Measuring Explicit Instruction using the *Classroom Observations of Student–Teacher Interactions* (*COSTI*)

Barbara Gunn, Keith Smolkowski, Lisa A. Strycker, and Caroline Dennis

Oregon Research Institute, Eugene, OR, USA

Barbara Gunn: barbarag@ori.org; Keith Smolkowski: keiths@ori.org

Lisa A. Strycker lisas@ori.org; Caroline Dennis: cdennis@ori.org

Address correspondence to: Barbara Gunn, Oregon Research Institute, 1776 Millrace Drive, Eugene, OR, USA 97403. E-mail: barbarag@ori.org ORCID: 0000-0002-1229-7513

Date of acceptance by *Perspectives on Behavior Science:* April 27, 2021

Journal peer review: https://www.springer.com/journal/40614/submission-guidelines

Declarations

Funding: The U.S. Department of Education Institute of Education Sciences, through Grant R305A150325, funded this study. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the U.S. Department of Education.

Conflict of Interest: The authors declare that they have no conflicts of interest.

Availability of data and material: Not applicable.

Code availability: Not applicable.

Author Contributions: All authors contributed to this paper. Barbara Gunn drafted the original manuscript. Barbara Gunn, Caroline Dennis, and Keith Smolkowski conducted the literature review. Keith Smolkowski performed data analyses for the studies the authors conducted and cited. The manuscript was revised by Barbara Gunn, with editing and review by Lisa A. Strycker, Caroline Dennis, and Keith Smolkowski.

Acknowledgments: The authors thank Kathy Jungjohann for her guidance on observing and coaching teachers, which led to the development of the *COSTI* instrument, and Susan Long for her expertise and assistance in preparing this manuscript.

Abstract

This paper describes *the development and technical adequacy of the Classroom Observations of Student Teacher Interactions (COSTI) instrument*, a tool for measuring the frequency and rate of explicit instructional interactions, such as those used in Direct Instruction curricula, for teaching children basic reading and math skills. *The COSTI was originally developed* to provide teachers with coaching feedback to improve their explicit reading instructional practices*, and has been shown in multiple studies to be a reliable and valid predictor of student gains in beginning reading and math skills.* This paper discusses potential uses of the instrument for training and coaching across curricula with varying instructional design features, and lays out a future research agenda to further improve the *COSTI* and related observation tools for studying explicit instructional practices and their contribution to student learning.

Keywords: basic skills; explicit instruction; Direct Instruction; classroom observation

Measuring Explicit Instruction using the *Classroom Observations of Student-Teacher Interactions* (*COSTI*)

Teaching and learning basic skills is a focal point of early elementary school, and a significant part of daily classroom instruction. The mastery of basic skills lays a critical foundation for future learning by giving students the tools they will need to learn and apply to more-advanced knowledge and skills throughout their academic and working careers. The critical importance of gaining an early and deep understanding of core concepts and procedures is reflected in the Common Core State Standards for English language arts and mathematics for kindergarten and the early elementary grades (CCSS; National Governors Association Center for Best Practices, 2010) and in the recommendations provided by the What Works Clearinghouse Practice Guides (What Works Clearinghouse, 2020).

To teach basic skills effectively, teachers must engage in frequent, explicit instructional interactions with their students. Frequent interactions give students repeated exposure to and practice with new content and skills, and explicit interactions use clear, concise, and consistent language to help students practice correctly. Principles of learning and retention, with evidence from the fields of education, psychology, and cognitive neuroscience, recognize the essential role of teachers and their instructional interactions with students in developing foundational skills (Archer & Hughes, 2011; Carver & Klahr, 2001; Clark et al., 2012; Deans for Impact, 2019; Shaywitz et al., 2008; Taylor et al., 2010; Twyman & Heward, 2016). Research findings show that these student–teacher interactions are crucial in teaching new content and skills to novice learners, that is, students who have little or no prior knowledge of content as well as very young learners who have limited experience in learning how to learn (Adams, 1990; Frye et al., 2013; Logan, 1997; Perfetti & Bolger, 2004; Rosenshine, 2012).

To provide teachers with feedback on their basic skills instruction, and to learn more about effective instructional practices, it is important to reliably document aspects of instructional delivery that have been shown to help students learn new content and skills. This paper describes the *Classroom Observations of Student Teacher Interactions (COSTI) instrument,* a tool for measuring the frequency and rate of explicit instructional interactions, such as those used in Direct Instruction curricula, for teaching children basic skills in the early elementary grades. We describe t*he development and research basis of the COSTI to assess teachers' beginning reading instructional practices, including the technical adequacy of the instrument.* We discuss implications, including potential uses of the instrument for providing teachers and instructional assistants with coaching feedback on their basic skills instruction, and lay out a future research agenda with the *COSTI* and related observation tools for studying explicit instructional practices and their contribution to student learning.

Teaching Basic Skills: The Role of Explicit Instruction

In this section, we define and provide a brief discussion of explicit instruction as a combination of teaching behaviors designed to promote student learning. We describe the origins of explicit instruction in the Direct Instruction Model. We discuss the distinction between the common use of the term explicit instruction, as it applies to commercially published curricula, and the explicit instruction found in the Direct Instruction programs developed by Engelmann and colleagues.

