RedFame

Dexterity, Deliberateness, And Disposition: An Investigation of Instructional Strength for Early Literacy

Philip Sirinides[1], Abigail Gray[2]

[1]Pennsylvania State University at Harrisburg, United States

[2]School District of Philadelphia, United States

Correspondence: Philip Sirinides, Pennsylvania State University at Harrisburg, United States.

Abstract

This mixed-methods study examines the relationship between an emergent conceptualization of teachers' instructional strength and their students' progress in early literacy. The conceptualization includes three components: a teacher's deliberateness, their instructional dexterity, and a set of teacher dispositions that catalyse and maximize these attributes. Based on the results of qualitative inquiry, the authors developed a measure of individual teachers' instructional strength according to this conceptualization. Regression analysis reveals that all three components are significant predictors of students' growth in early literacy. The study includes 318 teachers and 1,181 students from 227 schools across the country.

Keywords: literacy, reading, early literacy, teachers, instruction, effectiveness

1. Main Text Introduction

The quality of instruction provided by teachers is a critical and variable determinant of student learning (Aaronson, Barrow & Sander, 2007; Araujo et al., 2016; Darling-Hammond, 2000; Hanushek, 2011; Rockoff, 2004), and the "proximal cause of student progress" (Raudenbush, 2008, p. 207). The work presented here builds on existing educator-effectiveness research by conceptualizing and assessing instructional strength, defined as *the extent to which a teacher instructs for maximum learning in every lesson* (Authors, 2016). We then examine the association between this conceptualization of instructional strength and student achievement in literacy.

Our findings are the product of multiple years of mixed-methods research. Some of this research was conducted as part of a large-scale evaluation of Reading Recovery, which included a rigorous, large-scale randomized trial and a deep focus on implementation. The model of instructional strength we examine here was developed, in part, in the course of that research. It posits that strong instruction can by recognized by the extent to which it exhibits two characteristics: deliberateness, which we define as *an encompassing commitment to thoughtful practice;* and dexterity, which we define as *the flexible application of deep skill.* The model further suggests that deliberateness and dexterity can best flourish when teachers also demonstrate key dispositions that incline them to make the most of their training and instructional resources. This paper expands upon that earlier work by examining, for the first time, the relationship between our conceptualization of instructional strength and student learning. While findings have immediate relevance for Reading Recovery scholars and practitioners, who number in the tens of thousands worldwide, this framework may also offer a starting point for a widely applicable perspective on teaching that has implications for teacher preparation.

Prior research on teacher effectiveness

Our conceptualization of instructional strength builds on several growing literatures. First and foremost, it builds upon research on educator effectiveness. This literature is broad, spanning efforts to identify the attributes of strong teachers, to explore the nature and development of effective teaching practice, and to quantify the impacts of high-quality instruction on student learning. One major strand of the teacher-effectiveness research seeks to identify relationships between student achievement and proxy indicators of teacher quality--for instance, teacher characteristics like years of experience, training, and credentials (Cowan & Goldhaber, 2016; Darling-Hammond, 2000; Goldhaber & Anthony, 2007; Goldhaber, Lavery, & Theobald, 2015; Hattie, 2008; Huang & Moon, 2009; Kane, Rockoff, & Staiger, 2008; Rice, 2003; Shuls, 2018; Shuls & Trivitt, 2015; Vandevoort & Berliner, 2004). Few proxy measures have been demonstrated

to reliably predict student achievement (Burroughs et al., 2019).

A second strand of educator-effectiveness literature uses value-added models to identify teachers whose students perform better than expected on standardized tests (Chetty, Friedman, & Rockoff, 2014; Gordon, Kain, & Staiger, 2006; Hanushek, 2011; Hanushek, Kain, O'Brien & Rivkin, 2005; Koedel, Mihaly, & Rockoff, 2015; Lee, 2018; Parsons, Koedel, & Tan, 2018; Rivkin, Hanushek & Kain, 2005). In addition to its susceptibility to bias based on differences in teachers' student assignments and other factors (Rothstein, 2009), the value-added approach to identifying effective teachers offers scant insight into what, specifically, better teachers are doing to drive learning. Therefore, while value-added studies can arguably shed light on the contributions of teacher quality, they offer schools far less actionable information, and give us little to reflect on in efforts to improve teacher preparation.

A third strand of the effectiveness literature focuses on connecting teachers' instructional practice to student-level impacts. Growing interest in observation tools as a means of understanding teacher practice has spawned recent research connecting lesson observation data with student outcomes (Garrett & Steinberg, 2015; Hill & Grossman, 2013; Lekwa, Reddy, & Shernoff, 2019; Pianta & Hamre, 2009). With its focus on identifying features of strong instruction and on assessing their relationship to learning, the present study is well aligned with this literature. However, our work is not designed to validate or investigate a specific evaluative instrument. Rather, we join scholars interested in theoretical and empirical exploration of the "*what* and *how* of quality teaching" (Ingvarson & Rowe, 2007, p. 2) (Bietenbeck, 2014; Gitomer, 2019; Gage & Needels, 1989; Harbour, Evanovich, Sweigart, & Hughes 2015; Hidalgo-Cabrillana & Lopez-Mayan, 2018; Schwert & Wupperman, 2011). As such, this work answers calls for the development and testing of new conceptualizations of "what teachers *should know* and *be able to do*" (Ingvarson & Rowe, 2008, p. 5).

