

## Teacher formative assessment: the missing link in response to intervention

Rutgers University has made this article freely available. Please share how this access benefits you.

Your story matters. [\[https://rucore.libraries.rutgers.edu/rutgers-lib/49371/story/\]](https://rucore.libraries.rutgers.edu/rutgers-lib/49371/story/)

**Citation to Publisher** Reddy, Linda A., Dudek, Christopher M. & Shernoff, Elisa S. (2016). Teacher formative assessment: the missing link in response to intervention. In Shane R. Jimerson, Matthew K. Burns, Amanda M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The Science and Practice of Multi-Tiered Systems of Support (2nd ed.)* (607-623). New York: Springer. <http://dx.doi.org/10.1007/978-1-4899-7568-3>.

**Citation to *this* Version:** Reddy, Linda A., Dudek, Christopher M. & Shernoff, Elisa S. (2016). Teacher formative assessment: the missing link in response to intervention. In Shane R. Jimerson, Matthew K. Burns, Amanda M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The Science and Practice of Multi-Tiered Systems of Support (2nd ed.)* (607-623). New York: Springer. Retrieved from [doi:10.7282/T34B33DJ](https://doi.org/10.7282/T34B33DJ).

**Terms of Use:** Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

*Article begins on next page*

## **Teacher Formative Assessment: The Missing Link in Response to Intervention**

Linda A. Reddy, Christopher M. Dudek, and Elisa S. Shernoff  
Rutgers University

Chapter to appear in: S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of Response to Intervention: The Science and Practice of Multi Tiered Systems of Support* (2<sup>nd</sup> ed.).  
New York, NY: Springer Science.

### Author Note

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A080337 and U.S. Department of Education – Teacher Incentive Fund Program, through Grant S374A120060 to Rutgers University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Correspondence concerning this chapter should be addressed to Linda A. Reddy, Ph.D. Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ. 08854; E-mail: LReddy@scarletmail.rutgers.edu

### **Abstract**

Response to Intervention (RtI) focuses on the assessment, intervention, and progress monitoring of student academic performance and social behavior. Despite requiring highly-qualified personnel for successful implementation, the implementation of RtI has not focused on applying its foundational principles towards promoting teacher effectiveness through assessment, intervention, and progress monitoring of teacher classroom practice. Compounding this problem is the lack of availability of reliable and valid teacher assessments to apply in an RtI model for teacher professional development. This chapter provides a rationale for applying RtI principles to teacher professional development and how teacher formative assessment can improve educator effectiveness, student learning, and social behaviors. The Classroom Strategies Scale (CSS, Reddy & Dudek, 2014), a new multidimensional assessment of instructional and behavioral management practices is discussed as an example of one promising tool for promoting teachers professional development within an RtI model. We offer a synthesis of the theory, research, and evidence of reliability and validity of the CSS. The application of teacher formative assessment in job-embedded professional development/coaching models for schools is discussed. Finally, implications for practice and research are outlined.

Response to Intervention (RtI) is an education framework emphasizing multi-tiered systems for prevention, identification, and intervention for *all* students at risk for learning and behavioral challenges. RtI organizes educational resources and service delivery for general and special education into an integrated problem solving approach for promoting student achievement (Batsche et. al, 2005; [NCRI] National Center on Response to Intervention, 2010). Successful implementation of RtI requires collaboration between parents, educators, related service providers, and administrators to promote students' success (Jackson et al., 2009). RtI also provides school districts with an organizational framework for promoting student achievement, providing immediate learning supports as needed, reducing the number of special education referrals, and enhancing classroom instruction (Jackson et al., 2009).

Within an RtI framework, educators use progress monitoring tools to adapt instruction to meet diverse student learning needs. While RtI has several models, consensus exists among scholars on the core principles and overall framework (e.g., Fuchs & Fuchs, 2006; Klinger & Edwards, 2006; Mellard, McKnight, & Woods, 2009; Jackson, Pretti-Frontczak, Harjusola-Webb, Grisham-Brown, & Romani, 2005). National organizations, coalitions, and researchers have defined the key features of an RtI model to include: (a) utilizing a research-based core curriculum (b) universal screening, (c) a multi-tiered system of evidence-based interventions matched to student need, (d) data-based decision making regarding instructional practices, and (e) progress monitoring through formative and summative assessments (e.g., Batsche, Curtis, Dorman, Castillo, & Porter, 2008; Chun & Witt, 2008; Fuchs & Fuchs 2006; Mellard, McKnight, & Woods, 2009). These principles are woven into a three tiered service and resource delivery system (Fuchs & Fuchs, 2006; Klinger & Edwards, 2006).

Although RtI models have traditionally focused on identifying and supporting students at-risk for learning and behavioral difficulties, recent models emphasize supporting *all* learners in school systems (Bender & Shores, 2007; National Association of State Directors of Special Education, 2007; Jackson et al., 2005). The recent emphasis on supporting all learners is critical given RtI's potential impact on the important proximal and distal factors affecting student learning; in particular, the quality of education students receive. A key feature across RtI models includes the provision of evidence-based, high quality instructional practices linked to positive student learning outcomes (e.g., Fuchs & Fuchs, 2006; Klinger & Edwards, 2006; Griffiths, Parson, Burns, VanderHeyden, & Tilly, 2007). Likewise, RtI has been defined as providing high-quality instruction, matched to student need, that integrates ongoing assessment of performance to make educational decisions (Batsche et al., 2006).

The emphasis within RtI on evidence-based instruction is warranted given the corpus of literature documenting the link between effective instructional practices and curricula and student learning (National Reading Panel, 2000; National Mathematics Advisory Panel, 2008). Beyond curricula, there exists over 60 years of literature documenting effective instructional practices and the general features of effective instruction that promote student learning (e.g., Bennet, 1988; Creemers, 1994; Good & Brophy, 1980; Harris, 1998; Hattie, Biggs, & Purdie, 1996; Marzano, 1998; Scheerens, 1992; Walberg, 1986; Wang, 1991; Wang, Haertel, & Wahlberg, 1993). Thus, effective instruction is a core tenet in the three tiers of an RtI model (i.e., tier 1/universal interventions, tier 2/secondary interventions, tier 3/tertiary interventions) to maximize academic achievement for all learners.

Combining RtI's focus on effective instruction with its foundational principles, calls for highly qualified educators who possess the knowledge and skills to teach students in all tiers of

the system (Barnett, 2004; Espinosa, 2002, National Research Council, 2001). Teacher effectiveness has been documented as one of the most important variables influencing student learning and numerous studies have demonstrated that differential levels of effectiveness impact student achievement (e.g., Connor, Son, Himdman, & Morrison, 2005; Darling-Hammond, 2000; Leigh, 2010; Nye, Konstantopoulos, & Hedges, 2004; Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004; Sanders, 2000). The necessity of effective teachers for implementing RtI models is further underscored by the reauthorization of No Child Left Behind and IDEA, which called for highly-qualified teachers in every classroom (IDEA, 2005). National attention on promoting effective teachers continues to grow in response to increasing RtI program implementation, as well as federal and statewide teacher evaluation and accountability initiatives (Bender & Shores, 2007; Jimerson, Burns, Vanderheyden, 2007).

Despite the role that high quality instruction plays in the implementation of RtI, prevailing models continue to emphasize student assessment, student progress monitoring, and student outcomes to the exclusion of support for teachers. Limited attention has been devoted to applying RtI principles to improving teacher competencies and enhancing teacher effectiveness (e.g., Crawford et al., in press; Pianta, Hamre, & Downer, 2011; Reddy, Fabiano, & Jimerson, 2013; Reddy & Dudek, 2014). The organizational framework of RtI assists school districts in promoting and managing student achievement, yet systems for enhancing teacher effectiveness have been historically absent in school districts.

The popular *Rush to Judgment* report criticized current teacher evaluation systems for being "superficial, capricious" and often failing to address instructional quality, professional development, and student outcomes (Toch & Rothman, 2008, p.1). The *Widget Effect* similarly concluded that modern educator evaluator systems failed to accurately capture differences in

teachers' instructional performance or identify teachers' professional development needs (Weisberg, Sexton, Mulhern, & Keeling, 2009). Researchers and practitioners have similarly identified common themes of failure that include limited practical feedback, limited utility in improving instructional practices, and a lack of shared understanding regarding effective teaching process (e.g., Danielson & McGreal, 2000; Gersten, Chard, & Baker, 2000; Heneman & Milanowski, 2004; Johnson, 1990; Marzano, Frontier, & Livingston, 2011; Medley & Coker, 1987). Current teacher evaluation systems also do not support or collect (valid) information on specific teacher practices (Reddy, Kettler, & Kurz, in press) and instead offer qualitative information typically used in summative assessments.