Explicit instruction is a general term used to describe the combination of direct and unambiguous teaching behaviors needed to support successful student learning with any curriculum. It describes instructional delivery, or the "how" of effective instruction (Archer & Hughes, 2011; Brophy & Good, 1986; Clark et al., 2012; Gersten et al., 2000; Hughes et al., 2017; Klahr & Nigam, 2004; Rosenshine & Stevens, 1986). Often-cited examples of explicit instruction in this context include teacher modeling or think-alouds, scaffolded instruction, and frequent opportunities for practice with feedback. Explicit instruction contrasts with less-explicit approaches that rely on partial teacher guidance or student discovery for acquiring new knowledge and skills. Intervention research supports the effectiveness of explicit instructional approaches for teaching a range of literacy skills (Herrera et al., 2016; National Reading Panel [NRP], 2000) and math skills (Frye et al., 2013), and underscores the important role that explicit instructional delivery plays in teaching young learners basic skills (Clark et al., 2012; Reutzel et al., 2014; Stockard et al., 2018).

In their paper on the historic evolution of explicit instruction, Hughes et al. (2017) place its origins in the Direct Instruction Model developed by Siegfried Engelmann in the 1960s with colleagues such as Carl Bereiter and Wesley Becker. Engelmann and associates based Direct Instruction on three logical analyses: (a) an analysis of communications that transmit knowledge effectively ("faultlessly," or with a single interpretation) to students and allow the generalization of concepts through specific examples; (b) an analysis of knowledge systems applied to the classification of cognitive operations, from simple discriminations to complex functions, and the link between the classifications and instructional procedures that encourage efficient communication; and (c) an analysis of behavior to identify universal principles about how the environment influences behavior (Baer et al., 1987; Engelmann & Carnine, 1982, 1991). Boulton (2019) summarized the foundations of Direct Instruction this way:

> Rooted in behaviourism, almost everything is predicated on just two principles: the first is that the human mind possesses the ability to recognize a particular example as an instance of a more general, abstract concept; and the second is that the mind develops the

abstract generalisation after exposure to more than one example. (p. 24)

Relying on this basis for designing and delivering instruction, each Direct Instruction program is a comprehensive system of instruction that integrates explicit teaching techniques into the instructional design. Lessons encapsulate the central roles of mastery learning and the correct placement in the curriculum to ensure effective teaching and learning.

Within the Direct Instruction Model, explicit instructional strategies are but one element of an overall system designed to promote student learning. Scripted lessons guide teaching with an organized and carefully sequenced progression though curriculum content, thereby integrating curriculum design and effective teaching practices (Stein et al., 1998). Thus, "Direct Instruction includes both curriculum (what to teach) and instruction (how to teach), whereas explicit instruction focuses primarily on how to teach" (Hughes et al., 2017, p. 144). Direct Instruction and explicit instruction are similar in that explicit teaching strategies can be applied to any curricular content; Direct Instruction stands alone, however, as a comprehensive system of instruction intended to maximize generalization.

Classroom Observations of Student–Teacher Interactions

In this section, we describe the origins of the *COSTI* as a coaching tool, list four explicit student–teacher interactions documented with the instrument, each based on Direct Instruction principles and curriculum design standards, and provide supporting theory and research evidence (Aylward et al, 2003; Barbash, 2012; Engelmann & Carnine, 1991; Goswami, 2004; Shaywitz et al., 2008).

We based the original version of the *COSTI* on a coding system used by one of the authors to observe first-grade teachers using *Reading Mastery*, and to give them coaching feedback on the frequency and sequencing of four observable and measurable aspects of explicit instruction that commonly occur during beginning reading instruction: teacher demonstrations, student independent practice, student errors, and teacher corrective feedback. The systematic instruction in the scripted lessons helped teachers keep the language of their instruction clear, and provided them with programmed opportunities to monitor their students' understanding and adjust instruction to meet their needs. Data collected bimonthly from classroom observations were shared with teachers so that they could see how frequently they used explicit instructional strategies. With this feedback, teachers could reflect on their instructional patterns, and make changes in their instructional interactions with students to promote learning. For instance, the data might show that a teacher followed the script, but gave students few chances to practice independently—leading the teacher to decide to offer more practice opportunities, and thereby decrease disruptive behavior, increase task engagement, and accelerate student learning. Or the data could reveal a pattern of students making errors without the teacher giving corrective feedback—helping the teacher understand why some students were struggling. Teachers found the objective data more useful than general impressions because it helped them identify specific ways to promote student learning.

Teacher Demonstrations

A key feature of explicit instruction is teacher demonstrations (Archer & Hughes, 2011). In demonstrations, teachers give students new information or show students how to use a new skill. Demonstrations are especially important when students are learning a new skill, need additional practice, or make errors. To demonstrate, a teacher might say: "Watch and listen to me sound out the word <u>sun</u> … /sss/, /uuu/, /nnn/" or "The letter *s* makes the sound /sss/." Key phrases are "my turn," "listen to me," and "watch me." Teacher demonstration of a new skill allows students to practice correctly, which is faster and more effective than trial-and-error or discovery learning (Meltzoff et al., 2009; Smith, 1979).

Teacher demonstrations have been shown to help students acquire initial reading and math skills. Students of teachers who demonstrate new skills learn more than students of teachers who use less-explicit methods such as discovery learning (e.g., Clark et al., 2012; Foorman et al., 1998; Gunn et al., 201; Klahr & Nigam, 2004; Kulik et al., 1990).