Along with its contributions to the teacher-effectiveness literature, this work offers a new take on some established ways of thinking about instructional practice. The conceptualization of instructional strength we present here is characterized, in part, by the presence of two distinct manifestations: The first of these we call deliberateness, defined as *an encompassing commitment to thoughtful practice*. We define the second component, instructional dexterity, as *the flexible application of deep skill* (Authors, 2016). These central aspects of our model are indebted to the large body of existing research on reflective practice that descends from the work of Dewey (1933), Schön (1983, 1987, 1991), and Kolb (1984).

More specifically, this conceptualization of instructional strength incorporates aspects of two different types of reflection described in the literature: Reflection *on* action, which refers to the application of insights about past actions and results to decisions about future plans, and reflection *in* action, which takes place in the moment of activity (Griffiths & Tan, 1992; Mann & Walsh, 2013; Schön, 1983; Schön, 1987; Richert, 1992; van Menen, 1991). The former, reflection on action, informs our understanding of *deliberateness* as a primary component of instructional strength. Conversely, *instructional dexterity* builds upon the theory of reflection in action—an idea further elaborated in existing research on teacher thinking and decision-making (Ball & Forzani, 2009, Eraut, 1995; Burhan-Horasanlı & Ortaçtepe, 2016), and research on scaffolding (Rodgers, 2004; Rodgers, D'Agostino, Harmey, Kelly, & Brownfield, 2016), and research on instructional scaffolding (Johnson, 2019). Our definition of instructional dexterity grounds this theory in specific practices and qualities that are observable and, we hypothesize, teachable (Authors, 2016). We see Reading Recovery teachers engaged in decision-making "in the midst of action, without interrupting what [they are] already doing, and reshaping it at the same time" (Yanow & Tsoukas, 2009, p. 1340), and observe that dexterous Reading Recovery instruction exhibits "the capacity to reach into the center of confusing situations, to see itself, and to shift the base of its operations or pull up stakes altogether and follow the flow of the action" (Tremmel, 1993, p. 437). In elaborating the ways these ideas are manifested in strong Reading Recovery instruction, we answer calls for research that can ground the broad concept of reflective practice in specific and concrete elements of strong teaching (Korthagen and Wubbels, 1995; Mann & Walsh, 2013).

Along with positing instructional dexterity and deliberateness as two components or dimensions of instructional strength, our conceptualization includes a set of teacher dispositions that, we hypothesize, equip some teachers to be particularly receptive to training and to deliver particularly strong instruction. Teacher dispositions have been studied for decades (Getzels & Jackson, 1963; Richardson, 1996; Sockett, 2006), and research has associated particular dispositions or aspects of teachers' personalities with impactful instruction (Bastian, McCord, Marks, & Carpenter, 2017; Cheng & Zamarro, 2018; Fairhurst & Fairhurst,1995; Kim, Jorg, & Klassen, 2019; Klassen et al., 2017; Patrick, 2011; Rushton, Morgan, Richard 2007). However, scholars have long debated how best to understand dispositions as they pertain to teachers and teaching. Burant, Chubbuck & Whipp (2007) summarized the literature in this area as encompassing three general characterizations: Dispositions are variably understood as beliefs and/or attitudes, as aspects of personality (Jung & Rhodes, 2008), or as observable behaviors. Citing Callahan, Waciscko, & Wirtz (2004), Bolton & Reisboard defined dispositions broadly, as "attitudes, beliefs, self concepts, self efficacy, predispositions,

work ethics, feelings, morals, behaviors, skills, intentions, competence, or values" (2014, p. 25). Our own definition imagines teacher dispositions as *inclinations*—that is, as ways of approaching students and teaching that result from a combination of innate and learned perspectives and characteristics. We do not attempt, here, to discern the precise origins of these inclinations. Rather, we are interested in better understanding their manifestation in strong instruction and, ultimately, in the relationship between teaching and learning. In exploring all of these aspects of instructional strength, we address the following research questions: What are the characteristics of the strongest Reading Recovery instruction? What dispositions incline teachers to deliver strong instruction? And, does a teacher's instructional strength predict her students' growth in reading?

2. Context of the Research

Reading Recovery is a decades-old literacy intervention that is widely used in the US, Australia, New Zealand, the UK, and elsewhere. The intervention is delivered to struggling first-grade students by teachers specially trained to implement individualized teaching plans for each student (Clay, 2005), and unfolds over the course of several months of daily, individual, 30-minute lessons. Reading Recovery's basis is in constructivist learning theory (Bruner, 1966, 1983; Dewey, 1916; Piaget, 1967), which views learning as an active process in which students are responsible for constructing meaning based on their own beliefs and experiences.

Constructivist theorists—Reading Recovery's creator Marie Clay among them (Clay, 1991, 1998)—regard the teacher's role in the learning process as primarily concerned with providing the student with opportunities to learn. A skillful teacher, from this perspective, is adept at anticipating, identifying, and arranging learning opportunities (Clay, 2005).

Reading Recovery lessons follow a prescribed pattern. Each includes a series of activities that progress in a particular order. Within this structure, the teacher exercises her discretion-informed by Reading Recovery theory and training-in selecting the instructional foci, texts, and teaching moves that best suit each student. Students' progress and lesson activities are carefully tracked via running records and other lesson documentation (Authors, 2018). Because of this program model—which embeds idiosyncratic teaching within a consistent structure or frame-Reading Recovery offers an ideal setting for close study of variations in instructional technique.

An important role in this research is that of the Reading Recovery teacher coach-or "teacher leader" in Reading Recovery parlance. Teacher leaders teach the year-long university-based courses Reading Recovery teachers are required to complete as part of their training, and support the teachers through intensive, ongoing, on-site coaching that includes regular observation of instruction. Nearly all teacher leaders are experienced Reading Recovery teachers themselves, with deep knowledge of the program's lesson structure and theory. They are the primary sources of the data we used to conceptualize, operationalize, and measure instructional strength.

