Mediational Effects of Feedback Style on the Relation between Teachers' Depressive Symptoms and Classroom Quality in 3rd Grade

by

Leigh Ellen McLean

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Approved June 2015 by the
Graduate Supervisory Committee:

Carol McDonald Connor, Chair
Kathryn Lemery
Leah Doane
Kevin Grimm

ARIZONA STATE UNIVERSITY

August 2015

ABSTRACT

Described is a study investigating the feasibility and predictive value of the Teacher Feedback Coding System, a novel observational measure of teachers' feedback provided to students in third grade classrooms. This measure assessed individual feedback events across three domains: feedback type, level of specificity and affect of the teacher. Exploratory and confirmatory factor analysis revealed five factors indicating separate types of feedback: positive and negative academic-informative feedback, positive and negative behavioral-informative feedback, and an overall factor representing supportive feedback. Multilevel models revealed direct relations between teachers' negative academic-informative feedback and students' spring math achievement, as well as between teachers' negative behavioral-informative feedback and students' behavior patterns. Additionally, a fall math-by-feedback interaction was detected in the case of teachers' positive academic-informative feedback; students who began the year struggling in math benefitted from more of this type of feedback. Finally, teachers' feedback was investigated as a potential mediator in a previously established relation between teachers' self-reported depressive symptoms and the observed quality of the classroom environment. Partial mediation was detected in the case of teachers' positive academic-informative feedback, such that this type of feedback was accountable for a portion of the variance observed in the relation between teachers' depressive symptoms and the quality of the classroom environment.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my mentor, Dr. Carol Connor, for her careful, patient and dedicated guidance throughout my graduate education.

Additionally, I would like to acknowledge my parents, Syreece and Jim, my sisters, Annie and Mary, and my loving partner, Conor. Without the unwavering support of my family, this endeavor would not have been possible.

I would also like to thank the teachers, students and families who participated in this study.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

LIST OF SYMBOLS/NOMENCLATURE

Feedback Shorthand Guide

Feedback Type:

1. Ac: ‘Academic’
2. Bv: ‘Behavioral’
3. Tot: ‘Total’ (both academic and behavioral)

Feedback Specificity:

1. ID: ‘Identification’
2. Co: ‘Correction;
3. El: ‘Elaboration’
4. En: ‘Encouragement’
5. Ss: ‘Support Statement’

Collapsed Feedback Specificity

1. Info: ‘Informative’
2. Supp: ‘Supportive’

Feedback Affect:

1. Ex: ‘Enthusiastic/Exuberant’
2. Con: ‘Content”
3. Nu: ‘Neutral’
4. Sd: ‘Sad/Depressed’
5. An: ‘Angry/Frustrated’

Collapsed Feedback Affect:

1. Pos: ‘Positive’
2. Neg: ‘Negative

Examples:

- AcCoAn = ‘Academic Correction, Angry/Frustrated’
- BvRdCon: ‘Behavioral Redirection, Content’
- AcInfoPos: ‘Academic Informative, Positive’
- BvInfoNeg: ‘Behavioral Informative, Negative’

CHAPTER 1. INTRODUCTION & BACKGROUND LITERATURE

Statement of Purpose

Early academic achievement is one of the strongest predictors of later school and life success. Children who display patterns of low achievement in their early elementary years are at particularly high risk of remaining behind their peers throughout formal schooling and into adulthood (Duncan et al., 2007), with severe long-term implications including higher rates of school dropout and incarceration. Early literacy and mathematics skills are considered foundational to a child's academic success, and are strong indicators of school performance throughout the elementary years and into high school (NRP, 2000). Third grade is a particularly important year for both teachers and students, as in many states it is the first time that students' achievement is tested using high-stakes assessments. Recent large-scale evaluations of reading and mathematics achievement observed in young students across the U.S. paint a disconcerting picture, with only 34% of students reading at or above 'proficient' reading levels as of 2011, despite nationwide reading improvement initiatives that have been implemented across U.S. (NAEP 2011). Similar patterns have been observed in early mathematics achievement. The same 2011 NAEP report found that only 40% of US 4th graders display proficient levels in mathematics, with this percentage dropping to 23% for children living below the poverty line. The obvious deficits in early literacy and mathematics achievement seen in today's elementary students point to a need for research that clarifies what helps and/or hinders student learning. Given this, many efforts in the field of education center on investigating the classroom context and it's influence on

children's development, with a specific focus on instructional practices and teacher quality as primary sources of impact on student learning. Classroom observation in particular is an increasingly common tool for analyzing and improving instruction, with the intention of informing best practice within early classrooms, but many observation systems don't actually predict students' academic achievement outcomes. This gap in the field calls for investigation into what makes an observational tool successful in terms of accurately depicting what is going on inside early classrooms, as well as providing reliable information about how these early classrooms impact the academic growth of young students.

The availability of effective measurement tools in these settings is a crucial piece of the puzzle comprising how researchers and practitioners in the field can gain this valuable knowledge. Observing and analyzing teachers' feedback provided to students was of central interest in the present study, as it is a type of instructional move that is hypothesized to be both important to student learning and highly sensitive in nature, as it involves the direct judgment by the teacher of student learning attempts or behavioral actions. How this information is communicated to the student could have a wide range of implications, from the level of student learning that takes place within the feedback event to the impact of that event on the quality of the teacher-student relationship, with patterns of feedback over time potentially making significant contributions to the overall quality of the classroom environment. Thus, one purpose of this study was to test the feasibility and predictive value of the Teacher Feedback Coding System, a novel observational measure of teachers' academic and behavioral feedback to students, within early classrooms.

Additionally, this study sought to build on previous work by the investigator (McLean & Connor, 2015) that revealed a negative relation between teachers' self-reported depressive symptoms and the observed quality of the classroom environment in third grade classrooms. This relation was found to predict lower levels of mathematics growth in students who began the year struggling in math (an interaction effect), suggesting that students who were already at risk for school failure may have been particularly sensitive to their teachers' depressive symptoms and the contributions of these characteristics to classroom quality. The present study utilized data from the Teacher Feedback Coding System, which was applied to the same sample of students, to examine how these teachers' self-reported depressive symptoms related to their observed patterns of feedback provided to students, with overall classroom quality examined as the primary outcome of interest. It was predicted that teachers' feedback would act as a mediator in the already-established relation within this sample between teachers' depressive symptoms and classroom quality in early elementary settings.

Classroom Observation in Educational Research

The primary focus of this study was to apply a novel classroom observation tool, the Teacher Feedback Coding System, to a sample of early classrooms with the goal of assessing this tool's feasibility and potential usefulness as a measure of teacher feedback quality. In general, the use of classroom observation measures in educational research provides valuable information about the many processes that take place in the classroom, as it provides an ecologically valid approach to assessing teacher and children's behavior (Kontos & Keyes, 1999; Meisels & Atkins-Burnett, 2006; Neisworth & Bagnato, 2004).

These approaches allow researchers to examine how individuals in the classroom (both students and teachers) respond and calibrate to the situational demands of the day over time, rather than making judgments on single snapshots of observed or reported behavior (Volpe, DiPerna, Hintze, & Shapiro, 2005). In addition, many classroom observation systems provide us with a direct measure of the quality of students' classroom experiences, something that is absolutely essential for understanding how the classroom contributes to student learning (Barnett, Epstein, Friedman, Boyd, & Hustedt, 2008).

The usefulness of classroom observation tools has been demonstrated in multiple contexts. There is evidence that outcomes for students can be improved by identifying and improving aspects of the classroom environment, with observation being a crucial part of this process (Fraser,1998). For example, past studies have found that children with severe learning and behavior problems differ significantly from their typically-developing peers in their observable actions within the classroom (such as time on/off-task), suggesting that classroom observation can be used as a screening tool for such developmental difficulties (Forness & Esveldt, 1975). Additionally, observation is quickly becoming one of the primary methods used by schools to formally evaluate teachers' effectiveness in the classroom, as it provides school officials and administrators with a much more accurate view of a teacher's practices within the classroom than would be available by just assessing standardized tests or teacher value-added scores (Barnett, Epstein, Friedman, Boyd, & Hustedt, 2008). Having more reliable methods for evaluating teacher performance is crucial in today's educational contexts, as it has been recently shown that principals and school administrators are essentially incapable of

distinguishing effective teachers from ineffective teachers (Jacob & Lefgren, 2008, Strong, Gargani & Hacifazlioğlu 2011).

Another motivation for using direct observational methods to identify effective teaching strategies is that there have repeatedly been mixed or nonexistent associations between student outcomes and many of the standard proxies for teacher evaluation such as teacher's level of education or certification status (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Clotfelter, Ladd, & Vigdor, 2007; Early et al., 2007; Jepsen, 2005; Kane, Rockoff, & Staiger, 2008). While value-added modeling can be helpful in these situations (Rivkin, Hanushek, & Kain, 2005; Rockoff, 2004), this method provides little to no insight into how teachers are impacting student gains, and also fails to provide professional guidance to teachers about improving their practices. These gaps in teacher evaluation can be filled through the use of high quality and reliable measures of classroom observation (Burchinal et al., 2008; Connor, Son, Hindman, & Morrison, 2005; Howes et al., 2008; Mashburn et al., 2008; NICHD Early Child Care Research Network, 2002; Pianta, La Paro, Payne, Cox, & Bradley, 2002).

Finally, classroom observation is frequently used in professional development contexts to inform and support teachers' knowledge of effective instruction, student support, etc. (Pianta, Mashburn, Downer, Hamre, & Justice, 2008). The strong consensus within educational communities that teachers' professional development is crucial to their ongoing success in the classroom (Caspary, 2002) supports the importance of having effective strategies to help teachers with their continued development. Classroom observation measures have the potential to play a significant role in professional development by providing standardized systems for documenting teachers' strengths and

weaknesses in order to inform practitioners about whether or not certain professional development tactics are working (Pianta, 2003).

