elementary and Secondary Education Act.* Washington, DC: Center for Applied Linguistics. Retrieved June 10, 2010, from http://www.cal.org/topics/ell/ELL-Working-Group-ESEA.pdf

Chait, R. (2009). *Ensuring effective teachers for all students: Six state strategies for attracting and retaining effective teachers in high-poverty and high-minority schools.* Washington, DC: Center for American Progress.

Chamot, A. U. (2009). *The CALLA handbook: Implementing the cognitive academic language learning approach* (2nd ed.). White Plains, NY: Longman.

Council for Exceptional Children. (2009). *What every special educator should know: Ethics, standards, and guidelines* (6th ed.). Washington, DC: Author. Retrieved June 10, 2010, from http://www.cec.sped.org/Content/NavigationMenu/ProfessionalDevelopment/ProfessionalStandards/What_Every_Special_Educator_Should_Know_6th_Ed_revised_2009.pdf

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

DelliCarpini, M. (2009). Success with ELLs. *English Journal, 98*(5), 116–119.

Doolittle, P. (1994). Teacher portfolio assessment. *Practical Assessment, Research, & Evaluation, 4*(1). Retrieved June 10, 2010, from http://pareonline.net/getvn.asp?v=4&n=1

Dwyer, C. A. (1998). The evaluation of teachers. In B. Alvarez & M. Ruiz-Casares (Eds.), *Evaluation and educational reform: Policy options* (pp. 171–201). Washington, DC: U.S. Agency for International Development, Bureau for Latin America and the Caribbean, Office of Regional Sustainable Development, Education and Human Resources Division: ABEL2 Project, Academy for Educational Development.

Echevarria, J. A., Vogt, M. J., & Short, D. J. (2008). *Making content comprehensible for English learners: The SIOP model* (3rd ed.). Boston: Allyn & Bacon.

Eun, B., & Heining-Boynton, A. L. (2007). Impact of an English-as-a-second-language professional development program. *Journal of Educational Research, 101*(1), 36–48.

Feng, L., & Sass, T. R. (2009). *Special education teacher quality and student achievement.* Report prepared for the U.S. Department of Education. Teacher Quality Research.

Gandara, P., Maxwell-Jolly, J., & Rumberger, R. W. (2008). *Resource needs for English learners: Getting down to policy recommendations.* Santa Barbara, CA: University of California Linguistic Minority Research Institute.

Garcia, P., & Potemski, A. (2009). *Recruiting teachers for schools serving English language learners* (Key Issue). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved June 10, 2010, from http://www2.tqsource.org/strategies/recruit/recruitingTeachersfor SchoolsServingELLs.pdf

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2009). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students Placed at Risk, 10*(4), 363–385.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved June 10, 2010, from http://www.tqsource.org/publications/LinkBetweenTQandStudentOutcomes.pdf

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 23*, 79–86.

Greenwald, R. L., Hedges, V., & Lane, H. (1996). The effect of school resources on school achievement. *Review of Educational Research, 36*, 42–48.

Halverson, R., Kelley, C., & Kimball, S. (2003). Implementing teacher evaluation systems: How principals make sense of complex artifacts to shape local instructional practice. In W. K. Hoy & C. G. Miskel (Eds.), *Educational administration, policy, and reform: Research and measurement* (Vol. 3). Charlotte, NC: Information Age Publishing.

Harper, C., & de Jong, E. (2004). Misconceptions about teaching English language learners. *Journal of Adolescent and Adult Literacy, 48*(2), 152–162.

Harris, D. N. (2009). *Teacher value-added and credentials as tools for school improvement.* PowerPoint presentation given during the National Comprehensive Center for Teacher Quality webcast *Evaluating Teacher Effectiveness: The What, How, and Why of Educator Evaluation.* Retrieved June 10, 2010, from http://www.tqsource.org/webcasts/evaluateEffectiveness/Harris_VAM_WC.pdf

Harvard Graduate School of Education. (n.d.). *A user's guide to peer assistance and review.* Cambridge, MA: Author. Retrieved June 10, 2010, from http://www.gse.harvard.edu/~ngt/par/resources/users_guide_to_par.pdf

Individuals with Disabilities Education Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647. (2004). Retrieved June 10, 2010, from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=108_cong_public_laws&docid=f:publ446.108.pdf

Janzen, J. (2008). Teaching English language learners in the content areas. *Review of Educational Research, 78*(4), 1010–1038.