This research was conducted, in part, within the i3 evaluation of Reading Recovery. Funded by a grant from the U.S. Department of Education's Office of Innovation and Improvement's (i3) *Scaling Up What Works* program, the evaluation was conducted between 2011 and 2014. The current study grew out of the observation that Reading Recovery lessons vary in important ways that have meaningful consequences for student learning. This observation echoes with a major finding of the randomized trial component of the evaluation: While Reading Recovery students outperformed control-group students in reading overall, some students' growth was particularly remarkable (Authors, 2018).

3. Materials and Methods

Data for this study come from interviews, observations, and focus groups with Reading Recovery practitioners, as well as student-level data achievement data, were collected in the course of the i3-funded evaluation. Following the conclusion of the impact evaluation, we developed a measure of instructional strength and administered as an online survey to teacher leaders who worked as a mentor with teachers in the study. Finally, we conducted regression analysis to examine the extent to which these individual teachers' instructional strength scores, as calculated from their mentors' survey responses, predicted the reading achievement of their students.

Sample

Participants in the research included Reading Recovery teachers, teacher leaders, and university-based regional supervisors, as well as first-grade students who received Reading Recovery lessons during the 2013-2014 school year (See Table 1). All teacher leaders and regional supervisors involved in the i3 scale-up were contacted for data collection activities, and all who responded were included in interviews and focus groups. Teachers in the study were randomly selected for participation in interviews.

Table 1. Research activities and participants

Activity	Regional supervisors	Teacher leaders	Teachers (schools)	Students
Interviews*	19	24	83	--
Lesson Observations*	--	--	108	--
Focus Groups*	--	48	--	--
Survey**	--	120	--	--
Observation Survey (OS)*	--	120	318 (227)	1,181

* *Data collected during the i3 evaluation.*

** *Data collected following the conclusion of the i3 evaluation*

In the summer of 2014, after the conclusion of the impact evaluation, we surveyed 120 teacher leaders who had supervised 318 teachers during the 2013-2014 school year. These teachers and students were from 227 schools across the country. We then used the resulting data on these teachers' instructional strength, along achievement data from 1,181 of their students, to explore the relationship between instructional strength and student progress in reading.

The student outcome data used for this exploration were collected as part of the multi-site, delayed-treatment design of the RCT study. Once deemed eligible for Reading Recovery by virtue of their low scores on screening assessments, first-grade students in the schools in the study were matched in pairs and randomly assigned, within their pairs, to treatment and control. Students assigned to treatment received Reading Recovery in the first part of the school year; those assigned to control received it once their matched counterparts had completed the program. Both students were assessed at the beginning and end of their Reading Recovery intervention period (Authors, 2018). Student impacts used in the current study were estimated using data from students in both groups.

All participants in the research provided signed or verbal consent to the data collection. All consent procedures and data collection instruments were reviewed and approved by the University of Pennsylvania's Institutional Review Board. Student assessment data were collected with parental consent as a routine part of their participation in their schools' Reading Recovery programs, and then shared with the research team.

Interviews and observations

In the first two years of i3 study, we conducted semi-structured interviews with Reading Recovery teachers, teacher leaders, and the university faculty members who train teacher leaders and oversee Reading Recovery via its network of university-based regional hubs. At this stage, we used interview protocols designed to elicit general information about barriers and facilitators of Reading Recovery's implementation in schools. Interviews were recorded and transcribed. After reaching at least 80 percent inter-rater reliability, researchers coded the data using both general codes derived from the interview protocols and codes that emerged inductively via multiple reviews of the data (Ravitch & Carl, 2016). Through our constant-comparative analysis of these data (Miles, Huberman, & Saldaña, 2014), we identified "lesson quality" as a key implementation factor (Authors, 2016).

In the final two years of the evaluation, we conducted targeted inquiry around this theme of lesson quality. In year three, we revised the protocol for our interviews with university faculty to include questions about how Reading Recovery lessons vary from teacher to teacher and from day to day, and about what instructional practices and contexts best support student learning. Data from these interviews were prepared for analysis via the process described above; this time, however, researchers applied both general implementation codes and also sub-codes specific to instruction (e.g. "teacher prompting") within pertinent parent codes (e.g. "lessons"). The sub-codes related to instruction were developed collaboratively by the researchers through multiple reviews of the data. Through constant comparative analysis of the coded data, we identified the themes that ultimately informed our conceptualization of instructional strength.

This conceptualization was further informed by the research team's 108 field-based observations of Reading Recovery lessons. These observations were guided by a protocol we developed to record the sequence of activities in each lesson, when and how the teacher prompted students, techniques the teacher used to instruct students on specific skills, and the proportion of lesson time dedicated to each activity. Following each observation, researchers conducted a short interview with the teacher, who was asked to explain the specific instructional decisions she made during the lesson (Emerson, Fretz, & Shaw, 2011). When possible, researchers observed lessons in the presence of teacher leaders, who were also interviewed about the lesson. In combination with our interviews, these observations allowed the research team to develop deep familiarity with Reading Recovery lessons.

Focus Groups

To confirm and further refine our emerging conceptualization, in the final year of the i3 study we conducted five teacher leader focus groups. Each hour-long focus group was held in a different state and included teacher leaders affiliated with a different university-based training site. Focus groups included between six and twelve teacher leaders, each serving as instructor and coach for a group of Reading Recovery teachers involved in the evaluation. During the focus groups, we asked teacher leaders to describe, as concretely as possible, ways in which individual lessons vary and the characteristics of particularly strong lessons. Focus group transcripts were coded using deductive codes related to the hypothesized components of instructional strength, and open coded to ensure no themes were missed. We then applied subcodes for the specific manifestations and features of each dimension of instructional strength.