With the success of RtI systems hinging on high quality instruction, current educator evaluation processes and systems must emphasize ongoing, specific feedback that supports planning and implementation of effective instructional and classroom behavioral management practices. Educators need information about specific evidence-based effective instructional practice, how best to implement these practices in their classroom, and how to monitor their effectiveness (Reddy, Fabiano, Dudek & Hsu, 2013a; 2013b, 2013c). What educators need, is a *response to intervention that enhances their professional practices*.

Applying RtI principles to teacher evaluation offers a solution to current system failures. Promoting effective teaching, akin to student learning, requires the integration of assessment and intervention through ongoing progress monitoring of teacher classroom practices that can inform data-based decision making on how best to tailor interventions and supports to meet teachers' professional development needs. A multi-tiered formative assessment approach to teacher evaluation provides teachers with a mechanism for receiving timely, actionable, and data driven feedback to enhance their practice and promote positive student learning outcomes (Reddy,

Fabiano, Dudek, & Hsu, 2013c). An RtI system for teachers would include the same foundational principles but additionally integrating the focus on teachers. The system would include the following: (a) an evidence-based core curriculum of effective teaching practices, (b) universal screening through classroom observations and other measures of teacher practice, (c) a multi-tiered system of evidence-based interventions matched to teacher needs, (d) data-based decision making guiding professional development, and (e) progress monitoring of teacher classroom practice through formative and summative assessments.

This chapter describes a multi-tiered teacher formative assessment model embedded within an RtI framework and uses the Classroom Strategies Scale (CSS) as an illustration of one promising multi-dimensional measure of instructional and behavioral management practices designed to monitor teacher progress in meeting instructional goals. The theory, research, and evidence underlying the CSS is outlined in addition to its' application in a case example. Finally, directions for future research are described.

### **Multi-Tiered Teacher Formative Assessment Model**

A three-tiered model that embeds formative assessment at each tier can be utilized to enhance teacher instructional practices and provide additional support to teachers requiring more intensive or different support to improve their instruction and behavior management skills. Multi-tiered formative assessment of teachers provides a systematic process for examining the extent to which teachers are benefitting from instructional support and monitoring adaptations to the curriculum or instruction to support effective instruction and student learning. Within a public health framework, we hypothesize that Tier 1 assessment and support targets all educators rather than those at-risk for struggling with instruction and is designed to prevent or reduce new behavioral problems and/or academic failures from emerging in classrooms (Epstein et al., 2008;



Glover & DiPerna, 2007). Thus, Tier 1 formative assessment would focus on routine progress monitoring to promote teachers use of effective instructional and classroom management practices. Tier 1 formative assessment could also include the provision of high-quality instructional feedback that is integrated into routine teacher evaluation systems to promote effective instruction or changes in school wide policies to prevent behavior problems and promote effective instruction. Within this three-tiered model, we would expect the majority of teachers to effectively utilize Tier 1 formative assessment data to enhance their instruction and meet their instructional goals. However, for a subset of teachers, Tier 1 formative assessment will be necessary but insufficient to promote enhanced classroom practices. This may reflect level of experience, composition of students in the classroom, training, or other idiographic contextual variables that impact teacher practices (Shernoff et al. 2011a, 2011b).

Tier 2 assessment and support would be designed for teachers at-risk for struggling with behavior management or effective instruction and/or would target those teachers who make insufficient progress at the first tier. Tier 2 supports for teachers may include individualizing or small group professional development, including being paired with colleagues who are skilled in implementing specific evidence-based interventions designed to improve teachers' behavior or rate of learning a new skill. Mentor colleagues could use the CSS to monitor their peers' implementation of specific behaviors and skills, as well as provide targeted feedback to facilitate implementation. Tier 2 supports may also target teachers who are most at-risk for being socially isolated within the school because they are new (Shernoff et al., 2011b) and would benefit from being linked with experienced colleagues who are well positioned to support their professional development.

Tier 3 assessment and support would reflect increasing the duration or intensity of support to teachers who continue to struggle in meeting professional development goals. Visual performance feedback guided by CSS scores can be used to enhance teachers' use of evidence-based instructional and behavior management practices. Tier 3 support may include job-embedded instructional coaches who provide in vivo modeling, demonstration, and guided practice for those teachers needing more intensive classroom-based support (Joyce & Showers, 2002; Scott, 2012; Shernoff et al., 2011b). Coaching can embed a problem solving process that focuses on establishing mutually agreed upon goals and practical intervention plans to improve educators' classroom practices and student achievement. Although tier 3 supports are more expensive, they may be time limited and/or reserved only for those teachers who have not responded to Tier 1 and Tier 2 support.

### **The Classroom Strategies Scale: A Multi-Tiered Teacher Formative Assessment**

The CSS is a promising multi-dimensional measure of instructional and behavioral management practices designed to monitor teacher progress in meeting instructional goals. The theoretical and empirical basis for the CSS is described in detail below.

#### **Theoretical Foundation of the CSS**

The CSS was designed for school personnel to assess educators' use of evidence-based instructional and behavioral management strategies associated with positive student outcomes. The CSS includes two forms, an Observer Form, which is used by school personnel to evaluate teachers during classroom observations and the Teacher Form, which is used by teachers to self-reflect on their strategy usage following direct classroom observations. Both the Observer and Teacher Forms include similar information designed to: (a) enhance communication of evidence-

based instructional and behavioral management practices through complimentary perspectives and (b) identify specific practice goals linked to professional development and supports.

The CSS-Observer and Teacher Form constructs and strategies are grounded in over 60 years of effective teaching and behavioral management literature (e.g., Brophy & Good, 1986; Gage, 1978; Marzano, 1998; Marzano, Pickering, & Pollock, 2001; Wittrock 1986; Walberg, 1986). This body of work highlights the general features of effective instructional practice linked to positive student academic performance (e.g., Bennet, 1988; Creemers, 1994; Good & Brophy, 1980; Harris, 1998; Hattie, Biggs, & Purdie, 1996; Scheerens, 1992; Walberg, 1986; Wang, 1991; Wang, Haertel, & Wahlberg, 1993). The CSS scales and strategies are based on direct instruction, differentiated instruction, and constructivist models of teaching. The CSS includes two domains (instruction and behavior management) that encompass nine (subscales) dimensions related to positive student learning (see Table 1).

The CSS was developed based on several key theoretical assumptions. First, both instructional and behavioral management strategies are used dynamically to enhance classroom climate and student learning. Second, teachers use *sets of strategies* both in concert and in series to each other to foster student learning. Therefore, support to implement the evidence-based strategies should target sets of teaching strategies and the sequence of strategies to use in instruction. Third, the CSS dimensions (subscales) are assumed to be correlated based on prior research and theory that underscores that teaching is an interactive process that requires teachers to fluidly adapt strategies as lessons proceed (Clark & Peterson, 1986; Harris, 1998; Tomlinson & McTighe, 2006). Fourth, sets of strategies are context dependent to lesson format and student learning needs. The literature devoted to measuring differentiated instructional approaches has documented the effectiveness of tailoring learning interventions to meet individual student needs

(Ames, 1992). Finally, the fifth theoretical assumption asserts effective teaching is a continuous learning process which requires ongoing specific feedback and supports that inform the planning and implementation of effective instructional and behavioral management practices. Teachers want and need timely, specific, actionable, and data driven feedback to inform their instruction.

The CSS was developed as a user-friendly multidimensional assessment of instructional and behavioral management strategies. The CSS generates scores that: (1) assess educators' use of empirically supported instructional and classroom behavioral management strategies, (2) identify practice goals for improvement, (3) monitor educators' progress towards practice goals following intervention, (4) provide evidence for professional development and supports (e.g., professional learning committees), and (5) refine school-wide teacher professional development plans.

### **Development of the CSS**

The CSS-Observer and Teacher Forms were iteratively developed over the course of 4 years using contemporary test theory (e.g., Anastasi & Urbina 1997; Benson, 1998; Crocker & Algina, 1986; Kane, 2002; 2008). As noted, the CSS was designed specifically for use by school personnel in routine educational practice with the central goal to assess teachers' *use* of evidence-based instructional and behavioral management practices and to *inform* professional development and coaching efforts (e.g., Reddy & Dudek, 2014; Reddy, Fabiano, Dudek, & Hsu, 2013a).