In sum, observation has become one of the primary methods used by both researchers and schools to evaluate many classroom processes, particularly teachers' effectiveness in the classroom, as it provides a view of classroom practices that go beyond standardized tests or value-added scores. However, while these more in depth and reliable methods for evaluating these classroom processes are appealing in their contributions of knowledge to the field (Barnett, Epstein, Friedman, Boyd, & Hustedt, 2008), many commonly-used systems of observation fall short in their attempts to accurately capture what is going on in the classroom. Recent studies have shown that principals and school administrators are essentially incapable of distinguishing effective from ineffective teachers through direct observation (Jacob & Lefgren, 2008; Strong, Gargani & Hacifazlioglu, 2011). I conjecture that one reason for this may be that many of the current observational tools and training systems do not consider the individual and repeated interactions that take place between teachers and their students during classroom instruction, which is a central feature of the Teacher Feedback Coding System.

History of Classroom Observation

The use of observation as a measurement tool in the classroom has been a staple of educational research spanning the past three decades (Gage & Needels, 1989), and continues to be a valuable source of information regarding how the classroom environment influences student outcomes. Multiple types of observation systems have been developed for use in the classroom, many focusing on specific teacher behaviors (Cochran-Smith & Lytle, 1990; Smith, Waller, & Waller, 1982) and their impacts on

students' learning (Brophy, 1986). While varying approaches have been taken to using observation in the classroom, most have yielded outcomes that strongly suggest observation techniques can lead to the improvement of a wide range of student outcomes through modifying aspects of the classroom environment based on the findings of observation systems (Fraser, 1998).

A clear evolution of trends and tactics in classroom observation systems can be observed throughout the span of their use in educational research. The primary focus of classroom observation at it's beginning (in the early 1970's) was on classroom processes surrounding the teacher (Cochran-Smith & Lytle, 1990; Smith, Waller, & Waller, 1982). More specifically, these early systems looked closely at teacher practices and behaviors and how they related directly to student academic outcomes, observing the classroom as a whole and generating a global judgment of the teachers' performance (Good & Brophy, 1970). These systems were often simple in nature, utilizing frequency counts to link teacher behaviors to student outcomes. For example, Brophy & Good (1986) used frequency counts to investigate the relation between overall quantity of teaching and student learning. He also used frequency counts to examine teacher's allocation of time to specific educational subjects (Brophy & Evertson, 1976). Other simple, frequency-centered approaches to classroom observation from this time period looked at variables such as the number of pages presented to students within a curriculum (Borg, 1979; Good, Grouws, & Beckerman, 1978). These early approaches were important in establishing a solid foundation on which to build more complex systems of classroom observation, and paved the way for later researchers to create more elaborate observation schemes.

A shift from quantitative (frequency-centered) to qualitative methods in classroom observation in the early to late 1990s brought on an increased focus on using ethnographic approaches to gain in-depth information about teacher practices and student experiences, and the interaction between the two (Gudmundsdottir, 1991, Kontos & Keyes, 1999). These approaches have been an important step towards the development of new theories and constructs relating to the dynamic interactions that take place within the classroom. While results from these qualitative observation systems provided in-depth, highly descriptive data, they were difficult to generalize to larger populations, and therefore did not yield many concrete findings. Instead, they were used to drive the development of hypotheses regarding the specific mechanisms of classroom impacts on student learning.

More recent years have seen a strong shift towards the development of classroom observation systems that display high validity and reliability, in addition to providing large amounts of high-quality information to researchers (e.g., Cameron, Connor, & Morrison, 2005; Pianta et al., 2007). This increased focus on standardization has driven the field of classroom observation research in the direction of practitioner and program quality assessment. Many measures have been implemented statewide, and are relied upon for assessing the effectiveness of educational programs. In fact, 38 of 50 state-funded pre-kindergarten initiatives in place today have policies requiring observational monitoring of educational sites (Greenburg, 2004, Barnett et al., 2008). The standardization of classroom observation systems has allowed researchers to collect in-depth, high-quality data on the impacts of the classroom environment on children's learning that can be reliably generalized to larger populations.

Development and Use of Classroom Observation Measures

While the many practical uses of classroom observation tools are well documented, the truth remains that in order for observations to be useful they must utilize standardized observation protocols that minimize measurement error and promote generalizability (Pianta & Hamre 2009). While the reliability of classroom observation methods has been improving steadily over the past three decades, there are still some common problems in the development and application of observation measures that researchers warn about. In their discussion of the reliability and methods of analysis of classroom observation measures, Meyer, Cash & Mashburn (2011) address many of the common issues faced by researchers using observation tools. Firstly, they warn that variability observed in student achievement scores could very well reflect true changes in the quality of teacher/student interactions, or this variability could simply be due to measurement error. While this is true of any measure, observational or not, it is an important factor to keep in mind when trying to clearly define an observation scheme. Authors recommend that classroom observation measures should be 'designed to minimize measurement error while allowing measurable changes in the construct of interest'.

Other common sources of error seen in the use of observation tools are variations in levels of severity or leniency across observers, the selection of scoring criteria, and variation in the specific occasion of observation. Since classrooms are complex and dynamic systems in and of themselves, student-teacher interactions will vary throughout the day, week, months, and across the entire school year. This combined with the fact that one observer will inevitably judge the same situation slightly different than another

means that reliability between raters can be very difficult to achieve (McGaw, Wardrop, & Bunda 1972). Clear definitions of the goals and specifications of the observation's coding scheme, combined with meticulous training of observers that includes ample opportunity for practice can help with these issues, however the harsh reality is that these biases will always be a threat to the validity of classroom observation tools.

Review of Current Classroom Observation Measures

Many classroom observation measures exist today that are commonly used in educational research. These measures address multiple aspects of the classroom environment, from interactions between teachers and students (Burchinal et al., 2008; Pianta et al., 2008, Palermo, Hanish, Martin, Fabes, & Reiser, 2007), to teachers' educational background and certification (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Clotfelter, Ladd, & Vigdor, 2007; Early et al., 2007; Jepsen, 2005), and even the contribution of teachers' personal characteristics, such as stress level and self-efficacy, to student learning experiences (Mashburn, Hamre, Downer & Pianta, 2006; 2008; Li-Grining et al., 2010). While the majority of observation systems in use today primarily focus on a single aspect of the classroom environment, many systems can be thought of as 'hybrid' systems that measure different types of data at the same time (instructional tactics, emotional climate, interactions etc.).

Observational measures can be applied to the classroom environment in a variety of ways, from secondary analysis of videotaped classroom sessions to live observations that take place in the actual classroom environment. Coding of these systems can be done using computer technology, handwritten notes or the use of scales (or some combination of these). The wide variety of tactics employed by these systems has

resulted in researchers having available to them many options for classroom measurement that they can manipulate to fit the needs of the individual study. Following I will present a review of some of the more commonly used observational measures in educational research today.

Instruction-Centered Observation Systems

Instruction-centered classroom observation measures are perhaps the most widely used types of classroom assessment tools in educational research today. Such measures primarily focus on the instructional tactics utilized by the teacher within the classroom, with student learning as the outcome of interest. A wealth of research in the past two decades has illustrated the importance of effective instruction during the early grades, especially in the area of literacy (Connor et al., 2013; NICHD, 2000; Pianta, Belsky, Houts, Morrison, & NICHD-ECCRN, 2007; Snow, 2001). Three observational systems that provide solid examples of how observation can be utilized to asses classroom instruction are the Instructional Content Emphasis (ICE) system (Edmonds & Briggs, 2003), the Video Assessment for Teaching (VAT) system (Carlisle et al., 2011), and the Individualizing Student Instruction (ISI) system (Connor et al., 2009). These three systems take different approaches to assessing very similar classroom variables, namely teachers' implementation of instruction and students' grouping and participation within classroom activities. All are able to provide valuable information that has proven highly predictive in empirical studies.

The ICE observation system originally developed by Edmonds & Briggs provides a detailed framework for assessing instruction in the classroom. Six distinct categories of instruction are identified in this system, with an emphasis on literacy (oral text reading,

for example). There is also a secondary focus on other instructional aspects such as the materials used within lessons and the level of student engagement during activities. Grouping is categorized in this system, tracking how much time students spend in various types of groups within classroom learning opportunities (whole-class, small-group, etc.), and teachers' level of instructional involvement during these times. Studies utilizing this system have found that, in early elementary settings, teachers engage in whole-class, direct instruction the most. It has also been found that teachers emphasize reading comprehension in their instruction above and beyond other instructional categories such as phonological awareness or alphabetic knowledge, and that teachers' instructional techniques were directly related to student literacy outcomes (Vasadey and Sanders, 2008).

A more recent instructional observation tool developed by Carlisle and colleagues in 2011, the VAT, provides a more pedagogical approach to lesson observations. The intended purpose of the lesson is taken into account along with the specific techniques employed by the teacher and students' level of engagement. This measure also examines teachers' personal characteristics such as subject knowledge, education level and self-reported practices. Studies using this measure have found that time spent by teachers engaged in direct instruction with students, as well as higher levels of teacher support for student learning, are positively related to student academic outcomes. Additionally, it has been found that the longer individual lessons last, the more likely teachers are to participate in high quality instructional tactics (Carlisle et al., 2011).

The final classroom observation system focusing on instruction is the Individualizing Student Instruction system developed by Connor and colleagues in 2009.