Survey

The results of four years of observation and qualitative inquiry informed an emerging conceptualization of instructional strength that includes two distinct but related manifestations--instructional dexterity and deliberateness--and a set of underlying teacher dispositions. Our efforts to understand this conceptualization and its significance for student learning continued beyond the end of the study with the development and administration of a survey for teacher leaders who had supported Reading Recovery teachers in the i3 scale-up during the 2013-14 school year. Teacher leaders were identified as respondents because of both their deep expertise about early literacy instruction in general, and their familiarity with the idiosyncrasies of individual teachers' instruction.

Each of the three components of instructional strength was operationalized using multiple indicators that could be administered as items in an online survey instrument. Each item asked teacher leaders to evaluate the teacher relative to an average Reading Recovery teacher using a six-point Likert scale of agreement. Items were eliminated and/or refined through a process that included a pilot administration to a group of teacher leaders and university-based trainers, followed by cognitive interviews with respondents to ensure that all items were worded clearly and interpreted consistently. Ultimately, we administered and included in the analysis eight items designed to assess teachers' deliberateness, six items about instructional dexterity, and five items about dispositions. The survey was administered in Summer, 2014, after the impact study was finished. In completing the survey, each teacher leader was asked to respond to a series of items about every teacher she supported during the 2013-2014 school year.

4. Regression Analysis

The final step in our research was to examine the extent to which instructional strength—as represented by the individual teacher scores yielded by the survey—predicts student achievement in reading. The outcome measure was the Observation Survey of Early Literacy Achievement (OS) (Clay, 2002, 2005, 2016). The OS is the standard assessment used in Reading Recovery. It is an individually administered standardized assessment comprising six sub-scales that, combined, form a unidimensional measure of early reading achievement (D'Agostino, Rodgers, & Mauch, 2018). Multiple studies have supported the validity of the OS as a measure of early reading achievement (Denton, Ciancio, Fletcher, 2006; Gómez-Bellengé, Gibson, Tang, Doyle, & Kelly, 2007; Gómez-Bellengé, Rodgers, Wang, & Schultz, 2005; NCRTI, 2012; Tang & Gómez-Bellengé, 2007). The OS measure was administered to all students prior to receiving Reading Recovery, and again at the end of each student's intervention period (Authors, 2018). The date of each assessment and the number of Reading Recovery lessons provided, a measure of dosage, were also recorded for each student.

To estimate the impact of teachers' instructional strength on their students' reading achievement, we regressed student-level reading achievement data on teacher instructional strength scores from the survey. We used a hierarchical linear model (HLM) in which students were nested within teachers, and teachers nested within schools. The dimensions of instructional strength reflected in our conceptual model, while related, are distinct. Each looks different in practice and, we hypothesize, requires different types of supports or interventions to improve. For this reason, we estimate separate HLM models for the three dimensions.

To account for differences in teacher leaders' perceptions and reporting, and to reflect the fact that all the Reading Recovery teachers in a given school are typically supported by a single teacher leader, we included a fourth level in the model to nest schools within teacher leaders. In the interests of parsimony, we included only two covariates in the model: OS pre-test total score and the total number of Reading Recovery lessons provided to each student. This latter covariate was included to account for the fact that some students received more lessons—and therefore a higher dose of their teachers' instruction—than others. Table 2 shows the distribution of these variables. Based on national percentile ranks from the instruments U.S. norming sample (2012), the average scores at entry and exit are associated with the 17th percentile among students in the Fall, and the 39th percentile at mid-year.

Table 2. Dependent and independent variables

	Study Sample				National Percentile Ranks		
	Mean	StD	Min	Max	Fall	Mid-year	Year-end
Exit OS Total Score	502.9	41.5	325	609	90	39	8
Entry OS Total Score	381.0	36.8	245	480	17	1	1
Lessons	66.0	13.1	7	97			

5. Results

In this section, we first summarize the results of our qualitative inquiry into instructional strength, which was conducted during the i3 evaluation (Authors, 2018). We then discuss the findings of work completely following the larger study, including our examination of the factor structure of the instructional strength construct, and our exploration of the relationship between instructional strength and student learning.

5.1 Results of the Qualitative Inquiry

As noted, our qualitative research produced a hypothesized model of instructional strength that includes two manifestations of instructional strength, deliberateness, and instructional dexterity. While deliberateness and dexterity are related, we observed qualitatively that they are manifested in different ways and at different times. We observed, further, that each is expressed through a set of concrete behaviors or practices on the part of Reading Recovery teachers, and that the strongest instruction is characterized by high levels of both deliberateness and dexterity. Our research also revealed a set of dispositions that underlie both deliberateness and instructional dexterity, and that facilitate the demonstration and development of instructional strength.

In our conceptualization, deliberateness is defined as an encompassing commitment to thoughtful practice, and we see it in teachers' consistent engagement in three types of activities: the use of careful lesson documentation for purposes of analyzing individual students' progress and planning instruction, continual reflection on their own instruction, and dogged pursuit of knowledge that includes participation in a community of practice. These activities take place primarily in the time in between Reading Recovery lessons, rather than during instruction itself. More specifically, we observe that strong Reading Recovery teachers are not just faithful documenters of their students' progress; they are expert noticers-skilled at converting subtle observations about students' strengths, weaknesses, and interests into data that inform all aspects of their instruction. Further, strong teachers take their Reading Recovery training very seriously, and attend training activities faithfully. They also access the expertise of their teacher leaders regularly, and build and routinely engage with practitioner networks of teachers.