The development process was guided by several methods: (1) expert input, (2) consumer input, (3) extensive field-testing with over 500 classrooms, and (4) a set of data analytic methods. Domains and strategy/items were developed through input from a national advisory board of experts in instruction, behavior management, and measurement, and a comprehensive

review of peer-review publications and other related tests. Consumer input (advisory boards of principals, general, and special education teachers) provided important feedback on the specific domains and strategies/items, as well as item ambiguity and possible bias. Face/content validity of the CSS was established in part through the expert and consumer advisory boards (please see Reddy, Fabiano, Dudek, & Hsu, 2013a for a detailed description). The boards were also encouraged to provide feedback on new domains and strategies/items and the CSS intended use and score utility for assessing practices and informing changes practices (i.e., professional development). Additionally, numerous data analytic methods were employed to refine the CSS scales and strategies/items such as item to total correlations, pooled mean item variances across observation (level of disagreement), as well as confirmatory factor analysis within observation using recommended fit indices (Jackson, Gillapsy, & Purc-Stephenson, 2009) and information-theory-based indices of relative fit (Bowen & Guo, 2012; please see Reddy, Fabiano, Dudek, & Hsu, 2013a; 2013b for details).

### **The CSS Model**

The CSS-Observer Form consists of three parts (see Table 1). For Part 1 Strategy Counts, observers tally each time eight instructional and behavior management strategies are used by the teacher during a classroom observation (lesson) and whether the strategy used was for individual students or group of students (i.e., two or more students). For Part 2 Strategy Rating Scales, observers complete the Instructional Scale (IS) and Behavioral Management Strategies (BMS) scale after classroom observations. After the completion of the Part 2 Strategy Rating Scales, observers complete Part 3 the Classroom Checklist which includes 14 items that assess the presence of classroom structural procedures. The CSS-Teacher Form does not include the Part 1

Strategy Counts, but contains the exact same Part 2 Strategy Rating Scales items for IS and BMS, as well as the same Part 3 Classroom Checklist items.

**Part 1 Strategy Counts - Observer Form.** Strategy Counts includes eight strategies that are nested in the domains of instructional and behavioral management practices. Concept summaries, academic response opportunities, academic praise, and academic corrective feedback fall under instruction, while one to two step directives, vague directives, behavioral praise, and behavioral corrective feedback fall under behavior management.

Concept summaries are defined as the teacher summarizing or highlighting key concepts or facts (steps) taught throughout the lesson (not at the beginning or end of lessons). Examples may include “we learned today that a hypothesis is a scientist’s best guess about how an experiment will turn out” and “DeShawn, to find a cube’s volume we multiply the length times the width, times the height”. The use of concept summaries has been linked to enhance students’ understanding of the lesson and recall and organization of learning material (e.g., Brophy & Alleman, 1991; Brophy 1998; Hines, Cruickshank, & Kennedy, 1985; Rosenshine & Stevens, 1986). Techniques that reinforce key concepts and facts (concept summaries) are important for students with executive functioning impairments and/or auditory processing disorders as these students often require lesson content to be repeated and emphasized for successful integration and skill application (Reddy, Newman, & Verdesco, in press). A general guideline for educators is to provide brief concept summaries every two to three minutes during learning activities.

Academic response opportunities are an important strategy for teachers to encourage students to share their ideas and understanding of lesson content in class. Academic response opportunities can be verbal or nonverbal responses (e.g., explain answers, repeat key points, brainstorm ideas, and show answers on the board). Research on “opportunities to respond”

(OTR) among students with behavior disorders underscores the utility of teachers creating opportunities for their students to respond to questions and learning activities during instructional time (e.g., Partin et al., 2010; Sutherland, Alder, & Gunter, 2003; Stichter, Lewis, Whittaker, Richter, Johnson, & Trussell, 2009). Research has found that increasing OTR and praising students for effort and/or correct responses can lead to higher levels of both on task behavior, prosocial behavior, and correct student answers (Sutherland, Wehby, & Yoder, 2002). The recommendation is that teachers elicit 4 to 6 responses per minute when teaching new learning material and 9 to 12 responses per minute when reviewing previously taught learning material (practice or drill work).

Academic feedback (i.e., praise and corrective feedback) is important for instruction and student learning (e.g., Council for Exceptional Children; Bender, 2008; Gable, Hester, Rock, & Hughes, 2009; Tomlinson & Edison, 2003). Praise for academic performance is verbal or nonverbal statements or gestures provided by teachers to individual or groups of students immediately following academic responses. Likewise, corrective feedback for academic performance is verbal or nonverbal statements or gestures provided by teachers to individual or groups of students immediately following incorrect academic responses.

Praise statements should be frequent, immediate, enthusiastic, and specifically describe the behaviors of the student's success. Praise statements that are implemented consistently will orient students towards a better appreciation of students' own task-related behavior (Brophy, 1981; Gable et al., 2009). Similarly, academic corrective feedback for students' incorrect academic responses should be immediate and explain what is specifically incorrect about their answers (Hattie & Timperely, 2007). Simply telling children their answer is right or wrong is not enough. Students need to be told what specifically is correct or incorrect about their answers

(e.g., “Your answer is incorrect because you forgot to carry over the one when you added the tens column”). A common sequence of instructional strategies is to first provide an academic response opportunity followed by a praise statement or corrective feedback for academic performance. For all students, academic response opportunities and academic feedback (praise and correct feedback) are critically important for encouraging and monitoring task completion.

Effective directives are important for enhancing appropriate behavior and follow through on tasks (Kern & Clemens, 2007). Clear one to two step directives are brief verbal instructions that direct specific student behavior. Directives are clear, declarative statements (not questions or favors) that specifically describe the desired behavior in no more than two steps. Directives are most effective when phrased as “do” commands (i.e., telling children what they should do rather than what they should not). On the other hand, vague directives are verbal directives that are unclear, issued as questions, and include unnecessary verbalizations (more than two steps).

For the CSS, behavioral feedback consists of two strategies: praise for appropriate behavior and corrective feedback for inappropriate behavior. Praise for appropriate behavior is verbal or nonverbal statements or gestures provided by teachers to individuals or groups of students immediately following appropriate behavior. As with praise for academic performance, key features of effective praise statements for behavior are high frequency, immediacy, and specifically labeling the appropriate behavior. Research has shown that levels of on task behavior significantly increase when students were given specific praise about their behavior compared to simply positive praise (i.e., Good Job; Chalk & Bizo, 2004; Brophy, 1981). In contrast, corrective feedback for inappropriate behavior is verbal or nonverbal statements or gestures provided by teachers to redirect inappropriate behavior. Like praise, corrective feedback should be specifically labeled and given after inappropriate behavior is observed (Bangert-Down, Kulik,



Kulik, & Morgan, 1991). Research has recommended that teachers should provide approximately *one* corrective feedback statement for *every three* praise statements in the classroom (White, 1975; Stichter et al., 2008).

**Part 2 Strategy Rating Scales and Part 3 Classroom Checklist – Observer and Teacher Form.** Following classroom observations, observers [teachers] complete the Part 2 Strategy Rating Scales, IS and BMS scales (see Table 1). The IS scale includes 28 items that comprise a total scale, two composite scales, and five subscales. The Instructional Methods Composite scale (17 items; maximum frequency score of 119) consists of the Direct Instruction (8 items; maximum score of 56), Adaptive Instruction (4 items; maximum score of 20) and Student Focused Instruction (6 items; maximum score of 42) subscales. The Academic Monitoring/Feedback Composite scale (11 items; maximum score of 77) consists of the Promotes Student Thinking (5 items; maximum score of 35) and Academic Performance Feedback (6 items; maximum score of 42) subscales.

The BMS scale includes 26 items that composes a total scale, two composite scales, and four subscales. The Behavioral Feedback Composite scale (12 items; maximum frequency score of 84) consists of Praise (5 items; maximum score of 35) and Corrective Feedback (7 items; maximum score of 49) subscales. The Proactive Methods Composite scale (14 items; maximum score of 91) consists of Prevention Management (8 items; maximum score of 56) and Directives (6 items; maximum score of 42) subscales.

After classroom observations, observers [teachers] rate how often (*Frequency Rating*) teachers used specific instructional and behavioral management strategies on a 7-point Likert scale (1 “never used”, 4 “sometimes used”, 7 “always used”) and then rate how often the teachers *should have* used each strategy (*Recommended Frequency*) on a 7-point Likert scale (1

“never used”, 4 “sometimes used”, 7 “always used”). The Part 2 Rating Scales produce both frequency scores and discrepancy scores. For the Part 2 Strategy Rating Scales, item discrepancy scores are computed as follows:  $|\text{recommended frequency} - \text{frequency ratings}|$ . In sum, absolute value discrepancy scores indicate if any change (regardless of direction) was needed as measured by the observer [teacher] using the CSS. Larger discrepancy score values indicate greater amounts of change are needed in the practices measured by the CSS (see Reddy et al., 2013a for details on scoring). After completing Parts 1 and 2, the observer [teachers] completes the Classroom Checklist (Part 3). The Classroom Checklist assesses the presence of 14 specific items or procedures in the classroom (see Table 2).