This system is one of the most in-depth observation schemes used today to assess teachers' instruction during student learning opportunities. This system takes observation to the level of the activity for each student, categorizing observations across three dimensions: context (i.e. small group, whole-class, etc.), teacher/child interaction (i.e. who is managing the instruction, teacher or student?), and content (i.e. what type of lesson is being taught?). Any activity lasting longer than 15 seconds is documented based on all three of these dimensions, providing researchers with incredibly detailed data regarding what exactly is taking place within the classroom for each target child. While this system does not take into account quality of instruction, it provides more information than many other schemes as it incorporates well over 200 instructional codes, which can provide information about what is taking place in the classroom. Studies using this observational coding system have produced a wide variety of interesting results, ranging from the different types of instruction that tend to yield more positive student outcomes (Connor et al., 2011; Connor et al., 2013) to the different types and amounts of instruction received by different groups of students based on specific student characteristics such as race, gender, and exhibited behavior patterns (McLean & Connor, under review; Tani & Connor, under review).

Child-Centered Observational Measures

A second type of classroom observation measure commonly used in educational research is the child-centered observation measure. These measures focus primarily on the behavior and activities of the child in the classroom, without documenting or judging teacher's instructional techniques or other aspects of the classroom environment. One commonly used child-centered measure is the Behavioral Observation of Students in

Schools (BOSS; Shapiro & Heick, 2004). The BOSS measures student activity in the classroom in order to make highly informed inferences about students' individual behavior patterns. Classroom behavior is divided into on- or off-task, and these behaviors are further categorized to provide more specific information about how the child is spending their time during learning opportunities. This observation takes place inside the classroom, with the observer making detailed notes on the target child's activities every 15 seconds. While this measure does have a teacher instruction component, the primary focus is given to the behaviors of the child. Studies utilizing the BOSS have found that it is able to reliably discriminate between children with ADHD and their typically-developing peers, as well as between children with different levels of academic performance (DuPaul et al., 2004, Ota & DuPaul 2002).

Another widely used measure of child behavior in school settings is the Direct Observation Form (DOF; Achenbach, 1986), which is a part of the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach & Rescorla, 2001). A strength of the DOF is that it can be applied across multiple school settings, from the classroom to the cafeteria/lunchroom to the playground. While each observation is brief, about 10 minutes or so per child, each child receives a total of three to six observations, which are combined to provide a more stable estimate of child behavior. Observers use the 10-minute observation to inform their ratings of students' behavior across 97 target items. Items are scored on a 4-point likert scale and correspond highly with items on the Child Behavior Checklist (CBL), another popular measure of child behavior patterns (Achenbach & Rescorla, 2001). The DOF has been found to be highly discriminating, and is able to identify children with problem behaviors within samples that are matched

for age, grade and race (Reed & Edelbrock, 1983). More specifically, the sensitivity of the on-task, nervous/obsessive, internalizing, and depressed subscales of the DOF have been demonstrated to accurately identify children with these difficulties, resulting in their early identification for emotion-based assistance (McConaughy et al., 1998, 1999).

Teacher Feedback

The feedback provided by teachers to their students can be a powerful instructional tool to enhance student learning (Altwell 1998, Bratcher 2004; Hattie, 2009). While this has been supported through educational research, there is a lack of consistency across the field when it comes to defining the term 'feedback' (Van de Ridder, Stokking, McGaghie, & Ten Cate, 2008), as well as distinguishing high from low-quality feedback (Nicol & Macfarlane-Dick, 2006). Some past investigators have conceptualized feedback as a purely academic endeavor, relating it only to teachers' direct responses to student academic attempts with the express goal of helping the student improve their understanding and/or performance (Hattie, 1998). Studies defining feedback as more academic tend to focus on topics such as whether it is written or verbal (Hillerich 1985, Kulhavy & Stock, 1989; Smith 1989), effort or ability-focused (Craven et al., 1991; Dohrn & Bryan, 1994; Marsh, 1990), or the degree to which academic feedback is elaborative, as opposed to just providing the correct answer (Mandernach, 2005; Pridemore & Klein, 1995; Smits, Boon, Sluijsmans, & van Gog, 2008). Others have conceptualized feedback as a more emotion/relationship-driven process, with things such as teachers' positive and negative affect and levels of praise for student behavior being the main factors taken into account (Brophy, 1985, Hoyenga & Hoyenga, 1993).

Student outcomes of interest also differ greatly between these viewpoints, with academic feedback studies focusing on academic outcomes and behavioral feedback studies focusing on emotion-related outcomes. The Teacher Feedback Coding System incorporates both points of view, as both have demonstrated to be predictive in empirical studies. Both academic and behavioral feedback interactions will be observed and assessed for level of elaboration and affect of the teacher.

Academic Feedback

Feedback on students' academic performance is important to the learning process because it allows students to resolve discrepancies between actual and desired knowledge (Black & Wiliam, 1998), with content of the feedback message often being considered the most important part of the process. This content enables student recipients to correct informational errors, maintain correct responses, and in many cases can enhance learning through elaboration on a specific topic (Butler, Karpicke, & Roediger, 2008; Pashler, Cepeda, Wixted, & Rohrer, 2005). Given the suggested importance of this process, a primary goal observed across academic feedback studies is to identify what specific components make up a successful feedback interaction. It is generally agreed upon that, at the most basic level, a feedback message should include a judgment of the correctness of the student's initial response, paired with the correct answer if required (Pashler et al., 2005; Phye & Sanders, 1994; Whyte, Karolick, Neilsen, Elder, & Hawley, 1995). Within this study I predict feedback that is more elaborative in nature will prove more helpful to students in terms of learning outcomes.

Many argue that elaborative feedback, or feedback that provides more in-depth information to the student regarding why and how their information was correct or

incorrect, is helpful in improving students' understanding of the topic at hand (Corbett, Koedinger, & Anderson, 1997). Surprisingly, past studies that have compared elaborative feedback methods with basic 'correct answer only' methods have seen little to no difference in terms of student learning outcomes (Mandernach, 2005; Pridemore & Klein, 1995; Smits, Boon, Sluijsmans, & van Gog, 2008), with some evidence showing that corrective feedback may actually be more effective than elaborative for English language learners in their early acquisition of English (Kao, 2013). These findings point to the possibility that correct-answer and elaborative feedback messages may be differentially effective for varying groups of students. Butler, Godbole & Marsh (2013) investigated the differential outcomes of providing students with correct-answer feedback, elaborative feedback or no feedback and found that correct-answer and explanation feedback led to equivalent performance on repeated questions, but explanation feedback produced superior performance on novel inference questions. This finding suggests that the outcome variables used in academic feedback studies may play a large role in whether or not effects are seen. Overall, in reviewing the literature surrounding the characteristics and effectiveness of academic feedback tactics employed in the early elementary classroom, it becomes clear that there is still much to be discovered about which delivery methods are the most effective, and for which groups of students.

Behavioral Feedback

Research that focuses on behavioral feedback takes a very different approach. Most of these studies assess feedback messages solely within the context of student-teacher interactions, placing emphasis on more emotional aspects of the interaction such

as whether the feedback was positive or negative in nature (Brophy, 1985, Hoyenga & Hoyenga, 1993). They also utilize outcomes that have less to do with academics and more to do with students' self-concept (Chen et al., 2011), classroom/school satisfaction (Burnett, 2002), peer relations (White & Kistner, 1992), and the teacher-student relationship (Burnett, 2002).

Brophy (1985) observed that 5th grade boys received more negative teacher feedback concerning failure to follow directions, whereas girls received more positive feedback concerning compliance. Expanding on this study, Morgan (2001) randomly assigned 5th grade students to receive one of 5 feedback patterns, all of which primarily focused on positive competence-related feedback but varied in their patterns pertaining to the previously mentioned differences in gender. Students who received the more typically 'male' (negative) feedback patterns showed decreased interest in classroom activities, decreased perceived competence and less liking for the teacher. They also reported less willingness to work with the teacher. This is a perfect example of a system for feedback assessment that is purely behavior-focused, with little attention being given to academics.

Other studies delve deeper into the categorization of behavior-centered feedback, one prime example being the differentiation between ability feedback and effort feedback (Dohrn & Bryan, 1994). Ability feedback is feedback that focuses on a students' skills or intelligence ("good job, you're very smart"), while effort feedback focuses on the actual effort being applied to a task, rather than the skills behind that effort ("I can tell you're trying very hard today"). Mueller & Dwek's (1998) research using these conceptualizations of feedback found that ability feedback was more strongly associated

with students' interest in performance-type tasks (tasks that have a clear goal and reward), while effort feedback was more related to 'interest in learning' tasks. This study also revealed that students who received more ability feedback tended to lie more often about their school performance, while students who received more effort feedback were more truthful.

When these studies are taken into consideration along with the wide range of studies that exist examining academically focused feedback, it becomes clear just how varied the many conceptualizations of 'feedback' are across the field of early education. While the majority of these studies provide interesting results that are highly applicable to many of today's nationwide educational goals, the fact remains that we have yet to clearly define the construct of 'feedback' and apply this definition to a set of student outcomes that are both behavioral and academic in nature. While results of efforts to study the widely varying conceptualizations of teachers' feedback have been mixed, the general trend does seem to suggest that high quality feedback, even across these varying definitions, is potentially beneficial to students. However further study into how to define feedback and effectively measure it is warranted, all with the goal of revealing how this aspect of instructional interactions operates within the classroom system. Thus, the central goal of study 1 is to create a novel measure of teachers' feedback that utilizes both academic and behavioral indicators in it's definition of 'feedback', as both viewpoints have been shown to be highly predictive of student outcomes. It is hypothesized that behavioral and academic feedback will be highly related to each other (e.g. teachers will display similar patterns for both types of feedback), and that teachers' combined feedback quality will predict student outcomes, both academic and emotional.