The second manifestation of instructional strength, instructional dexterity, is defined as the flexible application of deep skill. Unlike deliberateness, instructional dexterity is primarily evident during lesson enactment. The first expression of dexterity that that we observe concerns the nature of the teacher's rapport with her students. The strongest instruction is characterized by warm but firm support, and the strongest teachers can judge when the time is right to provide encouragement, and when a student can be pushed to call upon and apply new skills independently. In doing so, these teachers help their students develop confidence and self-sufficiency. Another expression of instructional dexterity we identify is the ability to make instantaneous instructional decisions based on both prior information and in-the-moment observations. While Reading Recovery teachers are taught to make a plan ahead of each lesson, dexterous teachers can deviate from that plan in response to students' successes or their difficulties. Students' reading and writing errors present particularly rich teaching opportunities for dexterous teachers, as they know how to acknowledge the student's effort and quickly transition the mistake into new learning. Finally, we observe that dexterous teachers exhibit economy of language. They speak mostly to arrange learning opportunities or prompt students' application of skills, rather than to explain or instruct. As a result, their statements are brief, clear, and redirecting. Typical prompts by a dexterous teacher might consist of statements like "Again" or "Try a word that looks like that word." According to teacher leaders, this judicious use of language is an indication that a teacher is successfully identifying the student's strengths and development needs and arranging appropriate practice opportunities.

We also observe that underlying both deliberateness and instructional dexterity is a set of dispositions that facilitate the development of instructional strength. By and large, respondents described these attitudes and dispositions as attributes that interact with Reading Recovery training to support the behaviors we associate with deliberateness and dexterity. In our conceptualization, dispositions undergirding instructional strength include the following: Openness to change, interpersonal communication, work ethic, and belief in the ability of all students to learn.

5.2 Factor Structure

To investigate the composition and structure of the theorized dimensions, survey items were submitted to a confirmatory factor analysis (Anderberg, 1973; Harman, 1976). We found that all items had a higher squared multiple correlation with items in its theorized scale, than with the items of the other scale. Although the two scales were highly correlated (r=0.86), item membership was able to explain 82.7% of the total variance, and each scale had high internal consistency ≥.95. Item means, standard deviations, item-total correlations and conditional alphas are presented by scale in Table 3.

Table 3. Instructional Strength Survey Item Characteristics

	Mean	Std Dev	I/T corr	cond a
Deliberateness (alpha = 0.96)				
effectively evaluates his/her own performance and growth	4.24	1.31	0.91	0.95
effectively utilizes supports to improve instruction	4.31	1.26	0.90	0.95
effectively consults theory or literature related to Reading Recovery	3.96	1.32	0.88	0.95
effectively reflects on lessons	4.12	1.27	0.87	0.95
consistently uses lesson documentation to plan instruction	4.12	1.30	0.87	0.95
consistently documents lessons	4.23	1.27	0.80	0.96
exhibits commitment to continual improvement of practice	5.16	0.97	0.79	0.96
exhibits openness to change in practices/ways of thinking	5.11	1.05	0.73	0.96
Dexterity (alpha = 0.98)				
consistently choose impactful teaching points for each student	3.97	1.26	0.94	0.97
consistently adapts instruction moment to moment within the lesson	3.87	1.29	0.93	0.97
consistently corrects or interrupts students appropriately during lesson	3.89	1.21	0.92	0.97
consistently uses crisp language during lessons	3.88	1.31	0.91	0.97
consistently releases responsibility appropriately during lessons	3.84	1.25	0.91	0.97
consistently transitions smoothly between activities during the lesson	4.24	1.22	0.90	0.97
Dispositions (alpha = 0.93)				
exhibits high expectations for student learning	5.05	1.02	0.88	0.91
exhibits a natural gift for teaching	4.89	0.98	0.83	0.92
exhibits a belief in the ability of all students to learn to read well	4.99	1.07	0.83	0.92
exhibits a sense of urgency	4.89	1.14	0.83	0.92
exhibits a strong work ethic	5.26	0.94	0.75	0.93

Note: n=318 teachers; all items measured on a six-point Likert scale of agreement

Through qualitative inquiry and instrument development, our research suggests that both deliberateness and dexterity exist on continua for individual teachers. Based on results of the item analysis, teacher scores for both dimensions of instructional strength and teach attitude were calculated as within-scale item averages.

Relationship to student achievement

The study next examined associations between scales of instructional strength and student-level reading achievement data. Table 4 presents the fixed and random effects for a model that was estimated three times for each of survey scales. Random intercept coefficients were statistically significant for all four levels: Teacher leader, school, teacher, and student. Fixed effect estimates for baseline assessment and the number of lessons between measures were also statistically significant in the three models. The significant main effects for pretest and number of lessons indicates that students who started at higher reading levels, and who received greater number of one-to-one lessons, on average scored higher.