### **Training**

The CSS-Observer Form training consists of a five step process focused on content knowledge and direct observational skills for the specific components of the CSS. Training includes: (1) group training, (2) knowledge testing, (3) practice coding, (4) criterion testing, and (5) report interpreting, coaching training.

Training includes:

- (1) Didactic group trainings from a CSS Trainer/Master Coder which include discussions of definitions, criteria, and orientation to the scientific literature guiding the development of the CSS and the recommended frequencies of Part 2 strategies.
- (2) Completion of knowledge tests for CSS Part 1-3.
- (3) Practice coding classroom videos using the CSS and review of practice results by a CSS Trainer/Master Coder.

(4) Completion of five classroom video coding criterion tests using the CSS.

Certification as a qualified users is set by reaching the minimum inter-rater reliability level of 80% with CSS Trainer/Master Coders.

(5) CSS score report interpretation and coaching.

### **Psychometric Characteristics**

**Construct Validity.** For the CSS Parts 1 through 3, the constructs and strategies/item were extensively developed through on-going expert and consumer advisory board feedback, pilot testing, and advance statistical analyses. All three parts represent important instructional and behavioral management strategies associated with effective teaching linked to general and special education student academic and behavioral outcomes (e.g., Reddy & Dudek, 2014; Reddy, Fabiano, Barbarasch, & Dudek, 2012; Reddy, Newman, & Verdesco, in press a; in press b).

The Part 2 IS and BMS scales are theoretically and factor analytically derived (confirmatory factor analysis) within classroom observations. The CSS factor structure was examined with over 12 confirmatory factor analyses using generalized least squares estimation (SPSS's AMOS Version 19 software, Arbuckle, 2010). As described in Reddy, Fabiano, Dudek, and Hsu (2013a), several fit indices (e.g.,  $\chi^2/df$ , Root Mean Square Error of Approximation (RMSEA), adjusted goodness of fit index (AGFI), and goodness of fit index (GFI) recommended by Jackson, Jackson, Gillapsy, and Purc-Stephenson (2009) were used to test the fit to the data. As noted, the CSS IS and BMS subscales were conceptualized as being correlated factors based on theory and research that teachers use a blend of strategies and sets of strategies in the teaching and learning process. In addition, CSS first order 4-factor models (IS and BMS preferred factor models) were compared to alternative models including 5- and 6-factor models using

information-theory-based indices of relative fit (i.e., Akaike Information Criterion (AIC), Brown-Cudeck Criterion (BCC), Schwarz Bayesian Information Criterion (BIC)) described by Bowen and Guo (2012). Overall, results indicated that CSS four-, five- and six-factor models yielded good fit to the data and superior fit to the data in comparison to alternative models using information-theory-based indices of relative fit (see Reddy, Fabiano, Dudek, & Hsu, 2013a). The CSS-Teacher Form also yielded good fit to the data and superior fit to the data in comparison to alternative higher order factor models (Reddy, Dudek, & Fabiano, in progress).

**Reliability and Validity.** In the high-stakes era of teacher evaluation, evidence of reliability and validity for new teacher measures is critically important for evaluating teacher effectiveness and informing professional development/coaching models. The reliability of the CSS has been examined in four studies that assess the internal consistency, inter-rater reliability, test retest reliability, and freedom from item bias.

The CSS was found to demonstrate good internal consistency (Cronbach alphas of .92-.93) across Parts 1 through 3 (see Reddy et al., 2013a for further details). Inter-rater reliability data was found to have good inter-rater reliability estimates for Parts 1-3. For example, inter-rater reliability for the Part 1 Strategy Counts was  $r = .94$  (percent agreement 92%), Part 2 IS and BMS Strategy Rating Scales was  $r = .80$  and  $r = .72$  (percent agreement 92% and 88%), and Part 3 Classroom Checklist was  $r = .86$  (percent agreement 91%). The CSS inter-rater reliability estimates are consistent with accepted values for other classroom observation assessments such as the measures used in the Measures of Effective Teacher Project (Cantrell, 2013; Kane & Staiger, 2012) and the Classroom Assessment Scoring System (CLASS; Pianta, Le Paro, Hamre, 2008). Furthermore, the CSS has been found to have fair to good test-retest reliability (approximately 2 to 3 weeks). For example, an  $r$  of  $.70$  (percent agreement 81%) was found for

the Part 1 Total Behaviors,  $r$ s of .86 and .80 (percent agreement 93% and 85%) for the Stage 2 IS and BMS Total scales, and  $r$  of .77 (percent agreement was 81%) for the Stage 3 Classroom Checklist. Part 2 Strategy Rating Scales items evidenced freedom of item bias for teacher age, educational degree, and years of teaching experience using differential item functioning analyses (Reddy, Fabiano, Dudek, & Hsu, 2013a). Similar results have been found for the CSS-Teacher Form (Reddy, Dudek, Fabiano, in progress).

**Validity.** The CSS-Observer Form evidences convergent and discriminant validity, as well as predictive validity. The CSS was compared to the Classroom Assessment Scoring System (CLASS), a well-established measure of teacher and classroom quality (Pianta, La Paro, & Hamre, 2008). As hypothesized, the CSS corresponded with logically related CLASS domains (e.g., Behavior Management) and it did not correspond with domains hypothesized to be unrelated (e.g., Language Modeling), suggesting that the CSS yields good convergent and discriminant validity with the Classroom Assessment Scoring System (Reddy, Fabiano, & Dudek, 2013). Results highlight the unique features of the CSS for assessing teacher classroom practices. Similar results have been found with the CSS-Teacher Form (Reddy & Dudek, in progress).

The CSS also evidences good predictive validity with student academic outcomes. Using a series of two level hierarchical linear modeling, the CSS IS scale discrepancy scores uniformly predicted student mathematics and language arts statewide testing scores for 663 third, fourth, and fifth graders (Reddy, Fabiano, Dudek, & Hsu, 2013b). Findings revealed that IS scale discrepancy scores significantly predicted mathematics and English language arts proficiency scores: relatively larger discrepancies on observer ratings of what teachers did versus what should have been done were associated with lower proficiency scores. In sum, the greater the

need for changes in practices exhibited by a teacher, the greater number of students receiving lower proficiency scores.

For mathematics statewide assessments, odds ratios indicated that an increase in one standard deviation of IS discrepancy scores (i.e., Total, Instructional Methods and Academic Monitor/Feedback Composite Scales) reduced the odds of success (i.e., students scoring proficient or above proficiency) by approximately 25%. Likewise, for English language arts statewide assessments, odd ratios revealed that an increase in one standard deviation of IS discrepancy scores (i.e., Total, Instructional Methods Composite) reduced the odds of success by approximately 30%. Results support the predictive validity of the CSS and have important implications for professional development. Preliminary results offer support for the predictive validity of the CSS-Teacher Form (Reddy & Dudek, in preparation).

### **CSS Case Example**

The following case example demonstrates the utility of embedding teacher formative assessment within a tiered model to support teacher development and enhance instructional practices (see Table 2). The CSS provides an illustration of a progress monitoring instrument that supports and promotes data-based decision making regarding instructional practices and professional performance.

John is a 41 year old, Caucasian male teacher with a Bachelor's Degree in Elementary Education. He has 13 years of experience as a teacher in elementary school settings and has worked in his current position as a 5<sup>th</sup> grade teacher for the past ten years. He has predominantly taught the Language Arts component of the fifth grade curriculum at his district while his partner teacher instructs students in the Math Science and Social Studies components. John teaches to one group of students in the morning (9am to 11:30am) and to a second group of students in the

afternoon (12:30pm to 3pm). John and his partner teacher switch classrooms after recess. John's afternoon classroom is composed of twenty-five general education students. Although no students are classified with a specific learning disability, John noted there are six students evidencing academic and behavioral learning difficulties.

John's school employs an RtI model for its students and recognizes how this model can support teachers. To help promote effective instruction in the classroom and support teacher development, the school has recently incorporated this model into its comprehensive human capital management system (HCMS; Allen & Odden, 2008). In its first Tier, the system relies on routine educator evaluation for providing feedback to teachers on their instructional practices. This year John will receive a series of four observations, one per quarter, by a qualified evaluator utilizing the CSS.

For his first observation in October, John and the evaluator complete the CSS. John utilizes the CSS-Teacher Form and the evaluator observes John using the CSS-Observer Form. Following the observation, John and the evaluator discuss their findings. On the CSS-Teacher Form, John indicated that he utilized behavioral corrective feedback very frequently and thought there was significant need to use it at a high frequency due to the frequency of inappropriate behaviors in the classroom. The evaluator agreed with John's assessment but also noted on the CSS-Observer Form that John rarely used proactive methods of managing students' behavior and used behavioral praise at a low frequency. The evaluator discussed with John the benefits of using these strategies to help shape a positive learning environment and reduce inappropriate behaviors in the classroom. The evaluator and John devise an operationalized behavioral plan for implementing these strategies in the classroom and set a target goal for John's next evaluation in January.