Classroom Quality

The quality of the classroom-learning environment is strongly related to student development in early elementary school (Bronfenbrenner & Morris, 2006; Cameron, Connor & Morrison, 2005). Students in high-quality classrooms have been found to display more social/emotional and academic growth when compared to their counterparts in low-quality classrooms (Connor, Son, Hindman & Morrison, 2005; Hamre & Pianta 2007; Ponitz, Rimm-Kaufman, Grimm & Curby 2009). Further, it has been demonstrated that a high level of classroom quality helps buffer against the negative academic and social outcomes associated with high-risk student characteristics such as difficult temperament, low self-regulation, and low SES (Curby, Rudasill, Edwards & Perez-Edgar 2011; Hamre & Pianta, 2005).

Within this study, high-quality classrooms are conceptualized as learning environments that involve purposeful organizational, instructional, and management techniques facilitated by the teacher that aid in successful student learning. Organization includes both the physical characteristics of a classroom as well as the teacher actions used to promote efficient use of time (e.g., explaining upcoming activities, allowing opportunities for students to rehearse the behaviors necessary to complete assigned academic tasks; Pressley et al., 2001). Management refers to the teachers' use of intentional proactive and reactive actions in order to maintain a successful learning environment. This could include discipline, assignment of work groups, and strategic breaks during which students moved to expend excess energy (i.e., wiggle breaks). Lastly, instruction is conceptualized as the purposeful relaying of academic information from the teacher to the students.

These dimensions of classroom quality have been found to be individually predictive of student outcomes. For example, Bohn, Roehrig & Pressley (2004) found that when teachers demonstrated high levels of organization, students generally exhibited increased behavioral regulation skills. Further, it has been found that teachers who effectively establish routines and implement rules are observed to experience fewer student disruptions that may interrupt classroom learning (Borko & Niles, 1987, Epstein, Atkins, Cullinan, Kutash & Weaver, 2008). Purposeful organization and instruction on the part of the teacher has been found to be a particularly important part of the learning environment that impacts student outcomes (Cameron, Connor, Morrison & Jewkes, 2008; Connor et al., 2010). Additionally, there appears to be dynamic interplay among these dimensions (Cameron, Connor, & Morrison, 2005; Eccles & Gootman, 2002). Pressley et al. (2001) investigated the impacts of multiple classroom features concurrently on first grade students' literacy achievement and found that a combination of effective classroom management, purposeful support for student self-regulation, balanced and developmentally appropriate instruction, and tactful interweaving of literacy lessons throughout the day were highly and positively predictive of student achievement. Mashburn et al. (2008) found that teachers' instructional and emotional interactions with students were predictive of both academic and emotional outcomes. Important to note, teachers who were rated by observers as high on one of these dimensions tended to be highly rated on the others as well, further demonstrating the likelihood that these dimensions, when considered together, are indicative of the construct of classroom quality (Hamre & Pianta, 2007).

Since teachers are responsible for the implementation of the classroom factors discussed, successful implementation of these systems is likely important for supporting student learning. Connor et al (2010; 2014) found that teachers' level of success in implementing effective classroom management, organization, and instruction, combined with time spent by the student in meaningful literacy instruction may influence students' literacy gains in early elementary school. Given the wealth of past research pointing to the importance of students' mastery of academic skills early, it becomes clear that a high-quality classroom environment may be one of the key factors that supports growth in literacy and mathematics.

This study seeks to examine relations between data collected using the Teacher Feedback Coding System and the quality of the classroom environment. Expanding the research of teacher/classroom quality by taking into account the impact of teachers' feedback could provide education professionals with more knowledge on the impact that these repeated individual interaction patterns have in determining the overall quality of the classroom.

Teacher Characteristics

Recent research has identified the impact that teachers' personal characteristics can have on multiple aspects of the learning environment to which children are exposed (Mashburn, Hamre, Downer & Pianta, 2006, Hamre, Pianta, Downer & Mashburn, 2008). Teachers' feelings of self-efficacy and depressive symptoms have been found to contribute significantly to amounts of conflict observed within the classroom (Hamre et al., 2008). Factors such as job burnout and low amounts of perceived control have been

found to negatively impact over teachers' abilities to maintain positive and responsive in their instruction and interaction with young students (Chang, 2009). Foundational to the proposed second study, McLean & Connor (2015) found that teachers with more depressive symptoms were less able to maintain high-quality classrooms, and student mathematics performance suffered as a result. This impact on achievement was the strongest for students who began the year with weaker skills, suggesting that children at risk for academic failure are particularly sensitive to the influence of their teachers' depressive symptoms within the classroom environment. Further investigation into which characteristics aid or impede a teachers' ability to create a positive classroom learning environment is a promising direction that research in this area could take, as it has the potential to inform professional development and improve the experiences and outcomes of young students.

Teachers' Depressive Symptoms

Many of the personal challenges commonly faced by educators, such as low self-esteem and self-efficacy, feeling out of control or burnt out, difficulty managing one's stress level and emotions etc. are all strongly correlated with clinical depression (American Psychiatric Association, 2001). Clinical depression, or Major Depressive Disorder is recognized by the DSM V as a mental disorder and is most generally characterized by overall low mood, low self-esteem, and a loss of interest or pleasure in normally enjoyable activities (among quite a few other symptoms). It is considered a disabling condition that adversely affects all aspects of a persons' life, most importantly to note here, their professional life. Diagnosis of clinical depression is based primarily on a combination of self-reporting with the use of a standardized screening tool such as the

Major Depression Inventory or the Beck Depression Inventory (BDI), and evaluation by a trained clinical professional. A certain number of symptoms must be present in the individual for a specified, ongoing amount of time in order to qualify a diagnosis (American Psychiatric Association, 2001).

A large body of research exists that describes the effects of clinical depression on the individual, as well as on those closest to the individual. The adverse effects of maternal depression on child development have been especially well described. It has been found that children of depressed mothers generally exhibit poorer academic performance and social competence, as well as higher instances of behavioral problems than do children with non-depressed mothers (Murray & Cooper 1997; Supplee et. al 2004). High quality mother-child relationships have been found to act as a buffer for children against the effects of multiple risk factors, most notably poverty, which often lead to lower levels of academic achievement (NICHD ECCRN, 2002). Because depressed mothers are at high risk for developing low-quality relationships with their children, these buffering effects could potentially be unavailable for the children of depressed mothers.

While there is extensive research surrounding maternal depression, relatively little exist that examines the effects depressive symptoms in teachers on the quality of the classroom-learning environment, and on students' academic and emotional development (with the exception of McLean & Connor, 2015). Since, as previously mentioned, teachers play an influential role in the educational experiences of their students, it would stand to reason that patterns similar to those seen in depressed mothers and their children may emerge for teacher depression and student academic growth. Interestingly,

O'Connor & McCartney (2007) found that a strong, positive teacher-child relationship acted as a buffer against the negative effects of insecure patterns of attachment between children and their mothers. Since depressed mothers are at particularly high risk of developing negative/insecure attachment patterns with their children, this study points at the potential importance of teachers' emotional well-being in the classroom, especially for students who may be experiencing a home environment that puts them at risk of academic failure.

Two of the most common depressive characteristics experienced by teachers are burnout and job-related stress. This makes sense, as teaching has been consistently identified as one of the most stressful of occupations (Johnson et al., 2005; Travers, 2001). Stress and burnout are most often conceptualized concurrently as a significantly interconnected pair of characteristics (Maslach, Schaufeli, & Lieter, 2001). Past studies have found that ongoing job-related stress and feelings of burnout can take significant tolls on both the professional and personal lives of teachers. Physical and mental health suffer in terms of depression, anxiety, cardiovascular disease and high blood pressure (Dimsdale, 2008; Gunnar & Quevedo, 2007; Maslach Schaufeli, & Lieter, 2001; McEwen, 2008), while professional effectiveness suffers through increased absenteeism and diminished capacity to engage with students and apply high-quality instruction (Darr & Johns, 2008; Roeser, Skinner, Beers & Jennings, 2012). In perhaps one of the first studies of it's kind, Whitaker, Becker, Herman & Gooze (2013) found that reports of poor mental health were more prevalent among female head start teachers compared to US women with similar socio-demographic characteristics in other professions. This study provides foundational evidence for the high-risk nature of teaching as an

occupation in relation to practitioners' emotional well being. All things considered, the negative effects of teachers' stress and burnout (both depressive symptoms) on student achievement, perhaps through teachers' absenteeism or lack of effective instruction (Miller, Murnane, & Willett, 2007) are highly likely. In fact, past studies have strongly suggested that this relationship does indeed exist (Briner & Dewberry, 2007; Jennings & Greenberg, 2009).

The importance of considering teachers' psychological well-being has been well documented (Hamre & Pianta, 2004). The presence of depressive symptoms in teachers has the potential to affect multiple aspects of their professional performance, most importantly in this study their contribution to the quality of the classroom environment through teachers' daily interactions with students. In fact, Li-Grining et al. (2010) found a direct relation between personal stress in teachers and the quality of their classroom behavior management, in particular their ability to foster positive teacher-student and student-student interactions. Raver and colleagues (2008) found that teachers who were more successful at regulating their emotions were much more likely to "catch" and reinforce positive behavior in their students. Results of this same study also indicated that teachers who were unable to regulate their emotions often had a difficult time focusing attention on the multiple aspects of a classroom during times of conflict and discipline, leading to a more chaotic classroom environment and an inability to attend to the activity and behavior of the class as a whole.

Intervention researchers consistently highlight the important role of teachers' psychological characteristics in the successful implementation of in-class interventions (Baker et al. 2009). This is especially true for interventions targeting child socio-

emotional development, since teachers have the potential to serve as either positive or negative role models of social behavior (Jennings and Greenburg 2009). Although not directly tested, this research implies that a teacher exhibiting more depressive symptoms may not be able to implement interventions or curriculum changes as effectively as their colleagues who are more emotionally stable.