Table 4. Total OS score regressed on dimensions of Instructional Strength

	Deliberateness		Dexterity		Dispositions	
IS Scale [a]						
# items (reliability)	8 (0.96)		6 (0.98)		5 (0.93)	
Mean (SD)	4.41 (1.08)		3.95 (1.19)		5.02 (0.92)	
Random Effects (s.e.)						
Teacher Leader	103.8 (46.32)	*	101.7 (46.87)	*	105.2 (46.51)	*
School	108.7 (61.48)	*	115.6 (62.40)	*	100.7 (61.25)	*
Teacher	117.7 (49.77)	**	111.7 (49.52)	*	122.1 (50.18)	*
Student (residual)	708.1 (34.26)	***	708.2 (34.27)	***	708.1 (34.25)	***
Fixed Effects (s.e.)						
Intercept	175.6 (13.3)		177.2 (13.0)		171.2 (13.9)	
Entry Total OS	0.7 (0.03)	***	0.7 (0.03)	***	0.7 (0.03)	***
# Lessons	0.5 (0.08)	***	0.5 (0.08)	***	0.5 (0.08)	***
IS scale [a]	2.5 (1.22)	*	2.6 (1.12)	*	3.2 (1.44)	*

Notes: [a] IS scale name in column header; Analytic sample includes 1181 students, 318 teachers, 227 schools, 120 Teacher Leaders; * $p<0.05$, ** $p<0.01$, *** $p<.0001$

The estimated effects for Instructional Strength scales were also statistically significant. We estimated a 2.5 and 2.6 increase in Total OS scores for every 1-point increase on scales of deliberateness and dexterity, respectively. These gains are equivalent to the estimated improvements after five lessons (typically 1 week of individual instruction). Fully standardized coefficients associated with these effects are 0.06 and 0.07 standard deviations. Findings offer evidence that Reading Recovery teachers' instructional strength is associated with their students' improved reading achievement. The disposition scale was also a significant predictor of student reading achievement with effect size comparable to the IS scales.

6. Discussion

This study addresses both the nature of instructional strength and its relationship to reading achievement. Findings suggest that the strongest Reading Recovery lessons are similar to one another in observable ways and are conducted by teachers who behave in similar ways. The students who received strong instruction, as conceptualized in this study, were observed to make the greatest gains in reading. This suggests future utility of this conceptualization in terms of understanding instruction. More specifically, our findings contribute to the teacher-effectiveness literature by highlighting a theory that not only links instruction to learning, but also offers specificity on the instructional activities that make some teachers especially effective.

All three scales in our measure showed small but significant effects on student achievement. Along with confirming our theory about the instructional practices strong teachers use, this research therefore confirms our hypothesis that particular dispositions incline some teachers to leverage their training and resources more effectively than others. This represents a contribution to prior research on the role of teachers' personal characteristics in their students' learning.

Because the goal of the i3 grant was to scale up Reading Recovery nationally by training new teachers to deliver the intervention, the sample consists largely of teachers recently trained to implement the program. Thus, one limitation of the study is that we are unable to assess whether instructional strength looks different or produces different outcomes in teachers with more experience using Reading Recovery. A second limitation is that the research that informs our conceptualization of instructional strength focused specifically on Reading Recovery teachers, who have gone through a rigorous training process in order to enact a particular type of instruction. The generalizability of our findings to teachers with different training backgrounds is unknown and warrants further exploration.

Future research may seek to address the types of training that can best support the development of instructional dexterity and deliberateness Findings suggest that training models must attend to teacher attitudes and dispositions which improve receptiveness to training and lead to instructional strength for early literacy. We hope findings from this study will advance an understanding of both the practices and attributes of strong teachers in ways that are useful for educator preparation and development.

References

Authors (2016).

Authors (2018).

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135. https://doi.org/10.1086/508733

Anderberg, M. R. (1973). *Cluster analysis for applications*. New York, Academic Press.

Araujo, M. C., Carneiro, P., Cruz-Aguayo, Y., & Schady, N. (2016). Teacher quality and learning outcomes in kindergarten. *The Quarterly Journal of Economics, 131*(3), 1415-1453. https://doi.org/10.1093/qje/qjw016

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497-511. https://doi.org/10.1177/0022487109348479

Bastian, K. C., McCord, D. M., Marks, J. T., & Carpenter, D. (2017). A temperament for teaching? Associations between personality traits and beginning teacher performance and retention. *AERA Open, 3*(1). https://doi.org/10.1177/2332858416684764

Bietenbeck, J. (2014). Teaching practices and cognitive skills. *Labour Economics, 30*, 143-153. https://doi.org/10.1016/j.labeco.2014.03.002

Bolton, M., & Reisboard, D. (2014). Improving Pre-Service Teacher Dispositions. *Journal of Teaching Effectiveness and Student Achievement, 1*(1), 24-32.

Bruner, J. (1983). Play, thought, and language. *Peabody Journal of Education*, *60*(3), 60-69. https://doi.org/10.1080/01619568309538407

Bruner, J. S. (1966). *Toward a theory of instruction* (Vol. 59). Harvard University Press.

Burhan-Horasanlı, E., & Ortaçtepe, D. (2016). Reflective practice-oriented online discussions: A study on EFL teachers' reflection-on, in and for-action. *Teaching and Teacher Education*, *59*, 372-382. https://doi.org/10.1016/j.tate.2016.07.002

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). A Review of the Literature on Teacher Effectiveness and Student Outcomes. In *Teaching for Excellence and Equity* (pp. 7-17). Springer, Cham. https://doi.org/10.1007/978-3-030-16151-4_2

Cheng, A., & Zamarro, G. (2018). Measuring teacher non-cognitive skills and its impact on students: Insight from the Measures of Effective Teaching Longitudinal Database. *Economics of Education Review*, *64*, 251-260. https://doi.org/10.1016/j.econedurev.2018.03.001

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2013a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates (Working Paper No. 19423). https://doi.org/10.3386/w19423

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633-2679. https://doi.org/10.1257/aer.104.9.2633

Clay, M. M. (1991). *Becoming literate: The construction of inner control*. Portsmouth, NH: Heinemann.

Clay, M. M. (1998). *By different paths to common outcomes*. Portland, ME: Stenhouse Publishers.

Clay, M. M. (2005a). *Literacy Lessons Designed for Individuals: Why? when? and how?*. Portsmouth, NH: Heinemann.