Student Independent Practice

When teachers provide independent practice, students practice a new skill or strategy on their own, without any prompting or help. Students can practice individually or as a group, with choral responding. This method is distinct from guided practice, when teachers practice *with* their students. Guided practice is a valuable part of skills instruction initially, when students need scaffolding and support, but as students learn new content and skills, they need to practice independently so that teachers can monitor their understanding.

Conceptually, teacher-directed independent practice is similar to opportunities to respond (OTRs) and engagement in academic responding (Greenwood et al., 2002; Rosenshine, 1995). These methods combine student classroom behaviors such as reading aloud, answering teacher questions, and doing assigned tasks. However, teacher-directed independent practice is a more useful descriptor because it identifies the teacher as the person directly controlling instruction and practice, and clearly delineates what students are doing in measurable, observable terms. Independent practice has been associated with improved acquisition of new skills, a defining feature of proficiency in basic academic learning (Ericsson et al., 2007; Heward & Wood, 2015; Logan, 1997; Meltzoff et al., 2009; Sutherland et al., 2003). Independent practice promotes fluency in early reading (Cuticelli et al., 2016; Ellis & Worthington, 1994; Gunn et al., 2011; MacSuga-Gage & Simonsen, 2015; Vadasy et al., 2005) and math (Clarke et al., 2011; Gersten et al., 2009). According to Posner et al. (1997), “The idea that practice can automate a skill has been with us since the inception of psychology” (p. 267). In beginning reading instruction, a high rate of independent practice keeps students focused on the lesson (Carnine, 1976) and accelerates their development or automaticity with skills needed for skilled reading (Samuels & Flor, 1997; Share, 2008). In beginning math instruction, frequent practice of basic skills builds fluency, thereby enabling students to access basic knowledge quickly and automatically, freeing up mental energy to focus on the more challenging aspects of a complex problem (Morgan et al., 2010).

Student Errors

During independent practice, students err when they give either an incorrect response or no response. On the *COSTI*, both are coded as student errors. Students make errors for many reasons. They may not have acquired the skill at all, or they may have learned the skill incorrectly, or they may not understand the task. Although student errors do not directly measure a teacher’s instructional interaction or approach, we hypothesize that student error data are useful in measuring the effectiveness of instructional practices. The *COSTI* documents group and individual student errors to: (a) measure effectiveness of teachers’ instructional practices, (b) better understand the role of corrective feedback, and (c) quantify relations between student learning and student errors.

Teacher Corrective Feedback

When students practice independently, teachers can monitor their understanding, offer prompt feedback on errors, and give additional practice and re-teaching as needed. With good corrective feedback, teachers focus on the correct, rather than the incorrect, response (Kameenui & Simmons, 1990; Hattie & Timperley, 2007). Teachers give students the correct information or strategy immediately, and then have students practice correctly, thereby reinforcing correct information rather than having students repeat errors. Good corrective feedback matches the type of error the student makes, is provided in a positive and encouraging manner, and ends with student practice of the correct answer (Archer & Hughes, 2011). Research shows that these feedback techniques improve word recognition accuracy with new readers (e.g., Barbetta, Heward, Bradley, & Miller, 1994; Gardner, 1998; McCoy & Pany, 1988). Studies comparing early math core curricula instruction also support the positive impact of systematic and explicit feedback on student achievement (Clements et al., 2013).

The Observation Instrument

After our experience using the *COSTI* as a coaching tool, we field-tested it in kindergarten and first-grade classrooms. Our aim was to refine the instrument and the observer training so that the *COSTI* could be feasibly and reliably used by teachers, coaches, and researchers. Figure 1 shows a sample of the original paper and pencil tally sheet coaches use to summarize and share the data with the teachers; Figure 2 shows a sample of a coded *COSTI* form used for research purposes.

Using the original paper-and-pencil instrument, we developed a standardized coding form. The form has a cover page on which observers record general classroom information: grouping (whole class, small group), number of students, curriculum used, and lesson stop and start times. The rest of the instrument is divided into half-page sections. Each page of the coding form has a series of columns of bubbles for observers to record, in sequence, each instance of (a) explicit teacher demonstration, (b) student independent practice, (c) student errors, and (d) teacher corrective feedback.by filling in a bubble on the row allocated for the interaction. Observers use a new section of the form for each activity during a lesson. Each instance of a specific teacher-student interaction is defined and scored as one instructional step within an activity. For example "My turn, the sounds in **nest** are /nnn/ /eee/ /sss/ /t/ **nest**," is coded as one teacher demonstration. "Your turn, what are the sounds in **nest**? /nnn/ /eee/ /sss/ /t/," is coded as one student independent practice. At the end of the observation, rates of each type of interaction can then be calculated by dividing the total interactions by the total minutes. We initially used the refined instrument for observing beginning reading instruction. The serial method of coding provided a total count of each interaction and the sequence in which the interactions occurred. We used these data to identify patterns of instructional interactions, such as whether a teacher demonstration was followed by student independent practice, or whether the practice was an error—and, if so, whether that was followed by a correction. The coded example provided in Figure 1, shows how an observation is scored in rates per minute and then summarized for teachers to objectively see their instructional interactions with their students, and reflect on their teaching practices.