The studies cited so far all indirectly support our theory that the presence of a constellation of depressive symptoms in teachers, whether related directly to occupational stress or to other aspects of their lives (e.g., divorce, financial problems), may negatively affect their students' achievement, functioning through factors such as the teacher-student relationship or teachers' ability to effectively implement instruction. Teachers' warmth and responsiveness, discipline techniques, implementation of instruction, and organization style in the classroom are all factors that impact the classroom-learning environment (Connor, Son, Hindman & Morrison 2005; Connor, 2012). If a teacher is not able to maintain success in all of these areas simultaneously, it is very possible that students could suffer negative consequences. While many of the characteristics associated with clinical depression have been studied individually in the context of the school environment and have been found to be significantly interconnected, little work has been done combining such characteristics into one overall variable of 'depression'. It is quite possible that teachers experiencing a combination of these symptoms could in fact be suffering from undertreated or even undiagnosed clinical depression. Expanding the research of teacher quality in this direction could provide education professionals with more knowledge on the potential adverse effects that depression in teachers can have on their young students' development. Providing empirical evidence on this topic

could act as a catalyst for those teachers needing diagnosis and treatment for depression to seek it, as well as raise awareness on the part of school administrators of the importance of providing emotional support for teachers when needed, not only for their own personal benefit but for the educational benefits of their students.

Associations among Variables of Interest

Teachers' depressive symptoms have the potential to impact almost every aspect of the classroom-learning environment. Mashburn, Hamre, Downer & Pianta (2006), found direct links between teacher's psychological characteristics and their personal evaluations of levels of conflict with children. This finding clearly illustrates the link between teachers' psychological state and their perceptions of the teacher-student relationship. Teachers who exhibited more maladaptive psychological characteristics were more likely to rate their relationships with students as more hostile and characterized by conflict. While directionality is not determined here (it could be that these teachers do indeed have more hostile relationships, which exacerbates already-existing psychological problems). Further expanding on this line of research, Hamre, Pianta, Downer & Mashburn (2008) found that, while the majority of conflict perceived by teachers could indeed be attributed to levels of problem behaviors in children, teachers' depressive symptoms were highly predictive of their reports of conflict with students.

Past research has suggested that many of the daily stressors of teaching that have been shown to lead to burnout also have been found to erode teacher's ability to maintain a positive and responsive style of instruction (Chang, 2009; Curbow, 1990). Li-Grining et

al. (2010) expanded on this finding by investigating personal stressors in addition to professional stressors and their relations to teachers' performance in the classroom. They found that personal stressors (for example, being the sole source of income in a family) were associated with teachers' ability to effectively manage student behavior and initiate positive social interactions with students. Teacher's depressive symptoms and patterns of feedback provided to students may be interwoven within the context of the classroom environment, with student outcomes varying based on the successful or unsuccessful interactions between these variables. The proposed study seeks to elucidate the specific ways teachers' self-reported depressive symptoms and the nature of their feedback to students interact within the classroom environment to impact students' early educational experiences.

Theoretical Framework

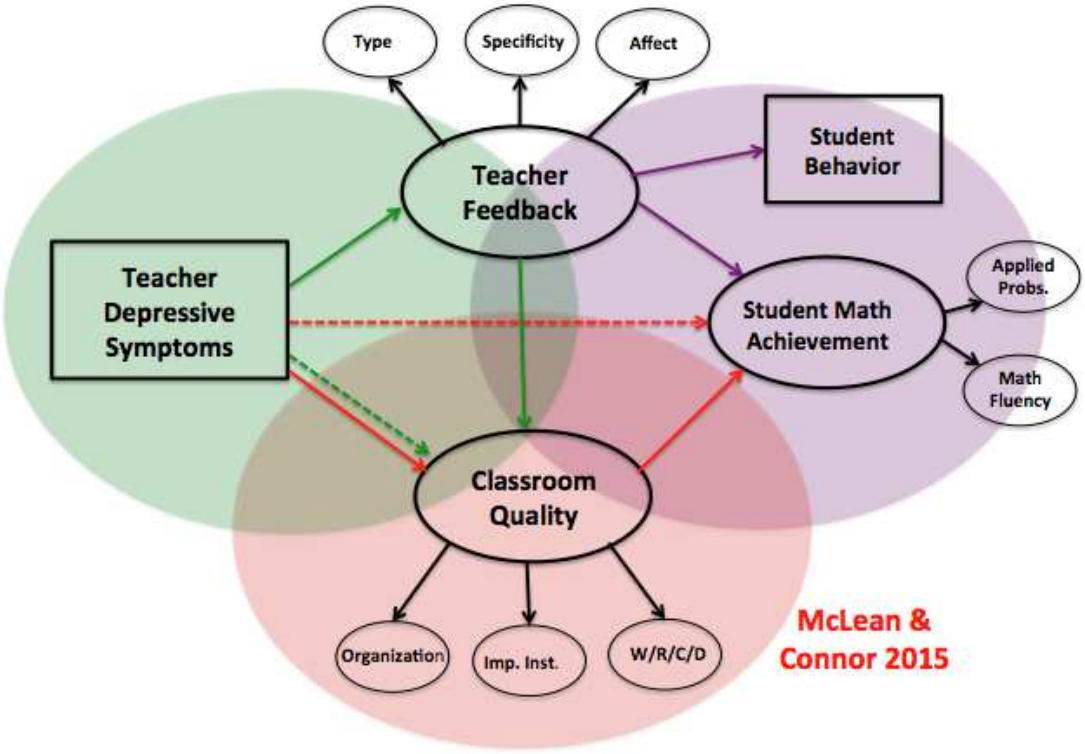
The Bio-Ecological Model of child development (Bronfenbrenner & Morris, 2006) defines significant developmental systems, both at the intra- and extra-individual levels, and characterized by interdependence and organization, as functioning systems of influence on developmental processes. Building on this, Dynamic Systems theories (Yoshikawa & Hsueh, 2001) assert that factors within these systems can interact to cause changes to developmental trajectories. Using this framework, I conceptualize the classroom as a complex microsystem of influence on student development, in which multiple factors interact synergistically to impact students' learning experiences and outcomes. This framework has been supported in empirical research, as strong connections have consistently been documented between classroom quality, teacher and

student characteristics, and student achievement in early elementary settings (NICHD ECCRN 2002, Connor et al., 2005, Connor et al., 2010; McLean & Connor 2015).

Based on this theoretical foundation, I first predict that the interactions between teachers and their students during feedback events will be directly indicative of student learning outcomes in mathematics. Further, I predict that the characteristics that teachers bring with them into the classroom (in this case, depressive symptoms), have strong potential to influence such operations within the classroom. As teachers are responsible for the implementation of multiple classroom factors that determine classroom quality, a logical conclusion is that the presence of depressive symptoms may hinder their ability to successfully implement the necessary systems to create an environment that is conducive to student learning. More specifically, one system that I predict that may be negatively influenced by teachers' depression is that of daily teacher-student interaction patterns, as observed through teachers' provision of feedback to students. The potentially negative interplay between teachers' depressive symptoms and their patterns of feedback provided to students may influence the overall quality of the classroom environment, with implications for student outcomes across multiple domains.

Figure 1

Theoretical model. Classroom quality mediates the relation between teachers' depressive symptoms and student math achievement (red; McLean & Connor 2015). The present study investigates direct effects of feedback on student outcomes (purple) and mediation effects of teacher feedback on the relation between teachers' depressive symptoms and classroom quality (green).



CHAPTER 2. METHODOLOGY

Study Aims & Hypotheses

The proposed study was guided by two primary aims; first to develop and apply a novel measure of teachers' feedback in 3rd grade classrooms and investigate relations between teachers' feedback and student math and behavior outcomes. Second, to test the predictive validity of this measure within a study that investigates the relations among teachers' self-reported depressive symptoms, their feedback patterns, and the quality of the classroom- learning environment. The first aim was be guided by the following research questions:

1) What factor structure best represents the data collected using the Teacher Feedback Coding System? I hypothesized that factors revealed in the data would be grouped by feedback type, specificity and affect during feedback events. Specifically I predicted that factor analysis would show differentiation in the data between academic and behavioral feedback, between levels of feedback specificity (simplistic vs. elaborative vs. supportive feedback) and between positive, neutral and negative affect during feedback events.
2) Do the primary factors revealed in RQ1 predict students' behavior patterns and mathematics achievement in 3rd grade? I hypothesized that teachers' feedback patterns would predict students' mathematics achievement and classroom behavior in the following ways;
    a. Academic feedback patterns identified as more elaborative and

positive would predict stronger student performance in mathematics, while feedback that is more simplistic and negative would predict weaker performance.

b. Behavioral feedback patterns identified as more elaborative and positive would predict fewer externalizing behavior problems, while feedback that is more simplistic and negative would be related to more problems.

The second aim used information gained in aim 1 to address the following research questions:

3) To what extent is there a relation between teachers' self-reported depressive symptoms and the nature of the feedback they provide to their students? I hypothesized that as teachers' self-reported depressive symptoms increased, the quality of the feedback they provide to students will suffer. Specifically, I predicted that teachers who reported more depressive symptoms would display feedback patterns characterized by more simplistic and negative feedback, and less elaborative/supportive and positive feedback.
4) To what extent does the quality of teachers' feedback impact the quality of the classroom- learning environment? I predicted that high-quality classrooms would be characterized by higher quality (more elaborative/supportive/positive) academic and behavioral feedback, while low-quality classrooms would be characterized by more simplistic and negative academic and behavioral feedback.
5) To what extent does the nature of teachers' feedback provided to students

mediate the relation between teachers' self-reported depressive symptoms and observed quality of the classroom-learning environment? I hypothesized that such a mediation effect would exist, such that as reports of depressive symptoms increased, feedback quality would decrease, resulting in a lower quality classroom learning environment.