For the second observation in January, the same procedure is repeated by John and the evaluator. The results were highly similar to the first observation: frequent behavioral corrective feedback but minimal use of proactive strategies and behavioral praise. During the post observation conference John and the evaluator discuss the target goals established for this observation. Although John attempted to utilize the strategies recommended in the behavioral plan, John agreed with the evaluator that these were unsuccessful and that he needs additional support.

As part of the second tier of the HCMS, the school district has a collaborative teaching mentor program that pairs both new and struggling teachers with an experienced teacher that excels at using evidence-based strategies. The evaluator helps John select a teacher that exemplifies effective classroom behavioral management strategies. Over the course of several weeks, the mentor teacher observes John's classroom and uses the CSS to gather information. John and the mentor teacher meet collaboratively after school several times to discuss effective behavior management strategies and to specifically work on proactive strategies and behavioral praise. For each meeting, the mentor teacher graphs John's use of behavioral management strategies on the CSS to help John monitor his progress in practicing and implementing these strategies. After several weeks, the graphs show John increasing his use of effective proactive strategies and behavioral praise. The graphs also show decreased use of behavioral corrective feedback.

For the third observation that year, both John and the evaluator notice improvement on the CSS. Through increased use of proactive strategies and behavioral praise, John was able to create a better learning environment in his classroom and reduce the amount of inappropriate student behaviors. The evaluator congratulates John on his efforts and encourages him to



continue implementing these evidence-based strategies. For his fourth observation, John received very positive reviews, especially in the area of behavior management.

Had efforts at the second tier of intervention not worked, John would have been assigned more intensive interventions and supports associated with the third tier of intervention. The third tier of interventions might have included working with a professional development coach external to the school district or enrolling John in professional development courses for effective teaching. Additionally, more intensive procedures like in-vivo modeling or shared teaching might also have been used to help facilitate John's competencies in these skills.

Routine progress monitoring in John school's district allowed for early identification and remediation in John's behavioral management practices. Had a more traditional evaluation system model been used, which historically requires one observation per year for tenured teachers, John's difficulties in behavior management competencies might not have been noticed until it was too late for remediation to yield gains in the classroom (Toch & Rothman, 2008; Weisberg, Sexton, Mulhern, & Keeling, 2009). Subsequently, increasing the level of support required by John allowed for tailored and specific professional development to occur. Ultimately, by implementing a multi-tiered formative assessment model for teacher development, both teachers and their students will benefit from the enhanced instruction in the classroom.

### **Future Research and Practice**

Given that multi-tiered teacher formative assessment models are slowly emerging, several areas for future research and practice are needed to better understand the impact of tiered assessment and support for teachers (Table 2). For example, further studies must establish reliable and valid decision-making criteria for determining progression through the three tiers and what constitutes adequate progress within an RTI assessment and intervention model for

teacher professional development (Glover & DiPerna, 2007). Given the prominent role that formative assessment models play in our educational system, further research in this area has implications for assessing student progress over time as well as teachers. Although, a significant potential benefit of an RTI framework is its utility in determining teacher responsiveness to support, it is unclear what the infrastructure of support should be for those teachers who continue to struggle despite intensive intervention across tiers. Given a corpus of research supports the importance of high-quality implementation of evidence-based practices, more research is needed to establish methods and measures that maximize the uptake, maintenance, and implementation of evidence-based interventions in schools (Becker, Bradshaw, Domitrovich & Jalongo, 2013). Several gaps in the coaching literature warrant further attention, including a better understanding of the functional components and key elements of coaching (e.g., Blachowicz, Fogelberg & Obrochta, 2005; Kretlow & Bartholomew, 2010) and the key instructional elements (e.g., modeling, demonstration, observation) utilized (Shernoff et al., under review). In addition, despite the important role that coaching can play in supporting teachers in need of tier 3 support, existing studies provide very limited insight into the feasibility of coach-based professional development and the extent to which such models can be sustained over time (Becker et al., 2013; Nadeem et al., 2013; Shernoff et al., under review).

## **Conclusion**

Multi-tiered teacher formative assessment models are a viable alternative to traditional, high stakes, summative assessment of teachers and are aligned with RTI models being adopted in the U.S. This chapter underscores the importance of teacher assessment, ongoing feedback, and providing a continuum of support for teachers in order to maximize teacher professional development and student learning. The CSS was illustrated as one promising tool that could be

used to develop meaningful evidence-based feedback for teachers and improve the quality of implementation of evidence-based interventions used in their classrooms. Coaching was described as one mechanism for providing more intensive, job-embedded support to those teachers requiring more intensive support.

## References

- Alber, S. R., Heward, W. L., & Hippler, B. J. (1999). Teaching middle school students with learning disabilities to recruit positive teacher attention. *Exceptional Children, 65*, 253-270.
- Ames, C. (1992). Classrooms: Goals, structures and student motivation. *Journal of Educational Psychology, 84*, 261-271.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall, Inc.
- Arbuckle, J.L. (2010). *IBM SPSS Amos 19 User's Guide, 587-590*. Crawfordville, FL: Amos Development Corporation.
- Avalos, B. & J. Assael (2006). Moving from resistance to agreement: The case of the Chilean teacher performance evaluation. *International Journal of Educational Research, 45*, 254-266.
- Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213-238.
- Barnett, W. S. (2004). *Better teachers, better preschools: Student achievement linked to teacher qualifications*. Retrieved November 30<sup>th</sup>, 2013, from <http://nieer.org/resources/policybriefs/2.pdf>
- Batsche, G., Curtis, J., Dorman, C. Castillo, J., & Porter, L. J. (2008). The florida problem solving/response to intervention model: Implementing a statewide initiative. In S. Jimerson, M. Burns, & A. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 378-395). New York: Springer.

- Batsche, G., Elliott, J., Graden, J. L., Grimes, J., Kovaleski, J. F., Prasse, D., et al. (2005). *Response to intervention policy considerations and implementation*. Reston, VA: National Association of State Directors of Special Education.
- Beaman, R., & Wheldall, K. (2000). Teacher's use of approval and disapproval in the classroom. *Educational Psychology, 20*, 431-446.
- Becker, K. D., Bradshaw, C. P., Domitrovich, C., & Ialongo, N. S. (2013). Coaching teachers to improve implementation of the good behavior game. *Administration and Policy in Mental Health and Mental Health Services Research, 48*(2), 482-493.
- Bender, W. N., & Shores, C. (2007). *Response to intervention: A practical guide for every teacher*. Thousand Oaks, CA: Corwin Press.
- Bennet, N. (1988). The effective primary school teacher: The search for a theory of pedagogy. *Teaching and Teacher Education, 4*, 19-30.
- Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 1*(7), 10-22.
- Bergan, J. R., & Kratochwill, T. R. (1990). *Behavioral consultation and therapy*. New York: Plenum.
- Bender, W. N. (2008). *Differentiating instruction for students with learning disabilities: Best teaching practices for general and special educators* (2nd ed.) City Thousand Oaks, CA: Corwin Press.
- Blachowicz, C. L., Obrochta, C., & Fogelberg, E. (2005). Literacy coaching for change. *Educational Leadership, 62*(6), 55.
- Bowen, N. K., & Guo, S. (2012). *Structural equation modeling*. New York: Oxford University Press.

Brophy, J.E. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, 51 (1), 5-32.

Brophy, J. (1998). *Motivating students to learn*. New York, NY: McGraw-Hill.

Brophy, J., & Alleman, J. (1991). A caveat: Curriculum integration isn't always a good idea. *Educational Leadership*, 49(2), 66.

Brophy, J. E., & Good, T. (1986) Teacher behavior and student achievement. In M.C. Wittrock (Ed.), *Handbook of research in teaching*. (3rd ed., pp. 328–375). New York: Macmillan.

Cantrell, S. (2013). Ensuring fair and reliable measures of effective teaching. Bill & Melinda Gates Foundation.

Cantrell, S., & Kane, T. J. (2013). *Ensuring Fair and Reliable Measures of Effective Teaching: Culminating Findings from the MET Project's Three-Year Study*. Retrieved from [www.metproject.org/downloads/MET\\_Ensuring\\_Fair\\_and\\_Reliable\\_Measures\\_Practitioner\\_Brief.pdf](http://www.metproject.org/downloads/MET_Ensuring_Fair_and_Reliable_Measures_Practitioner_Brief.pdf)

Carlisle, J.F., Correnti, R., Phelps, G., & Zeng, J. (2009). Exploration of the contribution of teachers' knowledge about reading to their students' improvement in reading. *Reading and Writing: An Interdisciplinary Journal*, 22, 459-486

Chalk, K., & Bizo, L. A. (2004). Specific praise improves on-task behavior and numeracy enjoyment: a study of 4 pupils engaged in the numeracy hour. *Educational psychology in Practice*, 20, 335-351.