Kennedy, M. M. (1992). The problems of improving teacher quality while balancing supply and demand. In E. E. Boe & D. M. Guildford (Eds.), *Teacher supply, demand, and quality: Policy issues, models, and data bases: Proceedings of a conference* (pp. 65–108). Washington, DC: National Academy Press.

Kimball, S. M., White, B., Milanowski, A. T., & Borman, G. (2004). Examining the relationship between teacher evaluation and student results in Washoe County. *Peabody Journal of Education, 79*(4), 54–78.

Koening, J. A., & Bachman, L. F. (2004). *Keeping score for all: The effects of inclusion and accommodations on large-scale educational assessments.* Washington, DC: National Academy of Sciences.

Koretz, D. (2008, Fall). A measured approach. Value-added models are a promising improvement, but no one measure can evaluate teacher performance. *American Educator*, 18–39.

Little, O. (2009). *Teacher evaluation systems: The window for opportunity and reform.* Washington, DC: National Education Association.

Long, M. H. (1997). *Focus on form in task-based language teaching.* Presentation at the Fourth Annual McGraw-Hill Teleconference in Second Language Teaching.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics, 29*(1), 67–101. Retrieved June 10, 2010, from http://www.rand.org/pubs/reprints/2005/RAND_RP1165.pdf

McGraner, K. (2009). *Preparing teachers of English language learners* (TQ Connection Issue Paper). Washington, DC: National Comprehensive Center for Teacher Quality. Retrieved June 10, 2010, from http://www.tqsource.org/publications/issuepaper_preparingELLteachers.pdf

McLaughlin, M. J., Embler, S., Hernandez, G., & Caron, E. (2005). No Child Left Behind and students with disabilities in rural and small schools. *Rural Special Education Quarterly, 24*(1), 32–29.

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*, 293–305.

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

Menken, K., & Atunez, B. (2001). *An overview of the preparation and certification of teachers working with limited English proficient (LEP) students.* Washington, DC: National Clearinghouse for Bilingual Education.

Milanowski, A. (2004). The relationship between teacher performance evaluation scores and student achievement: Evidence from Cincinnati. *Peabody Journal of Education, 79*(4), 33–53.

National Comprehensive Center for Teacher Quality. (2010). Survey of special educators conducted by the National Comprehensive Center for Teacher Quality with support from the Council for Exceptional Children [aggregated survey results]. Washington, DC: Author.

National Council on Teacher Quality. (2008). *State teacher policy yearbook 2008, national summary.* Washington, DC: Author.

Neuman, S., & Koskinen, P. (1992). Captioned television as comprehensible input: Effects of incidental word learning from context for language minority students. *Reading Research Quarterly, 27*(1), 95–106.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002). Retrieved June 10, 2010, from http://www.ed.gov/policy/elsec/leg/esea02/index.html

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417–528.

Nougaret, A. A., Scruggs, T. E., & Mastopieri, M. M. (2005). Does teacher education produce better special education teachers? *Exceptional Children, 71*(3), 217–240.

Odden, A., Borman, G., & Fermanich, M. (2004). Assessing teacher, classroom, and social effects, including fiscal effects. *Peabody Journal of Education, 70*(4), 4–32.

Office of Special Education Programs. (2001). *General education teachers' role in special education* (SPeNSE fact sheet). Gainesville, FL: Author.

Office of Special Education Programs. (2007). T*able 2-2. Students ages 6 through 21 served under IDEA, Part B, by educational environment and state: Fall 2007* [Spreadsheet]. Washington, DC: U.S. Department of Education. Retrieved June 10, 2010, from http://www.ideadata.org/TABLES31ST/AR_2-2.xls

Prince, C. D., Schuermann, P. J., Guthrie, J. W., Witham, P. J., Milanowski, A. T., & Thorn, C. A. (2006). *The other 69 percent: Fairly rewarding the performance of teachers of non-tested subjects and grades.* Washington, DC: U.S. Department of Education, Office of Elementary and Secondary Education.

Reeves, J. R. (2009). A sociocultural perspective on ESOL teachers' linguistic knowledge for teaching. *Linguistics and Education, 20*, 109–125.

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.

Rivkin, S., Hanushek, E., & Kain, J. (2002). *Teachers, schools, and academic achievement.* Dallas: University of Texas–Dallas, Texas Schools Project.