Participants

Individualizing Student Instruction Study

The data used for the present studies were collected during the 2010-2011 year as part of an ongoing longitudinal parent study investigating classroom instruction in early education (Connor et al., 2013), which began in 2005. Six hundred and twenty five third-grade students in 32 classrooms across 8 schools in a North Florida were recruited for this larger study. Participating schools presented a wide range of local family SES, measured by percentages of students qualifying for a Free and Reduced Lunch (FARL) program based on reported family income. The lowest-SES school displayed 92% student enrollment in FARL and the highest-SES school had 4% student enrollment. Forty-six percent of students were male, 82% were Caucasian, 7% were African American, 5% were Hispanic, and 6% were Asian or mixed-race. Age of students ranged from 7 to 11 years, with a mean of 8.6.

All teachers involved in this study met state certification requirements and had at least a bachelor's degree related to education. Teachers' years of experience ranged from 0 to 31 years, with a mean of 10.9 years. Teachers participated in one of two instructional interventions as part of the parent study, each focused on individualized instruction in

either literacy or mathematics. Teachers were randomly assigned to either the literacy or mathematics intervention conditions and all children in each teacher's classroom were exposed to the assigned intervention. These interventions were not focused on teachers' psychological characteristics or feedback methods, and exploratory analyses revealed no significant differences in levels of depression, nor in the rated quality of the CLE, between the intervention groups. Potential differences in feedback patterns between the two intervention groups were tested in preliminary analyses and accounted for in subsequent analyses.

Final Sample

Eight to twelve target children were selected from each classroom to be coded during video observations (described below) using various observational coding systems within the parent longitudinal study (Connor et al., 2009). Children in each classroom were stratified based on Fall reading and math achievement and target children were randomly selected from each strata to comprise a sub-group of target students representing a wide range of academic achievement levels. A total of 310 children were selected and subsequently coded (using both the ISI coding system and, for the purposes of this study, the Teacher Feedback Coding System), and these 310 children comprise the final sample used in the current study.

Of this final sample, 49% (152 students) were female and 51% (158 students) were male. Seventy-two percent were Caucasian, 6% were African American, 4% were Asian, 3% were Hispanic and the remaining 15% were other ethnicities such as Native American or Multiracial. Age of students ranged from 7 to 11 years, with a mean age of 8 years. Regarding the parent study's interventions, 53% (165 students) were in

classrooms assigned to the mathematics intervention and the remaining 47% (145 students) were assigned to the literacy intervention. Descriptive statistics for teachers remained the same as outlined above.

Measures

Classroom Video Observations

Video observations were taken for 31 of the 32 participating classrooms (one teacher declined to be videotaped). Three whole-day video sessions were recorded per classroom, one each in the fall, winter and spring of the 2010-2011 academic year. During these classroom video observations, two video cameras were used to record all classroom activities taking place. Trained videographers managed the cameras, as well as wrote physical descriptions of all students present and took detailed notes of classroom activities, paying special attention to events that may have been ambiguous later in the coding process, such as noting when certain students left the room or providing details about a worksheet used. Typically, one camera would be used to capture the overall activities of the classroom from a wider viewpoint and the other camera was designated for closer observation of the teacher and the students working directly with the teacher. Using this method, both whole-class and small-group instruction could be accurately captured, even when multiple instructional activities were taking place simultaneously. The cameras used were able to capture high-quality audio data as well as high-quality video data.

Each whole-day recording captured designated instruction blocks for reading, writing, mathematics and science within each classroom. McLean & Connor (2015)

found within this sample that teachers' depression impacted student performance in mathematics but not literacy. As the present study is a direct expansion of this foundational study, only math instruction was coded for teachers' feedback. The designated math instruction block for each classroom was identified within the larger videos and captured separately. All on-topic instructional activities were coded for this math block. Teachers' time in mathematics instruction ranged from 27 to 80 minutes, with a mean of 60 minutes (SD = 11 minutes). On two occasions, teachers switched their focus to literacy within the designated math block for one or two lessons. These literacy lessons were not coded for teachers' feedback and were not factored into the total minutes of math instruction.

Teacher Feedback

Teachers' feedback for 30 of the 32 participating teachers was assessed using the Teacher Feedback Coding System (see appendix A), a novel measure developed for use and validation in the present study. One teacher declined videotaping and thus had no available video data, and one classroom was led by a student teacher during the winter observation and was thus not coded for feedback as that data would not match that of the other teacher-level variables used in analyses. The Teacher Feedback Coding System (McLean & Connor, in preparation) is an observational coding system for use with classroom video data that documents and categorizes teachers' feedback to students at the student level, across multiple domains. Each 'feedback event', defined as any instance in which the teacher is providing reactive commentary on a students' academic or behavioral performance within the classroom, is assessed based on feedback type, specificity, and teacher affect during the event. Feedback type categorizes each feedback

event as either academic or behavioral, that is, a teacher is either reacting to a students' attempt at learning (academic feedback) or a students' behavioral action or lack of action (behavioral feedback). Feedback specificity captures the level of detail and effort put forth by the teacher during the feedback event. Following are examples of each type of feedback specificity, both academic and behavioral, outlined by the system:

- Academic Mistake Identification: The teacher points out a student mistake without providing the correct answer. Example: "No, that's not quite right. Who else can tell me?"
- Academic Correction: The teacher points out a mistake and provides the correct alternative, but does not provide any further information. Example: "No, that's not quite right, the correct answer is 13"
- Academic Elaboration: The teacher responds to an attempt at learning by providing information to the student with the express goal of increasing that student's understanding of the involved academic concept. Example: "You answered 36, but if you make three piles with 8 stones in each pile, you can count them and see that eight times three is 24."
- Academic Encouragement: The teacher offers encouragement in response to a students' attempt at learning. This is differentiated from a 'support statement' because it happens more quickly, on a smaller scale, is less pointed, and comes across as more generic rather than a significant observation on the part of the teacher. Example: teacher looks over a student's shoulder as he fills out a multiplication worksheet. After 5 seconds of observing his work, she says "good job, these are looking good" and moves on.

- Academic Support Statement: The teacher responds positively to a students' academic attempt, or 'catches' and points out positive academic actions, in a way that is both purposeful and elaborative. This is differentiated from an academic encouragement by the level of effort and enthusiasm put forth by the teacher. Example: "You did it, David! You recited the entire times table for the number 9! Last week you had some trouble with that but you did it perfectly today, I am so proud of you"
- Behavioral Redirection: The teacher acknowledges incorrect behavior in some way and redirects the student to a different behavior. This redirection can either be an overt command or suggestion, or could be unspoken but heavily implied during the feedback event. Example 1 (overt redirection): "That's not what we are doing right now. You need to return to your desk." Example 2: "Shhh, quiet!", followed by the teacher making purposeful eye contact with the student and then pointing to the students' desk, implying that they go sit down.
- Behavior Elaboration: The teacher both addresses unwanted behavior and elaborates on why that behavior is disruptive within the context of the classroom. Example: "Billy, I don't like that you're out of your seat right now. All that moving around is distracting to the other students who are trying to read."
- Behavioral Encouragement: The teacher offers encouragement in response to a students' positive behavior. This is differentiated from a 'support statement' because it happens quickly on a smaller scale, is less pointed, and comes across as more generic rather than a significant observation on the part of the teacher. Example: teacher looks up from her small group briefly to monitor the class. She says "good

job everyone, you're all so quiet!" and quickly returns to her instruction.

- Behavioral Support Statement: Teacher 'catches' and points out students' positive behavior in a more purposeful and exaggerated manner. While this may seem simple, it is considered the most complex in nature because the teacher must notice positive behavior and make the effort to support it proactively instead of waiting for a negative stimulus. Example: "I really appreciate how quiet you've been during silent reading, I can tell that you were all working hard during reading time, you all get bonus bucks after lunch!"

Lastly, teachers' observable affect during each feedback event is documented. Coders' judgments of teacher affect are based on the teachers' observable facial expressions, vocal tones, body posturing, and body movements during the feedback event. Each event is characterized by one of the following five affect judgments:

- Content: The teacher is relaxed, happy, engaged with the student(s) but tones/movements are not exaggerated in any way. The face is generally relaxed, mouth is smiling, and cheekbones are slightly elevated. Eyes are open and alert but not wide. The teacher is using calm, positive tones with a friendly intent. The words used are positive but not exaggerated, small elements of praise and encouragement that could easily be considered "fillers". The shoulders are relaxed, not elevated or slumped, and the teacher is usually facing the student. Body movements are fluid, relaxed and happy, and can be either moving or stationary.
- Enthusiastic/exuberant: The teacher is excited, highly engaged, and providing exaggerated vocal tones, words and/or body movements. Her eyebrows may be raised in interest, eyes may be wide, mouth open with happy expressions. The voice

may be raised in excitement, the teacher uses very positive tones and wording. The enthusiastic teacher will likely be facing the student directly, may use exaggerated gestures or engage in physical connection with student(s) such as a pat on the shoulder. The teacher will likely remain stationary or move towards student.

- Neutral: The teacher is displaying neither positive nor negative characteristics. She is engaged with students but is doing so without any discernable emotion. Eyebrows are likely flat, eyes relaxed but open, mouth and lips relaxed but not actively smiling, the face is likely not tense. There are no discernable positive or negative vocal tones or wording the teacher can be facing student(s) directly or turned away. The teacher is likely stationary or slowly circulating the room, and will likely not move markedly towards or away from student(s).
- Sad/Depressed: The teacher is showing visible signs of frustration, irritation, and sadness in regard to student(s) or the material being taught. Affect may also be interpreted as flat and/or disengaged. Visible expressions may include a furrowed brown (inner corners move upwards and together), droopy facial features, and lips drawn down in a frown. This teacher will likely use quieter tones, may sound flat and disconnected or noticeably sad, may sigh frequently or take long pointed pauses when addressing student(s). Body postures include shoulders that are slumped downwards, head directed downwards and/or to the side, away from the student or situation**.** Sadness/depression usually involves disengagement from, or movement away from, the student or situation. This could be very subtle, as in turning the head away and down, or it could be more obvious such as slowly walking away.
- Angry/Frustrated: The teacher is visibly and audibly upset with the student(s) and

uses harsh tones, wording and/or body posturing in her attempts at feedback. The inner corners of the brow may be lowered towards the center of the face, eyes may be tense or squinted, and the mouth may be tense with lips either wide open or tightly pressed together. This teacher may use harsh, negative tones and the voice may become louder. The teacher may also exhale sharply in exasperation (frustration). Wording will likely be direct and negative. The body may be tense, with raised shoulders. Usually the teacher will be facing the student directly, and movement will usually be towards the student.