In research studies, using the form depicted in Figure 2, observers began coding at the beginning of a lesson. Across the length of the lesson (typically 30 minutes), we computed the rates per minute for each explicit instructional interaction, which removed the influence of observation duration from the analyses. Counting these observable, measurable interactions with clear decision rules allowed observers to document teachers' literacy instructional practices objectively, with fewer judgment calls that might affect the reliability and interpretability of results.

Figure 3 provides operational definitions for each of the four instructional interactions that we observed, and the decision rules for what to count as examples of each. As noted, we focused on these aspects of explicit instruction because they are based on Direct Instruction principles and curriculum design standards and used for teaching children basic reading and math skills. For other observational research, definitions may need to be adjusted or modified for specific content areas, depending on the research questions and specific purpose of the observation. Categories could be expanded, such as splitting independent practice opportunities into individual versus group practice, or by adding a separate category such as guided practice. Codes could also be added as needed to document the instructional focus (e.g., phonological awareness, alphabetic understanding, decoding), or classroom contexts that might affect instruction (e.g., level of student engagement, lesson interruptions). These additional categories and codes would not be part of the COSTI per se, but may be useful for specific research questions.

Gathering Evidence for Best Teaching Practices

Although the *COSTI* was initially used for coaching with Direct Instruction programs, it has also been employed as a research tool in different contexts to determine the extent to which teachers use explicit instructional approaches to teach basic skills with other curricula. In earlier research on beginning reading instruction, the lead author (Gunn) trained and coached instructional assistants (IAs) to provide supplemental instruction with *Reading Mastery* and *Corrective Reading* to small groups of early elementary students at risk for reading difficulties (Gunn et al., 2005). The *COSTI* was used to document the frequency of student–teacher interactions and to provide IAs with data to improve these interactions. At the end of the intervention, students who received the supplemental instruction performed better than their matched controls on measures of entry-level reading skills (i.e., letter–word identification and word attack), oral reading fluency, vocabulary, and comprehension. The benefits of the instruction persisted 2 years after the program ended, with students in the supplemental instruction condition continuing to show greater growth in oral reading fluency. Results demonstrated the efficacy of supplemental instruction using a carefully designed curriculum, combined with high rates of student–teacher interactions, to help students at risk for reading failure develop fluent word-recognition skills.

In a subsequent study, the *COSTI* was validated with a larger sample of teachers and classrooms in an evaluation of the *Read Well Kindergarten* reading program, a beginning-reading curriculum based on direct instruction principles. We compared the efficacy of *Read Well Kindergarten* (Sprick et al., 2004) to typical kindergarten literacy instruction, and investigated whether the benefits of *Read Well Kindergarten* were maintained into first grade (Gunn et al., 2011). Twenty-six schools were randomized to treatment or control conditions, and nearly 1,500 students were tested at the beginning and end of kindergarten and at the beginning and end of first grade. We observed all teachers with the *COSTI* three times during the school year. We found an interaction between study condition and opportunities for independent student practice (measured by the *COSTI*), although neither the curriculum nor the rate of practice *alone* was sufficient to produce significant gains in students' performance on standardized measures.

For this study, we further refined and developed the observer training to ensure a high rate of interobserver agreement. Prior to the observations, observers were given an overview of the system, an explanation of the codes, and procedures for using the observation codebook. The trainer and the observers practiced coding and debriefing as a group with video clips. This was followed by practice coding with the trainer individually in kindergarten classrooms that were not in the study.

Technical Adequacy of the *COSTI*

Research results demonstrated that the *COSTI* could capture and reliably measure student–teacher instructional interactions. Rates of independent opportunities for practice and student errors were stable for teachers across the school year and produced reliable classroom-level means; also, independent practice opportunities were positively associated with literacy outcomes.

Interobserver Reliability

We examined *COSTI* reliability in two ways. Observers maintained an 80% or higher rate of interobserver agreement with an observation trainer throughout the study, measured by periodic retesting. Interobserver agreement does not express true reliability, however, due to its dependence on the base rate of the behavior coded (Mitchell, 1979). Intraclass correlation coefficients (ICCs) correct for chance agreement (Fleiss & Cohen, 1973). ICC values suggested a high level of reliability for independent practice opportunities (ICC = .86) and adequate reliability for student errors (ICC = .72). Reliabilities for corrections and demonstrations were acceptable (ICCs = .68 and .61, respectively). Because teachers sometimes used demonstrations to correct student errors, Smolkowski and Gunn, 2012 suggested that observer training could have been more clearly differentiated between instances when demonstrations were used to correct errors and when demonstrations introduced new material.

Teacher Stability and Reliability

Generalization from the *COSTI* and the use of observed means in research and practice assumes that student–teacher interactions remain stable across time within teachers. If teachers' behavior and their interactions vary markedly from day to day or over time, the results would have limited value for prediction and for coaching teachers to improve their instructional practices. Shoukri and colleagues (2004) showed that, as with reliability estimates, the reliability of an observed (aggregate) mean of multiple observations for a teacher depends on the number of observations collected on that teacher. Smolkowski and Gunn, 2012, tested teacher stability of the rates per minute of each interaction and demonstrated that independent practice opportunities were generalizable over time with two or three observation per year (ICC = .71). Student errors were stable enough to recommend four observations per year (ICC = .32). Teacher corrections and demonstrations varied considerably from one occasion to the next (ICCs = .21 and .10, respectively), which would require six or 14 observations, respectively, to obtain reliable means. Low interrater reliability on these latter two interactions, however, may have depressed their stability estimates.