Rothman, R. (2009). An inexact science: What are the technical challenges involved in using value-added measures? *Harvard Education Letter, 25*(2), 1–4. Retrieved June 10, 2010, from http://www.hepg.org/hel/article/164#home

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12*(3), 247–256.

Sanders, W., & Rivers, J. (1996). *Cumulative and residual effects of teachers of future student academic achievement.* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sindelar, P. T., Daunic, A., & Rennells, M. S. (2004). Comparisons of traditionally and alternatively trained teachers. *Exceptionality, 12*, 209–223.

Solórzano, R. W. (2008). High stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*(2), 260–329.

Stuhlman, M. W., Hamre, B. K., Downer, J. T., & Pianta, R. C. (n.d.). *A practitioner's guide to conducting classroom observations: What the research tells us about choosing and using observational systems.* Charlottesville: Center for Advanced Study of Teaching and Learning (CASTL), University of Virginia. Retrieved June 10, 2010, from www.wtgrantfdn.org/File%20 Library/Resources/Practitioners%20Guide.pdf

Texas Education Agency. (2001). *English as a second language standards.* Austin, TX: Author. Retrieved June 10, 2010, from http://www.sbec.state.tx.us/SBECOnline/standtest/standards/ allesl.pdf

Toch, T., & Rothman, R. (2008). *Rush to judgment: Teacher evaluation in public education.* Washington, DC: Education Sector.

Tollefson, N., Lee, S. W., & Webber, L. (2001). *The consistency of systematic classroom observations in urban schools.* Kansas City, MO: Ewing Marion Kauffman Foundation.

U.S. Department of Education. (2010). *A blueprint for reform: The reauthorization of the Elementary and Secondary Education Act.* Washington, DC: Author. Retrieved June 10, 2010, from http:// www2.ed.gov/policy/elsec/leg/blueprint/blueprint.pdf

Villegas, A. M., & Lucas, T. (2007). The culturally responsive teacher. *Educational Leadership, 64*(6), 28–33.

Walsh, K., Glaser, D., & Wilcox, D. D. (2006). *What education schools aren't teaching about reading and what elementary teachers aren't learning.* Washington, DC: National Council on Teacher Quality. Retrieved June 10, 2010, from http://www.nctq.org/nctq/images/nctq_reading_study_ app.pdf

Waxman, H., Padrón, Y. N., Franco-Fuenmayor, S. E., & Huang, S. L. (2009). Observing classroom instruction for ELLs from student, teacher, and classroom perspectives. *TABE Journal, 11*(1), 63–95.

What Works Clearinghouse. (2009). *WWC intervention report: Sheltered instruction observation protocol* (SIOP). Washington, DC: Institute of Education Sciences, U.S. Department of Education.

Wright, S. P., & Sanders, W. L. (2008). *Decomposition of estimates in a layered value-added assessment model.* Paper presented at the National Conference on Value-Added Modeling, Madison, Wisconsin. Retrieved June 10, 2010, from http://www.wcer.wisc.edu/news/events/ WrightSanders_Decomposition.pdf

ACKNOWLEDGMENTS

ABOUT THE NATIONAL COMPREHENSIVE CENTER FOR TEACHER QUALITY

The National Comprehensive Center for Teacher Quality (TQ Center) was created to serve as the national resource to which the regional comprehensive centers, states, and other education stakeholders turn for strengthening the quality of teaching—especially in high-poverty, low-performing, and hard-to-staff schools—and for finding guidance in addressing specific needs, thereby ensuring that highly qualified teachers are serving students with special needs.

The TQ Center is funded by the U.S. Department of Education and is a collaborative effort of ETS, Learning Point Associates, and Vanderbilt University. Integral to the TQ Center's charge is the provision of timely and relevant resources to build the capacity of regional comprehensive centers and states to effectively implement state policy and practice by ensuring that all teachers meet the federal teacher requirements of the current provisions of the Elementary and Secondary Education Act (ESEA), as reauthorized by the No Child Left Behind Act.

The TQ Center is part of the U.S. Department of Education's Comprehensive Centers program, which includes 16 regional comprehensive centers that provide technical assistance to states within a specified boundary and five content centers that provide expert assistance to benefit states and districts nationwide on key issues related to current provisions of ESEA.