Cowan, J., & Goldhaber, D. (2016). National board certification and teacher effectiveness: Evidence from Washington State. *Journal of Research on Educational Effectiveness, 9*(3), 233-258. https://doi.org/10.1080/19345747.2015.1099768

D'Agostino, J. V., Rodgers, E., & Mauck, S. (2018). Addressing inadequacies of the observation survey of early literacy achievement. *Reading Research Quarterly*, *53*(1), 51-69. https://doi.org/10.1002/rrq.181

Darling-Hammond, L. (2000). Teacher Quality and Student Achievement: A Review of State Policy Evidence. *Education Policy Analysis Archives, 8*(1), 1-44. https://doi.org/10.14507/epaa.v8n1.2000

Denton, C. A., Ciancio, D. J., & Fletcher, J. M. (2006). Validity, reliability, and utility of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly*, *41*(1), 8-34. https://doi.org/10.1598/RRQ.41.1.1

Dewey, J. (1933). *Democracy and education: An introduction to the philosophy of education.* New York, NY: Macmillan.

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226206868.001.0001

Eraut, M. (1995). Schon Shock: a case for refraining reflection-in-action?. *Teachers and teaching*, *1*(1), 9-22. https://doi.org/10.1080/1354060950010102

Estice, R. M. (1997). Using patterns of responding to "follow the child". *Running Record*, *9*(2), 1, 6-7.

Fairhurst, A. M., & Fairhurst, L. L. (1995). *Effective teaching, effective learning: Making the personality connection in your classroom*. Hachette UK.

Gage, N. L., & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. *The Elementary School Journal, 89*(3), 253-300. https://doi.org/10.1086/461577

Getzels, J. W., & Jackson, P. W. (1963). The teacher's personality and characteristics. In Gage (Ed.) *Handbook of research on teaching*, 506-582.

Gitomer, D. H. (2019). Evaluating instructional quality. *School Effectiveness and School Improvement*, *30*(1), 68-78. https://doi.org/10.1080/09243453.2018.1539016

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *The Review of Economics and Statistics*, *89*(1), 134-150. https://doi.org/10.1162/rest.89.1.134

Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational researcher*, *44*(5), 293-307. https://doi.org/10.3102/0013189X15592622

Gómez-Bellengé, F., Gibson, S., Tang, M., Doyle, M., & Kelly, P. (2007). *Assessment and identification of first-grade students at risk: Correlating the dynamic indicator of basic early literacy skills and an Observation Survey of early literacy achievement*. Retrieved from https://www.idecweb.us/

Gordon, R., Kane, T. J., & Staiger, D. O. (2006). Identifying Effective Teachers Using Performance on the Job. The Hamilton Project Policy Brief No. 2006-01. *Brookings Institution*.

Griffiths, M., & Tann, S. (1992). Using reflective practice to link personal and public theories. *British Journal of Teacher Education, 18*(1), 69-84. https://doi.org/10.1080/0260747920180107

Hanushek, E. A. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479. https://doi.org/10.1016/j.econedurev.2010.12.006

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality* (No. w11154). Cambridge, MA: National Bureau of Economic Research. https://doi.org/10.3386/w11154

Harbour, K. E., Evanovich, L. L., Sweigart, C. A., & Hughes, L. E. (2015). A brief review of effective teaching practices that maximize student engagement. *Preventing School Failure: Alternative Education for Children and Youth*, *59*(1), 5-13. https://doi.org/10.1080/1045988X.2014.919136

Harman, H. H. (1976). *Modern factor analysis*. University of Chicago Press.

Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge: London. https://doi.org/10.4324/9780203887332

Hidalgo-Cabrillana, A., & Lopez-Mayan, C. (2018). Teaching styles and achievement: Student and teacher perspectives. *Economics of Education Review*, *67*, 184-206. https://doi.org/10.1016/j.econedurev.2018.10.009

Hill, H. C., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, *83*(2), 371-384. https://doi.org/10.17763/haer.83.2.d11511403715u376

Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation and Accountability, 21*(3), 209-234. https://doi.org/10.1007/s11092-009-9074-2

Ingvarson, L., & Rowe, K. (2008). Conceptualising and evaluating teacher quality: Substantive and methodological issues. *Australian journal of education*, *52*(1), 5-35. https://doi.org/10.1177/000494410805200102

Johnson, E. M. (2019). Choosing and Using Interactional Scaffolds: How Teachers' Moment-to-Moment Supports Can Generate and Sustain Emergent Bilinguals' Engagement with Challenging English Texts. *Research in the Teaching of English*, *53*(3), 245-269.

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review, 27*(6), 615-631. https://doi.org/10.1016/j.econedurev.2007.05.005

Kim, L. E., Jörg, V., & Klassen, R. M. (2019). A Meta-Analysis of the Effects of Teacher Personality on Teacher Effectiveness and Burnout. *Educational Psychology Review, 31*(1), 163-195. https://doi.org/10.1007/s10648-018-9458-2

Klassen, R. M., Durksen, T., Kim, L., Patterson, F., Rowett, E., Warwick, J., et al. (2017). Developing a proof-of-concept selection test for entry into primary teacher education programs. *International Journal of Assessment Tools in Education, 4*, 96-114. https://doi.org/10.21449/ijate.275772

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*, *47*, 180-195. https://doi.org/10.1016/j.econedurev.2015.01.006

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. New York, NY: Pearson Education.