After judging each feedback event on its type, specificity and affect, the coder assigns the appropriate code. There are 45 possible codes, each specifying every domain discussed. For example, a feedback event may be given a code for "Academic Correction: Angry/Frustrated", or alternately "Behavioral Support Statement: Enthusiastic/Exuberant".

Teacher Depressive Symptoms

Twenty-seven of the 32 participating teachers completed an adapted version (see appendix B) of the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff 1977; alpha = .85) in the winter of 2010. All teachers who were given this scale completed and returned it, however five teachers were unavailable during the short window of time this measure was given, all due to prolonged holiday-related absences (the measure was given in mid-December). This scale includes 20 questions that ask subjects to report the frequency of their depressive symptoms. Scores range from 0 to 60 with 16 or higher indicating possible clinical depression. The adapted measure added 18 of the 20 questions to a larger self-efficacy survey to alleviate concerns about teachers'

sensitivity to a formal measure of depression. The Likert-scale was increased from 3 to 5 points in order to capture more nuanced levels of depressive symptoms, with a score of '1' indicating complete absence of a symptom and a '5' indicating constant presence of a symptom. Depression risk questions were scored separately from self-efficacy questions to determine each teacher's level of self-reported symptomatology. Scores within the teacher sample on this measure ranged from 22 to 62 with a mean score of 36 and a standard deviation of 9. Although few teachers reported markedly high levels of depressive symptoms, there was enough variability among teachers to continue with analyses. The adapted measure displayed acceptable reliability at alpha = .75. Teachers involved in this study were not professionally assessed for clinical depression nor did the questionnaire ask about any diagnosis of depression. We consider our measure an assessment of general risk for depression based on the presence of self-reported symptoms. It was beyond the scope of the study to base evaluation of depression on actual diagnosis.

Classroom Quality

Quality of the classroom-learning environment (Q-CLE) was assessed for 31 of the 32 classrooms (again, one teacher declined to be videotaped) in the winter of the academic year using the Classroom Learning Environment Rubric (Connor et al., 2011; 2014), an observational measure used within the parent study (see appendix C). Raters demonstrated adequate levels of inter-rater reliability (Cohen's Kappa = 0.73; Landis and Koch, 1977) upon initial assessment, and this level of reliability was maintained after recoding a randomly selected 10% of the videos. In other studies, this measure has predicted students' achievement outcomes (e.g., Connor et al., 2014), providing evidence

for its validity (Ochs, 1979). The three dimensions of the CLE rubric were moderately to highly correlated with each other (correlations ranged from .33 to .58, $p < .001$).

This scale assesses classrooms across three dimensions: implementation of individualized instruction, organization/planning, and teacher warmth/responsiveness. Scores on each dimension range from 1 to 6, with a score of 6 indicating exemplary practice on the part of the teacher and a score of 1 indicating weaker practice. This rubric is conceptualized to represent classroom quality, as opposed to teacher quality, as it takes into account students' and teachers' reciprocal interactions during instruction as well as the developmental appropriateness of the context of the educational environment. An exemplary rating (scored '6') on the 'organization/planning' dimension would indicate a classroom that is "well organized in its physical systems and instruction, with evident classroom routines and efficient transitions". An exemplary rating on the 'teacher warmth/responsiveness' dimension would indicate a classroom that "consistently offers a positive learning environment with clear expectations for students' behavior as a member of the learning community". Finally, an exemplary score in 'implementation of individualized instruction' would refer to a classroom in which "the content of literacy/math instruction is differentiated" and "the entire language arts/math block is spent in meaningful literacy/math activities."

Student Behavior

Students' patterns of externalizing behavior problems were assessed using 'Problem Behavior Scale' portion of the Social Skills Rating System (SSRS; Gresham & Elliott, 1990), a teacher-report measure that spans multiple domains of student characteristics. This measure asks teachers to report on their individual students'

behaviors by answering questions about their observable behavior patterns in class, such as whether they fight with others, bully or appear lonely. It has yielded alpha reliability coefficients of .87 for girls and .88 for boys in past studies. Of note, the higher the standard score, the worse the problematic behavior reported. Of the 310 students in the sample, 289 were assessed using this measure.

Student Mathematics Achievement

Students' math achievement was measured at three time points across the school year using the Woodcock-Johnson III tests of achievement (Woodcock, McGrew & Mather, 2001) Math Fluency and Applied Problems subtests. In the fall, 282 and in the spring, 279 students of the present study sample received mathematics assessments. The Math Fluency task is a timed measure of basic math skills that asks students to perform foundational math functions (addition, subtraction, and multiplication) with increasing difficulty, it has shown high reliability at alpha = .90. The Applied Problems task is a measure of students' mathematics processing skills. Children are presented with word-problems of increasing difficulty and asked to solve them using mental reasoning and pen and paper. This task has shown high reliability at alpha = .93.

Procedures

Data Collection

Assessment and video data were collected at three time points within the parent study across the school year: once each in the fall, winter and spring. Student academic achievement (of relevance to this study, the WJ math tasks) was assessed at each of these time points, and full-day video recordings were collected as well. Classroom-level

variables, namely classroom quality and teachers' depression, were assessed once in the winter of the school year when classrooms were at their most stable. The Teacher Feedback Coding System was applied only to the winter observations, again as an attempt to capture the most stable time of the school year. Further, since foundational work investigating teachers' depression in this same sample found effects for student mathematics achievement (McLean & Connor, 2015), only mathematics instruction was coded for teachers' feedback.

Teacher Feedback Coding System: Training & Reliability

One additional coder was trained by the primary investigator on the Teacher Feedback Coding System, and after training both coders carried out the coding of video data. Approximately 15 hours of training took place over the course of ten days, and opportunities for refinement and clarification of the coding manual continued through formal coding of video data as the two coders (the PI and the secondary coder) assessed the video data. Training consisted of two phases: first review, discussion and clarification of the coding system, and second practice coding using videos of classrooms from separate waves of the longitudinal parent study (2nd grade classrooms from academic year 2009-2010 in a separate district were used). Coders watched and co-coded one video together, and then practice coded another video separately. After practice coding, the two coders came together to compare results and discuss discrepancies.

Once training was completed, inter-rater reliability, or the extent to which multiple raters assign the same score to the same variable, was assessed. A randomly selected 10% of the videos (3 videos) from the pool of videos used in the present study were assessed by both coders and data were compared to ascertain the degree to which

coders were in agreement in their assessment of feedback events. Two strategies were used to assess reliability: percent-agreement and kappa (see Table 1). Percent-agreement, or the percentage of codes matched between coders out of the total number of codes recorded, was assessed at three levels: feedback type only (academic or behavioral), feedback type and specificity (ex. academic-elaborative vs. academic mistake ID), and feedback type, specificity and affect all considered together (ex. academic-elaborative-content vs. academic-elaborative-neutral). It was found that the two coders reached a percent-agreement of 81% in relation to their judgments of feedback type across the three videos. Further, coders displayed 79% agreement in their judgments of feedback type and specificity, and finally 81% agreement were reached in relation to judgments of feedback type, specificity and affect (also referred to as "code-level" percent agreement).

Table 1

Percent agreement across feedback levels, kappa inter-rater reliability estimates across reliability videos.

Percent Agreement	
Code level Percent Agreement	.81
Spec. Level Percent Agreement	.79
Type Level Percent Agreement	.81
Kappa	
Video 1 Kappa	.8
Video 2 Kappa	.76
Video 3 Kappa	.71
Average Kappa	.76

While this level of percent-agreement is considered high in the literature (Hill, Charalambous & Kraft, 2012), this method has been criticized for it's lack of ability to account for instances of chance agreement. In order to remedy this, inter-rater reliability was also evaluated using Cohen's Kappa (Cohen, 1960). Kappa is commonly used in observational coding systems as the assignment of a frequency code (i.e. the documentation that an event happened vs. no documentation for the absence of an event) is analogous to the assignment of a categorical variable. In general, kappa values above .75 are considered ideal, although in observational data a strong case can be made for leniency in the "acceptable" cutoff for this value. Kappa is usually calculated based on a finite and pre-determined set of judgments to be made by raters, for example assessing a psychological patient across 50 dimensions yields 50 opportunities for judgment. In contrast, observational data such as the videos used in the present study involve an

infinite number of potential judgments. Coders reached separate kappa values of .8, .76 and .71 on each of the three videos, for an average kappa of .76 across reliability videos. These values and the high percent-agreement values calculated show highly acceptable levels of inter-rater reliability on the Teacher Feedback Coding System.

Analytic Strategy

Factor scores, all of which center around zero and have a standard deviation of 1, were created for both student mathematics achievement (separate factor scores for fall and spring) and classroom quality. Principal Components Analysis in SPSS was used to confirm the consistency of the "math achievement" and "classroom quality" factors within this data. Student scores on the WJ Applied Problems and Math Fluency subtests both loaded strongly onto one factor, Math Achievement, each with loadings of .86 across both fall and spring. The three categories comprising the Q-CLE rubric, Organization/Orientation, Individualized Instruction, and Warmth/Responsiveness, all loaded onto a single factor, Classroom Quality, with loadings ranging from .68 to .83.