Shrout and Fleiss (1979) suggest that ICCs represent the average correlation between any two randomly chosen pairs of observations, which indicates considerable consistency in the rates of independent practice opportunities over time. Smolkowski and Gunn, 2012 noted that observations were spaced several months apart and that students frequently changed instructional groups, which means that, on this particular instructional skill, “teachers who tended to give more practice opportunities tended to do so consistently” (p. 324).

The reliabilities of observed (aggregate) classroom means are a function of the ICCs and the number of observations collected (Shoukri et al., 2004). With three observation per year, Smolkowski and Gunn, (2012) reported reliabilities of .81 for independent practice, .59 for student errors, and .45 for teacher corrections. Reliabilities were lower for teacher demonstrations (.24). The reliabilities of the classroom means showed that the *COSTI* can provide reliable information on the rate of students’ independent practice opportunities as well as reasonably reliable data on student errors and corrections. Reliability improves with more observations per year. Reliability also may improve within environments having more stable levels of instructional interactions or by observing only when teachers engage in specific activities, such as the introduction of new skills, rather than differing activities over time (e.g., introducing new skills, review of past material, seatwork).

Predictive Validity

The *COSTI* demonstrated strong predictive validity. (Smolkowski & Gunn, 2012) reported that the rate of independent practice opportunities accounted for more than 35% of the classroom-to-classroom variability in gains on measures of decodable- and sight-word reading, 25% of the variability in phonological processing, about 20% in the classroom variance of letter names and sounds, and 17% of the classroom variability in oral reading fluency. The authors noted that independent practice opportunities and classroom-level student reading in first grade were related to skills taught in kindergarten.

Use of COSTI across Studies

To date, the *COSTI* has been used in studies of beginning reading and math instruction (e.g., Clarke et al., 2015; 2017; Doabler et al, 2016; Gunn et al., 2011; Fien et al., 2021; Nelson-Walker et al., 2013; Smith et al., 2016; Smolkowski & Gunn, 2012). Findings from these studies support the reliability, generalizability, and predictive validity of the *COSTI* for observing beginning reading and math instruction. See Smolkowski (nd) for additional references for papers that collect and use the COSTI for intervention evaluation research or that otherwise collect and report on the *COSTI*.

Discussion and Future Research Directions

The *COSTI* was designed to serve two purposes: to give teachers feedback on their instruction in basic skills and to provide researchers with data on the frequency, or rates, of specific student–teacher interactions that impact student learning. The original version of the *COSTI*, based on the Direct Instruction Model (Englemann & Carnine, 1991), was used by one of the authors to observe first-grade teachers implementing the *Reading Mastery* curriculum, and to provide them with coaching on the frequency and sequencing of their use of demonstrations, student independent practice, monitoring errors, and giving corrective feedback. The scripted curriculum gave teachers a practical framework for presenting new skills and strategies. The teachers' instructional delivery was equally important. Data collected from the observations were used to show teachers the patterns in their instructional delivery while implementing the curriculum. Teachers were encouraged to think about how those patterns might affect their students' learning, and to make changes to improve their instruction.

The *COSTI* offers a practical, objective way to give teachers and instructional assistants feedback on their instructional delivery. Although it was beyond the scope of our work, the instrument has the potential to be used by teachers and specialists within a school system to observe and coach each other in a peer coaching format. It also has the potential to be used at the district level as part of in-service training on effective implementation of basic skills curricula. Data generated from the *COSTI* also provide research evidence that training and coaching on explicit instructional interactions can support the development of basic skills in the early elementary grades. The *COSTI* yields information about instructional interactions not captured by other observation instruments. Future intervention studies may use the *COSTI* to further clarify the role of student–teacher interactions in basic skills instruction, and to help teachers improve the quality of program implementation.

Additional research is needed to further develop and improve the *COSTI* and related observation tools for measuring the use of explicit instructional approaches for teaching and learning basic skills. The use of such instruments helps researchers form a more complete picture of how curriculum design and content influence teachers' instructional delivery. Although explicit teaching strategies can be applied to any curriculum regardless of program design, they are more effective when anchored to a carefully designed evidence-based curriculum (Engelmann & Carnine, 1982, 1991; Reutzel et al., 2014; Stein et al., 1998). Teachers are typically required to use district-adopted curricula, with varying degrees of evidence-based content and guidance on optimal instructional delivery. Future observational research should study the frequency and rate of explicit instructional interactions teachers use across curricula to document the influence of the curriculum on teaching and learning. Results from this work could inform teachers, curriculum developers, and teacher educators on curriculum design and optimal instructional practices.