Chunn, M., & Witt, J. (2008). System to enhance educational performance. Retrieved November 25, 2013, from <http://www.isteep.com/index.html>.

Clark, C.M., & Peterson, P.L. (1986). Teachers' thought processes. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3<sup>rd</sup> ed.). New York: Macmillan.

- Connor, C.M., Son, S.H., Hindman, A., & Morrison, F.J. (2005). Teacher qualifications, classroom practices, and family characteristics: Complex effects on first graders' language and early reading. *Journal of School Psychology, 43*, 343-375.
- Crawford, A.D., Zucker, T.A., Williams, J.M., Bhavsar, V., & Landry, S.H. (in press). Initial validation of the pre-kindergarten classroom observational tool and goal setting system for data-based coaching. *School Psychology Quarterly*.
- Creemers, B.P.M. (1994). *The effective classroom*. London: Cassell.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt, Rinehart, and Winston.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Educational Policy Analysis and Archives, 8*(1).
- Danielson, C., & McGreal, T.L. (2000). *Teacher evaluation to enhance professional learning*. Princeton, NJ: Educational Testing Service.
- Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Education Evaluation and Policy Analysis, 24*, 81-112.
- Dwyer, C. A., & Stufflebeam, D. (1996). Teacher Evaluation. In d. Bergliner & R. Calfee (Eds.), *Handbook of Educational Psychology*. New York: Macmillan.
- Epstein, J. L., Sanders, M. G., Sheldon, S. B., Simon, B. S., Salinas, K. C., Jansorn, N. R., . . . & Williams, K. J. (2008). *School, Family, and Community Partnerships: Youth Handbook for Action, Third Edition*. Thousand Oaks, California: Corwin Press.

- Espinosa, L. M. (2002, November). *Preschool policy matters; High-quality preschool: Why we need it and what it looks like*. Retrieved November 30, 2013, from <http://nieer.org/resources/policybriefs/1.pdf> .
- Farrell, P., Jimerson, S., Kalambouka, A., & Benoit, J. (2005). Teachers' perceptions of school psychologists in different countries. *School Psychology International*, 26 (5), 525-544.
- Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly*, 41, 93-99.
- Gable, R., Hester, P., Rock, M., & Hughes, K. (2009). Back to basics: Rules, praise, ignoring, and reprimands revisited. *Intervention in School & Clinic*, 44(4), 195.
- Gage, Nathaniel L. (1978). *The scientific basis of the art of teaching*. New York: Teachers College Press.
- Goldenberg, C., & Gallimore, R. (1991). Changing teaching takes more than a one-shot workshop. *Educational Leadership*, 49 (3), 69-72.
- Goldschmidt, P. & Phelps, G. (2010). Does teacher professional development affect content and pedagogical knowledge: How much and for how long? *Economics of Education Review*, 29, 432-439.
- Good, T. L., & Brophy, J. (1980). *Educational psychology: a realistic approach (2nd ed.)*. New York, NY: Holt, Rinehart, & Winston.
- Gersten, R., Chard, D.J., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities* , 33, 453.
- Harris, A. (1998). Effective teaching: A review of the literature. *School Leadership & Management*, 18, 169-183.



- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 1, 81-112.
- Hattie, J., Biggs, J., Purdie, N. (1996). Effects of learning skills interventions on students learning: A meta-analysis. *Review of Educational Research*, 66 (2), 99-136.
- Heneman, H.G.III, & Milanowski, A. (2004). Alignment of human resource practices and teacher performance competency. *Peabody Journal of Education*, 79(4), 108-125.
- Hines, C. V., Cruickshank, D. R., & Kennedy, J. J. (1985). Teacher clarity and its relationship to student achievement and satisfaction. *American Educational Research Journal*, 22, 87-99.
- Horner, R.H., Sugai, G., Todd, A.W., & Lewis-Palmer, T. (2000). Elements of behavioral support plans: A technical brief. *Exceptionality*, 8, 205-215.
- Horner, R.H., Sugai, G., Todd, A.W., & Lewis-Palmer, T. (2005). School-wide positive behavior support. In L. Bambara & L. Kern (Eds.) *Individualized supports for students with problem behaviors: Designing positive behavior plans*. (pp. 359-390) New York: Guilford Press
- Individuals with Disabilities Education Act of 2005* 70 (118) U.S.C. 1221 e-3.
- Isore, M. (2009). Teacher Evaluation: Current practices in OECD countries and a literature review. *OECD Education Working Papers* No. 23. Paris: OECD Publishing. Retrieved from: <http://www.oecd.org/edu/workingpapers>
- Jackson, D.L., Gillapsy, J.A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14, 6-23.
- Jackson, S., Pretti-Frontczak, K. Harsjusola-Webb, S. Grisham-Brown, & Romani, J, M. (2009).

- Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (2007). *Handbook of response to intervention: The science and practice of assessment and intervention*. New York: Springer.
- Johnson, S.M. (1990). *Teachers at work: Achieving success in our schools*. New York: Basic Books.
- Kane, M. (2008). Terminology, emphasis, and utility in validation [Peer Commentary on the paper A suggested change in terminology and emphasis regarding validity and education]. *Educational Researcher*, 37 (2), 76-82.
- Kane, M. (2002). Validating high-stakes testing programs. *Educational Measurement: Issues and Practice*, 21, 31-41.
- Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. MET Research Paper. Seattle, Washington: Bill & Melinda Gates Foundation (Retrieved July 16, 2012, from [http://www.metproject.org/downloads/MET\\_Gathering\\_Feedback\\_Research\\_Paper.pdf](http://www.metproject.org/downloads/MET_Gathering_Feedback_Research_Paper.pdf).)
- Kauchak, D., Peterson, K., & Driscoll, A. (1985). An interview study of teachers' attitudes toward teacher evaluation practices. *Journal of Research and Development in Education*, 19 (1), 32-37.
- Kern, L., & Clemmens, N.H. (2007). Antecedent Strategies to Promote Appropriate Classroom Behavior. *Psychology in the Schools*, 44, 65-75.
- Klinger, J.K., & Edwards, P.A. (2006). Cultural considerations with Response to Intervention models. *Reading Research Quarterly*, 41, 108-117.

- Kounin, J. S. (1970). *Discipline and group management in classrooms*. New York, NY: Holt, Rinehart, and Winston.
- Leflot, G., van Lier, P.A.C., Onghena, P., Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology*. 38 (6), 869-882.
- Leigh, A. (2010). Estimating teacher effectiveness from two-year changes in students' test scores. *Economics of Education Review*, 29, 480-488.
- Margo, J., Benton, M., Withers, K., & Sodha, S. (2008). *Those Who Can?* Institute for Public Policy Research (IPPR) Publications.
- Marzano, R. J. (1998). *A theory-based meta-analysis of research on instruction*. Aurora, CO: Mid-continent Research for Education and Learning. (Eric Document Reproduction Service No. ED 427 087).
- Marzano, R.J., Frontier, T., & Livingston, D. (2011). *Effective supervision: Supporting the art and science of teaching*. Alexandria VA: ASCD.
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Medley, D.M., & Coker, H. (1987). The accuracy of principals' judgments of teacher performance. *Journal of Educational Research*, 80(4), 242-247.
- Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice*, 24, 4, 186-195.
- Nadeem, E., Gleacher, A., & Beidas, R. S. (2013). Consultation as an implementation strategy

for evidence-based practices across multiple contexts: Unpacking the black box. *Administration and Policy in Mental Health and Mental Health Services Research*, 530-540.

National Association of State Directors of Special Education. (2007). *Policy considerations and implementation and response to intervention: NASDSE and CASE white paper on RTI*. Available from [www.nasdse.org](http://www.nasdse.org).

National Center on Student Progress Monitoring (2006). Retrieved March 9, 2013, from [www.studentprogress.org](http://www.studentprogress.org).

National Center on Response to Intervention. (2010). *Essential components of RTI – A closer look at response to intervention*. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Council of Teachers of Mathematics. Executive Summary of Principles and Standards for School Mathematics. [http://www.nctm.org/standards/12752\\_exec\\_pssm.pdf](http://www.nctm.org/standards/12752_exec_pssm.pdf)

National Reading Panel (2000). *Report of the national reading panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. (NIH Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Nye, B., Konstantopoulos, S., & Hedges, L.V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26 (3), 237-257.