Korthagen, F. A., & Wubbels, T. (1995). Characteristics of reflective practitioners: Towards an operationalization of the concept of reflection. *Teachers and Teaching: Theory and Practice*, *1*(1), 51-72. https://doi.org/10.1080/1354060950010105

Lee, S. W. (2018). Pulling back the curtain: Revealing the cumulative importance of high-performing, highly qualified teachers on students' educational outcome. *Educational Evaluation and Policy Analysis*, *40*(3), 359-381. https://doi.org/10.3102/0162373718769379

Lekwa, A. J., Reddy, L. A., & Shernoff, E. S. (2019). Measuring teacher practices and student academic engagement: A convergent validity study. *School Psychology*, *34*(1), 109. https://doi.org/10.1037/spq0000268

Mann, S., & Walsh, S. (2013). RP or 'RIP': A critical perspective on reflective practice. *Applied Linguistics Review, 4*(2), 291-315. https://doi.org/10.1515/applirev-2013-0013

Miles, M. B., & Huberman, & A. M. Saldana (2014). *Qualitative data analysis: A methods sourcebook, Third edition*. SAGE: Thousand Oaks, CA.

National Center on Response to Intervention. (2012). Screening Charts. Retrieved April 3, 2012, from http://www.rti4success.org/

Parsons, E., Koedel, C., & Tan, L. (2019). Accounting for Student Disadvantage in Value-Added Models. *Journal of Educational and Behavioral Statistics*, *44*(2), 144-179. https://doi.org/10.3102/1076998618803889

Patrick, C. L. (2011). Student evaluations of teaching: Effects of the Big Five personality traits, grades and the validity hypothesis. *Assessment & Evaluation in Higher Education, 36*, 239-249. https://doi.org/10.1080/02602930903308258

Piaget, J. (1971). *Biology and knowledge*. Chicago, IL: University of Chicago Press. (First French ed., 1967).

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119. https://doi.org/10.3102/0013189X09332374

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

Ravitch S. M., & Mittenfelner Carl, N. M. (2016). *Qualitative research: bridging the conceptual, theoretical, and methodological*. Thousand Oaks, CA: Sage Publications.

Reynolds, D., & Daniel, S. (2018). Toward contingency in scaffolding reading comprehension: Next steps for research. *Reading Research Quarterly*, *53*(3), 367-373. https://doi.org/10.1002/rrq.200

Rice, J. K. (2003). Teacher quality: Understanding the effectiveness of teacher attributes. Washington, DC: Economic Policy Institute.

Richardson, V. (1996). From Behaviorism To Constructivism In Teacher Education. *Teacher Education and Special Education*, *19*(3), 263-271. https://doi.org/10.1177/088840649601900324

Richert, A. E. (1992). Writing cases: A vehicle for inquiry into the teaching process. *Case methods in Teacher Education*, 155-174.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic

achievement. *Econometrica, 73*(2), 417-458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review, 94*(2), 247-252. https://doi.org/10.1257/0002828041302244

Rodgers, E. M. (2004). Interactions that scaffold reading performance. *Journal of Literacy Research*, *36*(4), 501-532. https://doi.org/10.1207/s15548430jlr3604_4

Rodgers, E., D'Agostino, J. V., Harmey, S. J., Kelly, R. H., & Brownfield, K. (2016). Examining the nature of scaffolding in an early literacy intervention. *Reading Research Quarterly*, *51*(3), 345-360. https://doi.org/10.1002/rrq.142

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education finance and policy*, *4*(4), 537-571. https://doi.org/10.1162/edfp.2009.4.4.537

Rushton, S., Morgan, J., & Richard, M. (2007). Teacher's Myers-Briggs personality profiles: Identifying effective teacher personality traits. *Teaching and Teacher Education*, *23*(4), 432-441. https://doi.org/10.1016/j.tate.2006.12.011

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126). New York, NY: Basic Books.

Schön, D. A. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey-Bass Inc.

Schön, D. A. (1991). *The reflective turn: Case studies in and on educational practice*. New York, NY: Teachers Press, Columbia University.

Schwerdt, G., & Wuppermann, A. C. (2011). Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review*, *30*(2), 365-379. https://doi.org/10.1016/j.econedurev.2010.11.005

Shuls, J. V. (2018). Raising the Bar on Teacher Quality: Assessing the Impact of Increasing Licensure Exam Cut-Scores. *Educational Policy*, *32*(7), 969-992. https://doi.org/10.1177/0895904816682315

Shuls, J. V., & Trivitt, J. R. (2015). Teacher effectiveness: An analysis of licensure screens. *Educational Policy*, *29*(4), 645-675. https://doi.org/10.1177/0895904813510777

Sockett, H. (2006). Character, rules, and relations. *Teacher dispositions: Building a teacher education framework of moral standards*, 9-25.

Tang, M., & Gómez-Bellengé, F. X. (2007, April). *Dimensionality and concurrent validity of the observation survey of early literacy achievement*. Paper presented at the annual meeting of the American Educational Research Association in Chicago, Illinois.

Tremmel, R. (1993). Zen and the art of reflective practice in teacher education. *Harvard Educational Review, 63*(4), 434-459. https://doi.org/10.17763/haer.63.4.m42704n778561176

Van Manen, M. (1991). *The tact of teaching: The meaning of pedagogical thoughtfulness*. Albany, NY: Suny Press.

Vandevoort, L. G., & Berliner, D. C. (2004). National Board Certified teachers and their students' achievement. *Education Policy Analysis Archives*, *12*(46), 1-117. https://doi.org/10.14507/epaa.v12n46.2004

Yanow, D., & Tsoukas, H. (2009). What is reflection-in-action? A phenomenological account. *Journal of management studies*, *46*(8), 1339-1364. https://doi.org/10.1111/j.1467-6486.2009.00859.x