Although the *COSTI* has been used in a range of studies focused on beginning reading and math skill acquisition, more work is needed to establish the predictive validity of the instrument for basic skills acquisition in other content areas, such as learning a second language. Future observational research could also clarify which student–teacher interactions, at which frequency or rate, are most effective, and examine the degree to which classroom variables, such as grouping for instruction and student characteristics, might interact with explicit instructional practices. For example, when students master basic skills and progress to more complex content, the optimal instructional approaches and documentation of those approaches will probably change. We suggest that more specific knowledge on how on how best to teach, as students' knowledge and skills grow, would be valuable to the field. Finally, research pairing the *COSTI* with other instruments, such as observations of the general classroom environment and other measure of teaching quality such as classroom management, could provide a more complete picture of the dynamics of basic skills instruction, and may be especially enlightening in intervention studies aiming to increase the effectiveness of student–teacher interactions

(Raudenbush, 2008).

In conclusion, initial results support the usefulness of the *COSTI* as a tool for both research and practice. Data from this instrument help teachers identify patterns in their instructional delivery, and make changes to improve their instruction and enhance student learning. Data from the *COSTI* also help researchers learn more about how explicit student–teacher interactions help students learn new knowledge and skills. Future research is needed to replicate and extend these findings, with the overarching goal of improving the effectiveness of early elementary classroom instruction.

Fig.1

Example Tally Sheet for Use with the COSTI for Coaching

Note. This example shows how a coach could use the COSTI system to record student-teacher interactions with paper and pencil. The sections with the shaded background show the portion of the form where a coach has recorded interactions in sequence with the symbols shown. The coach counted the symbols for each interaction, determined the total time, and calculated rates. The coach also examined the codes for important patterns as indicated in the note. Figure adapted from (Smolkowski & Gunn, 2012).

Fig.2

Example Coding Form Cover Sheet and COSTI Code Sheet for Research

Note. **Part a** shows the first two pages of a COSTI coding form. The cover page is used to code information about the observation (e.g., ID numbers, date and time) and classroom features. The second and subsequent pages are used to code student-teacher instructional interactions. **Part b** shows an excerpt of the COSTI section, which has two sets of rows. Coders begin with the top set of rows, using one column at a time and continue to the second four rows as needed. The letters printed between the two sets of rows are provided for reference. In this completed example, the coder began at the top left and observed a teacher demonstration (column a), three opportunities for independent practice (b–d), a student error (e), a teacher correction (f), and so on. After completing the first set of rows, the coder moved to the second set. If more space were needed, the coder would have continued into another section of the form for the same activity. Figure adapted from (Smolkowski & Gunn, 2012).

Fig.3

COSTI Observation Manual

Note. This page provides operational definitions for each of the four instructional interactions documented with the *COSTI*, and the decision rules for what to count as examples of each interaction.

References

Adams, M. (1990). *Beginning to read: Thinking and learning about print.* MIT Press.

Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching.* Guilford.

Aylward, E. H., Richards, T. L., Berninger, W. E., Nagy, W. E., Field, K. M., Grimme, A. C., Richards, A. L., Thomson, T. M., & Cramer, S. C. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology, 61*(2), 212–219. https://doi.org/10.1212/01.wnl.0000068363.05974.64

Baer, D. M., Wolf, M. M., & Risley, T.R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 20*(4), 313–327. https://doi.org/10.1901/jaba.1987.20-313

Barbash, S. (2012). *Clear teaching: With direct instruction, Siegfried Engelmann discovered a better way of teaching.* Education Consumers Foundation.

Barbetta, P. M., Heward, W. L., Bradley, D. M., & Miller, A. D. (1994). Effects of immediate and delayed error correction on the acquisition and maintenance of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis, 27*(1), 177–178. https:// dx.doi.org/10.1901/jaba.1994.27-177

Boulton, K. (2019). What was Project Follow Through. In A. Boxer (Ed.), *The researched guide to explicit and Direct Instruction: An evidence-informed guide for teachers* (pp. 15–28). John Catt.

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

Carnine, D. (1976). Effects of two teacher-presentation rates on off-task behavior, answering correctly, and participation. *Journal of Applied Behavior Analysis*, *9*(2), 199–206. https://doi.org/10.1901/jaba.1976.9-199

Carver, S., & Klahr, D. (2001). *Cognition and instruction: Twenty-five years of progress.* Erlbaum.

Clark, R., Kirschner, P. A., & Sweller, J. (2012). Putting students on the path to learning: The case for fully guided instruction. *American Educator, Spring,* 6–11.

Clarke, B., Baker, S., Smolkowski, K., Doabler, C., Strand Cary, M., & Fien, H. (2015). Investigating the efficacy of a core kindergarten mathematics curriculum to improve student mathematics learning outcomes. *Journal of Research on Educational Effectiveness, 8*(3), 303–324. http://dx.doi.org/10.1080/19345747.2014.980021

Clarke, B., Doabler, C. T., Kosty, D., Kurtz Nelson, E., Smolkowski, K., Fien, H., & Turtura, J. (2017). Testing the efficacy of a kindergarten mathematics intervention by small group size. *AERA Open, 3*(2), 1–16. https://dx.doi.org/10.1177/2332858417706899

Clarke, B., Smolkowski, K., Baker, S. K., Fien, H., & Chard, D. (2011). The impact of a comprehensive Tier I core kindergarten program on the achievement of students at risk in mathematics. *The Elementary School Journal*, *111*(4), 561–584. http://dx.doi.org/10.1086/659033

Clements, D. H., Agodini, R., & Harris, B. (2013, September). *Instructional practices and student math achievement: Correlations from a study of math curricula* (NCEE Evaluation Brief ED- 04-CO-0112/0003). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/pubs/20134020/pdf/20134020.pdf