Organization for Economic Co-Operation and Development. OECD. (2009). *Teacher evaluation: A conceptual framework and examples of country practices*. Paris: OECD.

Office for Standards in Education. (2006). *The logical chain: continuing professional*

- development in effective schools*. United Kingdom: OFSTED Publications.
- Partin, T., Robertson, R., Maggin, D., Oliver, R., & Wehby, J. (2010). Using teacher praise and opportunities to respond to promote appropriate student behavior. *Preventing School Failure*, 54, 172–178.
- Peterson, D. R. (1995). The reflective educator. *American Psychologist*, 50, 975-983.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System [CLASS] Manual: Pre-K*. Baltimore, MD: Brookes Publishing.
- Pianta, R. C., Hamre, B. K., & Downer, J. (2011). Aligning measures of quality with professional development goals and goals for children's development. In M. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Eds.), *Quality measurement in early childhood settings* (pp. 297-315). Baltimore, MD: Paul H. Brookes Publishing Co.
- Pochard, M. (2008). Livre vert sur l'évolution du métier d'enseignant. Rapport au ministre de l'Education nationale, La Documentation française, Collection des rapports officiels.
- Porter, A.C., Youngs, P., & Odden, A. (2001). Advances in teacher assessment and their uses. In V. Richardson (Ed.), *Handbook of Research on Teaching* (4<sup>th</sup> ed., pp.259-297).
- Riley-Tillman, & T.C., Eckert, T. (2001). Generalization programming and school based consultation. *Journal of Educational and Psychological Consultation*, 12, 217-241.
- Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- Reddy, L.A., & Dudek, C. (2014). Teacher progress monitoring of instructional and behavioral management practices: An evidence-based approach to improving classroom practices. *International Journal of School and Educational Psychology*.

- Reddy, L., & Fabiano, G., & Jimerson, S. (2013). Assessment of General Education Teachers' Tier 1 Classroom Practices: Current Science and Practice. *School Psychology Quarterly*.
- Reddy, L., Fabiano, G., Barbarasch, B., & Dudek, C. (2012). Behavior management of students with Attention-Deficit/Hyperactivity Disorders using teacher and student progress monitoring. In L.M. Crothers & J.B. Kolbert (Eds). *Understanding and managing behaviors of children with psychological disorders: A reference for classroom teachers*. New York: Continuum International Publishing Group, Inc.
- Reddy, L.A., Fabiano, G., & Dudek, C. (2013). Concurrent validity of the Classroom Strategies Scale-Observer Form. *Journal of Psychoeducational Assessment*. 31 (3), 258-270.
- Reddy, L.A., Fabiano, G., Dudek, C., & Hsu, L. (2013c). Instructional and behavior management practices implemented by elementary general education teachers. *Journal of School Psychology*.
- Reddy, L.A., Fabiano, G., Dudek, C., & Hsu, L. (2013a). Development and construct validation of the Classroom Strategies Scale-Observer Form. *School Psychology Quarterly*, 28, 317-34.
- Reddy, L.A., Fabiano, G., Dudek, C.M., & Hsu, L. (2013b). Predictive validity of the Classroom Strategies Scale-Observer Form on statewide testing. *School Psychology Quarterly*, 28, 301-316.
- Reddy, L.A., Kettler, R.J., & Kurz, A. (in press). School-wide educator evaluation for improving school capacity and student achievement in high poverty schools: Year 1 of the school system improvement project. *Journal of Educational and Psychological Consultation*.

- Reddy, L.A., Newman, E., & Verdesco, A. (2014). Students with ADHD: Using self-regulated learning interventions and student and teacher formative assessment. In T. Cleary (Ed). *Self-Regulated learning interventions with at-risk populations: Academic, mental health, and contextual considerations*. American Psychological Association Press: Washington, DC.
- Reddy, L.A., Newman, E. & Verdesco, A. (2014). Attention-Deficit Hyperactivity Disorders: Using cognitive behavioral interventions and teacher formative assessment in schools. In R. Flanagan & C. Allen (Eds). Cognitive behavioral therapies for children in schools. Springer Publishing.
- Rockoff, J.E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2) 247-252.
- Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education*, 8, 5–35.
- Sanders, W.L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329-339.
- Scheeler, M. C., Bruno, K., Grubb, E., & Seavey, T. L. (2009). Generalizing teaching techniques from University to K-12 classrooms: Teaching preservice teachers to use what they learn. *Journal of Behavioral Education*, 18, 189-210.
- Schloss, P.J. & Smith, M.A. (1998). *Applied behavior analysis in the classroom*. Boston, MA: Allyn and Bacon.
- Rosenfield, S. (2008). Best practice in instructional consultation and instructional consultation teams. *Best practices in School Psychology V*, 1645-1660.

- Rosenshine, B., & Stevens, R. (1986). Teaching functions. In M. C. Witrock (Ed.), *Handbook of research on teaching* (pp. 376–391) (3rd ed.). New York, NY: Macmillan.
- Scheerens, J. (1992). *Effective Schooling: Research, theory, and practice*. London: Cassell.
- Scott, S. E., Cortina, K. S., & Carlisle, J. F. (2012). Understanding coach-based professional development in reading first: How do coaches spend their time and how do teachers perceive coaches' work? *Literacy Research and Instruction, 51*(1), 68-85. doi: <http://dx.doi.org/10.1080/19388071.2011.569845>
- Shapiro, E. S. (2006). Are we solving the big problems? *School Psychology Review, 35*(2), 260-265.
- Shernoff, E. S., Marinez-Lora, A., Frazier, S. L., Jakobsons, L. J., Atkins, M. S., & Bonner, D. (2011a). Teachers Supporting Teachers in Urban Schools: What iterative research designs can teach us. *School Psychology Review, 40*, 465-485.
- Shernoff, E. S., Mehta, T. G., Atkins, M. S., Torf, R., & Spencer, J. (2011b). A qualitative study of the sources and impact of stress among urban teachers. *School Mental Health, 3*(2), 59-69. doi:10.1007/s12310-011-9051-z
- Shernoff, E.S., Jakobsons, L., Lakind, D., Frazier, S.L. Marinez-Lora, A., Hamre, B.K., Patel, D., Atkins, M.S., Parker Katz, M., Neal, J. & Smylie, M. (under review). Enhancing urban early career teachers' effectiveness and connectedness: A mixed method study. Manuscript submitted for publication to *American Education Research Journal*.
- Sparks, G. M. (1983). Synthesis of research on staff development for effective teaching. *Educational Leadership, 41* (3), 65–72



- Stage, S.A. & Quiroz, D.R. (1997). A meta-analysis of interventions to decrease disruptive classroom behavior in public education settings. *School Psychology Review, 26*, 333-368.
- Sutherland, K. S., & Wehby, J. H. (2001). Exploring the relationship between increased opportunities to respond to academic requests and the academic and behavioral outcomes of students with EBD: A review. *Remedial and Special Education, 22*, 113–121.
- Sutherland, K. S., Adler, N., & Gunter, P. L. (2003). The effect of varying rates of opportunities to respond to academic requests on the classroom behavior of students with EBD. *Journal of Emotional and Behavioral Disorders, 11*, 239–248.
- Sutherland, K. S., Wehby, J. H., & Yoder, P. J. (2002). Examination of the relationship between teacher praise and opportunities for students with EBD to respond to academic requests. *Journal of Emotional and Behavioral Disorders, 10*(1), 5-13.
- Toch, T., & Rothman, R. (2008). Rush to judgment: Teacher evaluation in public education. *Education Sector*, Washington, D.C.
- Tomlinson, C.A., & Edison, C.C. (2003). *Differentiation in practice: A resource guide for differentiating curriculum, grades K-5*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Tomlinson, C.A., & McTighe, C.A. (2006). *Integrating differentiated instruction (DI) and understanding by design (UbD)*. Association for Supervision and Curriculum Development.
- Training and Development Agency for Schools (2007). *Professional standards for teachers*. London: The Training and Development Agency for Teachers. Retrieved March 9, 2013, from [www.tda.gov.uk/upload/resources/pdf/s/standards\\_a4.pdf](http://www.tda.gov.uk/upload/resources/pdf/s/standards_a4.pdf)

- Walberg, H. J. (1986). Synthesis of research on teaching in Wittrock M (ed.) *Handbook of Research on Teaching*, (3<sup>rd</sup> ed.). New York, NY: Macmillan.
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2003). *Antisocial behavior in school: Evidence-based practices (2nd ed)*. Belmont, CA: Wadsworth/Thomson Learning.
- Wang, M. C. (1991). Productive teaching and instruction: assessing the knowledge base. *Phi Delta Kappan*, 71, 470-478.
- Wang, M. C., Haertel, G. D., & Walhberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research*, 63, 249 – 294.
- Ward, B. (1985). Teacher development: The challenge of the future. In S.. Hord, S. O' Neal & M. Smith (Eds.), *Beyond the looking glass* (pp. 283-312). Austin: The Research and Development Center for Teacher Education, The University of Texas.
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Retrieved June 2, 2009, from <http://widgeteffect.org/>
- Wittrock, M.C. (1986). *Handbook of research on teaching* (3rd ed.). New York: Macmillan.
- White, M. A. (1975). Natural rates of teacher approval and disapproval in the classroom. *Journal of Applied Behavioral Analysis*, 8, 367–372.
- Yoon, K. S., Duncan, T., Lee., s. w., Scarloss, B., & Shapley, K. L. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No. 033). Washington, DC: U.S. Department of Education, Institute of Education Science, National Center for Education and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from [http://ies.edu/ncee/edlabs/regions/southwest/pdf/refl\\_2007033.pdf](http://ies.edu/ncee/edlabs/regions/southwest/pdf/refl_2007033.pdf)

Ysseldyke, J. E., Burns, M., Dawson, M., Kelley, B., Morrison, D., Ortiz, S., et al. (2006).