Cuticelli, M., Collier-Meek, M., & Coyne, M. (2016). Increasing the quality of Tier 1 reading instruction: Using performance feedback to increase opportunities to respond during implementation of a core reading program. *Psychology in the Schools*, *53*(1), 89–105. http://dx.doi.org/10.1002/pits.21884

Deans for Impact. (2019). *The Science of early learning.* Deans for Impact. https://deansforimpact.org/wpcontent/uploads/2017/01/The_Science_of_Early_Learning.pdf

Doabler, C. T., Clarke, B., Kosty, D. B., Baker, S. K., Smolkowski, K., & Fien, H. (2016). Effects of a core kindergarten mathematics curriculum on the mathematics achievement of Spanish-speaking English learners. *School Psychology Review, 45*(3), 343–361. http://dx.doi.org/10.17105/SPR45-3.343-361

Ellis, E. S., & Worthington, L. A. (1994). *Research synthesis on effective teaching principles and the design of quality tools for educators.* Technical Report No. 5. National Center to Improve the Tools of Educators, University of Oregon.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications.* Irvington Publishing, Inc.

Engelmann, S., & Carnine, D. (1991). *Theory of instruction: Principles and applications* (rev. ed.). ADI Press.

Ericsson, K., Roring, R., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High Ability Studies*, *18*(1), 3–56. https://dx.doi.org/10.1080/13598130701350593

Fien, H., Nelson, N. J., Smolkowski, K., Kosty, D., Pilger, M., Baker, S. K., Smith, J. L. M. (2021). A conceptual replication study of the Enhanced Core Reading Instruction MTSS-reading model. *Exceptional Children*, 87*(3),* 265–288. https://doi.org/10.1177/0014402920953763

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement, 33*(3), 613–619. https://doi.org/10.1177/001316447303300309

Foorman, B. R., Francis, D. J., Fletcher, J. M., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology, 90*(1), 37–55. https://doi.org/10.1037/0022-0663.90.1.37

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). *Teaching math to young children: A practice guide* (NCEE 2014–4005). Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/PracticeGuides

Gardner, H. A. (1998). *The role of error correction in working with emergent readers* (Report No. CSO13298). School of Education, Dominican College (ERIC Document Reproduction Service No. ED430207). https://files.eric.ed.gov/fulltext/ED430207.pdf

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202–1242. http://dx.doi.org/10.3102/0034654309334431

Gersten, R., Schiller, E. P., & Vaughn, S. (2000). *Contemporary special education research: Synthesis of the knowledge base on critical instructional issues.* Lawrence Erlbaum.

Goswami, U. (2004). Neuroscience, education and special education. *British Journal of Special Education,* 31(4), 175-183.

Greenwood, C. R., Horton, B. T., & Utley, C. A. (2002). Academic engagement: Current perspectives on research and practice. *School Psychology Review, 31*(3), 328–349. https://doi.org/10.1080/02796015.2002.12086159

Gunn, B., Smolkowski, K., Biglan, A., Black, C. & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction: Results for Hispanic and non-Hispanic students. *Journal of Special Education, 39,* 66–85. https://dx.doi.org/10.1177/00224669050390020301

Gunn, B., Smolkowski, K. & Vadasy, P. (2011). Evaluating the effectiveness of Read Well Kindergarten. *Journal of Research on Educational Effectiveness*, *4*(1), 53–86. https://dx.doi.org/10.1080/19345747.2010.488716

Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research,* 77(1), 81-112. DOI: 10.3102/003465430298487

Herrera, S., Truckenmiller, A. J., & Foorman, B. R. (2016). *Summary of 20 years of research on the effectiveness of adolescent literacy programs and practices* (REL 2016-178). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast. http://ies.ed.gov/ncee/edlabs

Heward, W. L., & Wood, C. L (2015). *Improving educational outcomes in America: Can a low-tech generic teaching practice make a difference?* http://www.winginstitute.org/uploadedFiles/News_And_Events/Summits/2013WingSummit.pdf

Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit instruction: Historical and contemporary contexts. *Learning Disabilities Research & Practice, 32*(3), 140–148. http://dx.doi.org/10.1111/ldrp.12142

Kameenui, E. J., & Simmons, D. C. (1990). *Designing instructional strategies: The prevention of academic learning problems.* Merrill Publishing Company.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of Direct Instruction and discovery learning *Psychological Science, 15*(10), 661–667. https://doi.org/10.1111/j.0956-7976.2004.00737.x

Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning Programs: A meta-analysis. *Review of Educational Research, 60*(2), 265–299. http://dx.doi.org/10.3102/00346543060002265

Logan, C. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading and Writing Quarterly, 13*(2), 123–146. https://doi.org/10.1080/1057356970130203

MacSuga-Gage, A. S., & Simonsen, B. (2015). Examining the effects of teacher-directed opportunities to respond on student outcomes: A systematic review of the literature. *Education and Treatment of Children, 38*(2), 211–239. http://dx.doi.org/10.1353/etc.2015.0009

McCoy, K. M., & Pany, D. (1986). Summary and analysis of oral reading corrective feedback research. *Reading Teacher, 39,* 548–554.