*School psychology: A blueprint for training and practice III*. Bethesda, MD: National Association of School Psychologists.

Ysseldyke, J. (2005). Assessment and decision making for students with learning disabilities:

What if this is as good as it gets? *Learning Disability Quarterly*, 125-128.

Table 1 Definitions of the Three Part CSS Assessment

Part 1 Strategy Counts <sup>a</sup>	Definitions
Concept Summaries	A teacher summarizes or highlights key concepts or facts taught during the lesson. Summarization statements are typically brief and clear. This teaching strategy helps students organize and recall material taught.
Academic Response Opportunities	A teacher creates opportunities for students to share their understanding of the lesson content with the teacher or class. These opportunities can be verbal or nonverbal responses (e.g., explain answers, repeat key points, brainstorm ideas, and show answers on the board).
Academic Praise Statements	A teacher gives a verbal or nonverbal statement or gesture to provide feedback for appropriate academic performance.
Academic Corrective Feedback	A teacher gives a verbal or nonverbal statement or gesture to provide feedback for incorrect academic performance.
Clear One to Two Step Directives	A teacher gives a verbal instruction that specifically directs a behavior to occur immediately. These directives are clear and they provide specific instructions to students to perform a behavior. They are declarative statements (not questions), describe the desired behavior, and include no more than two steps.
Vague Directives	A teacher gives a verbal instruction that is unclear when directing a behavior to occur immediately. These directives are vague, may be issued as questions, and often include unnecessary verbalizations or more than two steps.
Behavioral Praise Statements	A teacher gives a verbal or nonverbal statement or gesture to provide feedback for appropriate behavior.
Behavioral Corrective Feedback	A teacher gives a verbal or nonverbal statement or gesture to provide feedback for inappropriate behavior.
Total	The sum of the frequency of the eight teacher behaviors.
Part 2: Instructional Strategies Scales	Definitions
Total Scale	The Total Instructional Strategies scale reflects the overall use of Instructional Methods and Academic Monitoring/Feedback.
Instructional Methods composite scale	How classroom instruction occurs. Measures teachers' use of teacher directed student directed methods, or differentiated instruction. This includes how a teacher incorporates active learning techniques such as hands on learning and collaborative learning in the presentation of lessons as well as how a teacher delivers academic content to students.
Adaptive Instruction Subscale	Strategies teachers use to respond to their students' learning needs while teaching. These practices reflect teacher flexibility and responsiveness to students' needs, as well as methods of differentiated instruction.
Student-Directed Instruction subscale	Strategies teachers use to actively engage students in the learning process. These practices encompass constructivist and hands-on instructional techniques, linking lesson content to prior learning, personal experiences, and cooperative learning.
Direct Instruction Subscale	Strategies teachers use to deliver academic content or convey information to students. These practices include direct instruction techniques, modeling, identifying, and summarizing.
Academic Monitoring/Feedback composite scale	How teachers monitor students' understanding of the material and provide feedback on their understanding. These strategies assess students' thinking and encourage students to examine their own thought processes. Teachers guide students understanding by encouraging students, affirming appropriate application of the material, and correcting misperceptions.

Promotes Students' Thinking Subscale	Strategies teachers use to activate students' thinking about the lesson material. These practices assess teachers' efforts to get their students to think about their thinking process (i.e., open-ended, what, how, and why).
Academic Performance Feedback Subscale	Strategies teachers use to provide specific feedback to their students' on their understanding of the material. These practices assess teachers' efforts to explain what is correct or incorrect with student academic performance.
<hr/>	
Part 2: Behavioral Management Strategies Scales	Definitions
<hr/>	
Total Scale	The Total Behavioral Management Strategies scale reflects the overall use of Proactive Methods and Behavior Feedback.
Preventative Methods composite scale	Strategies teachers use to promote positive behaviors in the classroom and reduce the likelihood of negative behaviors. These strategies include prompts, routines, reviewing rules, and presenting instructions or requests in a clear manner.
Proactive Methods Subscale	Verbal and nonverbal strategies teachers use to prevent student disengagement, and problem behaviors from occurring in the classroom. These practices assess how teachers create a positive classroom environment.
Directives subscale	Strategies teachers use for issuing directions or instructions to students and behavioral expectations in the classroom.
Behavioral Feedback composite scale	How teachers respond to students appropriate and inappropriate behaviors. This includes the usage of praise to encourage positive behaviors and corrective feedback to redirect negative behaviors.
Praise subscale	Verbal and nonverbal strategies teachers use to positively reinforce specific appropriate behaviors in the classroom. These practices assess how teachers respond to positive behavior in the classroom.
Corrective Feedback Subscale	Verbal and nonverbal strategies teachers use to correct students' inappropriate behavior. These practices assess how teachers respond to negative behavior in the classroom.
<hr/>	
Part 3 Classroom Checklist Items	
<hr/>	
1. Different methods/mediums of instruction are present in the classroom (e.g., blackboard, overhead projector, smart board, student clickers).	
2. Learning aids are present in the classroom (e.g., number chart, vocabulary list, critical thinking questions).	
3. Learning materials are present in the classroom (e.g., pencils, rulers, construction paper).	
4. Learning materials and areas in the classroom are labeled.	
5. A procedure or routine exists for students to organize their desks, backpacks, or learning materials.	
6. Classroom (e.g., floors, walls, table) is clean and uncluttered.	
7. Tables/desks are arranged for students to easily view and participate in the lesson.	
8. Classroom lesson or activity schedules are clearly posted.	
9. Assignments (e.g., homework, readings, tests) are clearly posted.	
10. Student work, artwork, and accomplishments are displayed in the classroom.	
11. Methods for tracking student academic and/or behavioral progress (e.g., homework-tracking chart, rule-following chart, sticker/star chart) are present.	
12. Classroom-wide reward system is present (e.g., ticket bin for a pizza party).	
13. Classroom rules are posted.	
14. Classroom rules specify positive behaviors that students "should do" rather than "not do".	
<hr/>	

Note <sup>a</sup> – CSS-Teacher Form does not include Part 1 Strategy Counts – includes Part 2 and 3. Table 1 is published Reddy & Dudek (2014).

Table 2. Using Teacher Formative Assessments: Implications for Practice

RtI Components	Teacher Support	Organizational Support
Tier 1 Formative Assessment	Routine progress monitoring to promote continuous learning opportunities	Inform district-wide professional development (PD) needs and goals
	Promote reflection, dialogue, and ongoing PD	Link individual performance with school-wide and district-level performance goals
	Increase use of general evidence-based instructional practices	Inform school-wide policies to prevent behavior problems
	Increase use of best practices for teacher assessment linked to PD	Inform school-wide policies to enhance instructional goals
		Promote organizational culture that supports ongoing learning and continued PD
Tier 2 Formative Assessment	Enhance relationships and supports for at risk teachers	Promote workforce development for teachers at risk
	Mentor at risk teachers via indigenous school supports	Identify organizational goals to support teachers at risk for struggling with behavior management or effective instruction
	Increase use of targeted evidence-based instructional practices	Promote sustainability of teacher support and PD
	Increase use of targeted practices for teacher assessment linked to PD	
Tier 3 Formative Assessment	Provide intensive individualized PD for identified teachers	Enhance communication between leadership teams and teachers regarding PD goals and benchmarks
	Provide intensive individualized teacher assessment linked to PD for identified teachers	Detail action plans to enhance school-wide knowledge, skill, and learning  Facilitate the link between classroom/individual learning and school-wide/organizational growth