Meltzoff, A. N., Kulh, P. K., Movellan, J., & Sejnowski, T. J. (2009). Foundations for a new science of learning. *Science, 325,* 284–288. https://dx.doi.org/10.1126/science.1175626

Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin, 86*(2), 376–390. https://doi.org/10.1037/0033-2909.86.2.376

Morgan, P. L., Farkas, G., & Wu, Q. (2011). Kindergarten children’s growth trajectories in reading and mathematics: Who falls increasingly behind? *Journal of Learning Disabilities, 44*(5), 472–488. https://dx.doi.org/10.1177/0022219411414010

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards.* http://www.corestandards.org/

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the*

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Pub. No. 00-4769). U.S. Department of Health & Human Services, National Institutes of Health, National Institute of Child Health and Human Development. https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf

Nelson-Walker, N. J., Fien, H., Kosty, D. B., Smolkowski, K., Smith, J. L. M., & Baker, S. K. (2013). Evaluating the effects of a systemic intervention on first-grade teachers' explicit reading instruction. *Learning Disability Quarterly, 36*(4), 215–230. http://dx.doi.org/10.1177/0731948712472186

Perfetti, C. A. & Bolger, D. J. (2004). The brain might read that way. *Scientific Studies of Reading, 8*(3), 293–304. https://doi.org/10.1207/s1532799xssr0803_7

Posner, M. I., DiGirolamo, G. J., & Fernandez-Duque, D. (1997). Brain mechanisms of cognitive skills. *Consciousness and Cognition, 6*(2-3), 267–290. https://doi.org/10.1006/ccog.1997.0301

Raudenbush, S. W. (2008). Advancing educational policy by advancing research on instruction. *American Educational Research Journal, 45*(1), 206–230. https://doi.org/10.3102/0002831207312905

Reutzel, D. R., Child, A., Jones, C. D., & Clark, S. K. (2014). Explicit instruction in core reading programs. *The Elementary School Journal, 114*(3), 406–430. http://dx.doi.org/10.1086/674420

Rosenshine, B. (1995). Advances in research on instruction. *Journal of Educational Research, 88*(5), 262–268. https://dx.doi.org/10.1080/00220671.1995.9941309

Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American Educator*, 12-39.

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

Samuels, S. J. & Flor, R.F. (1997). The importance of automaticity for developing expertise in reading. *Reading and Writing Quarterly: Overcoming Learning Difficulties*, 13, 107-121.

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. *Advances in Child Development and Behavior, 36,* 31–82. https://dx.doi.org/10.1016/S0065-2407(08)00002-5

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology, 59*(1), 451–475. http://dx.doi.org/10.1146/annurev.psych.59.103006.093633

Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of reliability study: Review and new results. *Statistical Methods in Medical Research, 13*(4), 251–271. https://doi.org/10.1191/0962280204sm365ra

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Smith, D. (1979). The improvement of children’s oral reading through the use of teacher modeling. *Journal of Learning Disabilities, 12*(3), 172–175.

http://dx.doi.org/10.1177/002221947901200309

Smith, J. L. M., Nelson, N. J., Smolkowski, K., Baker, S. K., Fien, H., & Kosty, D. (2016). Examining the efficacy of a multitiered intervention for at-risk readers in grade 1. *The Elementary School Journal, 116*(4), 549–573. http://dx.doi.org/10.1086/686249

Smolkowski, K. (n.d.) References to the *Classroom Observations of Student-Teacher Interactions* (COSTI). https://homes.ori.org/keiths/bibliography/measures-costi.html

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the *Classroom Observations of Student–Teacher Interactions* (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly, 27*(2), 316–328. http://dx.doi.org/10.1016/j.ecresq.2011.09.004

Sprick, M., Jones, S. V., Dunn, R., & Gunn, B. (2004). *Read Well Kindergarten: Critical foundations in beginning reading*. Sopris West.

Stein, M., Carnine, D., & Dixon, R. (1998). Direct instruction: Integrating curriculum design and effective teaching practice. *Intervention in School and Clinic, 33*(4), 227–234. https:// dx.doi.org/10.1177/105345129803300405

Stockard J., Wood, T. W., Coughlin, C., & Khoury, C. R. (2018). The effectiveness of direct instruction curricula: A meta-analysis of a half century of research. *Review of Educational Research, 20*(10), 1–19. https://doi.org/10.3102/0034654317751919

Sutherland, K. S., Alder, 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*(4), 239–248. https://dx.doi.org/doi:10.1177/10634266030110040501

Taylor, J., Roehrig, A. D., Hensler, B. S., Connor, C. M., & Schatschneider, C. (2010). Teacher quality moderates the genetic effects on early reading. *Science, 328*(5977), 512–514. http://dx.doi.org/10.1126/science.1186149

Twyman, J.S. & Heward, W.L. (2016). How to improve student learning in every classroom now. *International Journal of Educational Research, 87,* 78–90. http://dx.doi.org/10.1016/j.ijer.2016.05.007

Vadasy, P. F., Sanders, E. A. & Peyton, J. A. (2005). Relative effectiveness of reading practice or word-level instruction in supplemental tutoring: How text matters. *Journal of Learning Disabilities, 38*(4), 364–380. https://dx.doi.org/10.1177/00222194050380041401

What Works Clearinghouse. (2020). *Resources for educators, About practice guides.* https://ies.ed.gov/ncee/wwc/Resources/ResourcesForEducators