Exploratory Factor Analysis (EFA) in SPSS and Confirmatory Factor Analysis (CFA) in MPlus were used to investigate and confirm the factor structure of the Teacher Feedback Coding System data. Specifically, Maximum Likelihood extraction with Direct Oblimin Rotation was used in EFA. Factor loadings reported in the pattern matrix were the primary referents for determining potential factor structure. In CFA, indications of fit used included the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI), the Root-Mean-Squared Error of Estimation (RMSEA) and the Standardized Root Mean Square Residual (SRMS).

Multilevel modeling was used to investigate the relations between feedback

factors revealed from EFA/CFA and student behavior and mathematics achievement. As is the case with most educational data, the student participants in these data are nested within classrooms and this nesting warrants the use of multileveled analyses to account for variance attributable to between-classroom differences. Unconditional models, level-1 random-intercept predictor models, and level-2 means-as-outcomes predictor models were run to investigate the research questions posed. Child-level total scores, reflecting the total amount of each type of feedback received per child, were used to calculate classroom-level means and child-level deviations from their respective classroom mean, and these were the variables used in both multileveled and mediation analyses. Because teachers varied in the amounts of math instruction coded within each video, and because teachers were also involved in instructional interventions as a part of the longitudinal parent study, both of these factors were investigated in their impacts on target variables within the present study, as well as controlled for in all analyses.

Finally, basic linear regression analysis was used to investigate a potential mediation relation of teachers' feedback on the relation between teachers' depression and classroom quality. All variables in this mediation analysis were at the classroom level, using the classroom means for each feedback factor, teachers' reported depressive symptoms and observed classroom quality for each classroom. Each pathway of the mediation model was tested separately and then a final mediation analysis was performed that included both teachers' depression and each valid feedback factor as predictors of classroom quality.

CHAPTER 3. DATA ANALYSIS AND RESULTS

Factor Analyses

Exploratory (EFA) and confirmatory (CFA) factor analyses were performed to determine the factor structure of the feedback data for use in later analyses investigating whether teachers' feedback would be related to student math and behavioral outcomes. All factor analyses were performed at the child level, interpreted as the total number of feedback events experienced by the child.

Exploratory Factor Analysis

EFA with Maximum Likelihood extraction was performed to estimate the likely factor structure of the data. Academic and behavioral feedback data were first assessed separately to determine each type's unique factor structure, and then together to determine if academic and behavioral feedback were operating separately from each other as opposed to potentially operating together with a more generalized "overall feedback" factor structure. Analyses were first run at the code level to assess which individual codes seemed to be grouping together strongly enough to be collapsed. Subsequent analyses were run in stages after collapsing codes into larger variables.

Academic Feedback

For the code-level analysis, codes that had no data (i.e. that code was never observed during any video coding) were trimmed prior to performing factor analysis. This included the codes for "academic correction-sad", "academic elaboration-sad", "academic support statement-sad" and "academic support statement-angry". After trimming these variables, a code-level factor analysis was run to investigate how the

individual codes might be relating to each other. Eigenvalues greater than 1 (a common cutoff) were retained, and the analysis was rotated using Direct Oblimin rotation, which assumes the variables are correlated with each other.

Results of this first factor analysis estimated nine factors that accounted for approximately 67% of the variance in the data (see Table 2). Factor loadings suggested that "neutral", "angry" and "sad" affect ratings were grouping together, indicating a broader affect rating of "negative affect". Further, the more positive affect ratings of "exuberant" and "content" showed evidence of potential grouping, although not as strongly as angry, sad and neutral. Based on these findings, variables were collapsed within each category of feedback specificity by these affect groupings. Sad, angry and neutral affect codes were collapsed into "negative affect" and enthusiastic/exuberant and content affect codes were collapsed into "positive affect". Because academic encouragements and academic support statements were inherently positive, these data were summed without consideration of affect.

Table 2.

EFA Factor loadings for code-level academic feedback analysis.

	1	2	3	4	5	6	7	8	9
AcElNu	.73								
AcIDNu	.72								
AcIDAn	.7								
AcEnNu	.64						.21		
AcEnEx		.93							
AcSsEx		.87							
AcElCon			.86						
AcEnCon			.74				-.23		
AcSsCon			.63			.29			
AcIDSd				-.72					
AcEnAn				-.71					
AcElAn				-.7	.26		.21		
AcCoAn					.84				
AcEnSd					.8				
AcCoEx			-.2			.88			
AcIDCon			.26			.67	-.34		
AcCoCon			.25				-.67		
AcCoNu	.3		-.29	-.26			-.6		
AcElEx		.42						-.76	
AcIDEx		.58						.62	
AcSsNu									.98

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

A follow-up factor analysis was run using these collapsed variables. Each specificity category was now comprised of two types of affect, rather than five as was the case previously. Again, eigenvalues greater than one were retained, and the analysis was rotated using Direct Oblimin rotation. This analysis revealed three factors that accounted for 63% of the variance in the data (see Table 3). The loadings presented in the pattern matrix strongly supported the factors of "academic support", comprised of all academic

encouragement and support statement codes as well as positive academic elaborations and with factor loadings ranging from .68 to .76, "academic-informative-positive", comprised of positive academic identifications and corrections with factor loadings ranging from .74 to .89, and "academic-informative-negative", comprised of negative academic identifications, corrections and elaborations and with factor loadings ranging from .65 to .86.

Table 3.

EFA Factor loadings for academic codes collapsed by affect.

	1	2	3
AcSs	.76		
AcEn	.71		.24
AcElPos	.7		
AcIDNeg		.86	
AcElNeg		.76	
AcCoNeg	.33	.65	.23
AcCoPos			.89
AcIDPos			.74

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Behavioral Feedback

The same approach as outlined above was taken in determining the potential factor structure of the behavioral feedback data. Code-level variables with no data were trimmed from analyses, this time "behavioral encouragement-sad", "behavioral encouragement-angry", "behavioral support statement-sad" and "behavioral support statement-angry". The initial factor analysis run included all other code-level variables. Eigenvalues greater than 1 were retained and the analysis was rotated using Direct

Oblimin rotation as these variables were again assumed to be correlated.

This analysis estimated five factors accounting for approximately 62% of the variance in the data (see Table 5). Factor loadings closely followed the patterns revealed in the analysis of the academic feedback data in regard to affect: sad, angry and neutral affect codes seemed to be grouped together and enthusiastic/exuberant and content affect codes were grouped together. Code-level variables were again grouped by affect to create "positive" and "negative" affect groups for each feedback specificity category. Again, encouragements and support statements were inherently positive (and there was no data for the "sad" and "angry" codes for these specificity categories) so these codes were just summed with no regard for affect.

Table 4
EFA Factor loadings for code-level behavioral feedback analysis.

	1	2	3	4	5
BvRdCon	.89				
BvRdEx	.89				
BvElEx	.81				
BvElCon	.81				
BvRdAn		.85			
BvElAn		.83			
BvRdNu	.3	.61			-.35
BvSsNu			.65		
BvElNu			.58		-.39
BvSsCon			.52		
BvEnCon			.46		
BvEnNu			.39		-.36
BvSsEx				.92	
BvEnEx				.9	
BvElSd					-.8
BvRdSd		.36			.39

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

A follow-up factor analysis using these collapsed affect groups revealed three factors that accounted for approximately 72% of the variance in the data (see Table 5). These factors closely mirrored those detected in the academic feedback factor analysis: a "behavioral support" factor comprised of behavioral encouragements and support statements with factor loadings ranging from .64 to .74, a "behavioral-informative-positive" factor comprised of positive redirections and elaborations with factor loadings ranging from .82 to .92, and a "behavioral-informative-negative" factor comprised of negative redirections and elaborations with factor loadings ranging from .61 to .863 . Cross-loadings suggested some shared variance among the factors on certain variables,

however the strongest loadings were strong enough to retain assignment to each variable's estimated primary factor.

Table 5.

EFA Factor loadings for behavioral codes collapsed by affect.

	1	2	3
BvRdPos	.92		
BvElPos	.82		
BvEn		.74	
BvSs	.5	.64	
BvRdNeg	.3		-.86
BvElNeg	-.27	.47	-.61

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Feedback Type Differentiation

Because both analyses revealed the same factor structure, the final step in determining the likely factor structure of the data was to analyze academic and behavioral feedback together to see if the two types of feedback were differentiated from each other. An analysis was run with the same parameters as above that included all collapsed support variables, both academic and behavioral, to investigate whether academic and behavioral support were two separate factors. The result was a two-factor model that explained 63% of the variance with both behavioral and academic support variables loading together in a group (see Table 6). However, the only valid loading on the second factor was that of behavioral encouragements. Because previous analyses using only the behavioral factors showed strong evidence that behavioral encouragements were grouped strongly with behavioral support statements, and because in this combined analysis

behavioral support statements loaded strongly onto the "overall support" factor with all other variables, investigators decided to continue analyses assuming one factor for "overall support".

Table 6.

EFA Factor loadings for academic and behavioral support modeled together

	1	2
AcEn	.82	
BvSs	.74	
AcSs	.69	.2
AcElPos	.63	-.35
BvEn		.93

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

This same method was used to investigate whether the academic-informative and behavioral-informative factors previously identified could be differentiated from each other. An analysis was run using all variables that comprised these four factors. The resulting model revealed four factors accounting for approximately 72% of the variance in the data (see Table 7). The four factors exactly mirrored those detected in previous analyses, with academic and behavioral feedback factors strongly differentiated from each other. Thus, the resulting five factors to be confirmed using Confirmatory Factor Analysis were: Overall Support, Academic-Informative-Positive, Academic-Informative-Negative, Behavioral-Informative-Positive, and Behavioral-Informative-Negative.

Table 7.

EFA Factor loadings for academic and behavioral informative factors modeled together.

	1	2	3	4
AcIDNeg	-.84			-.21
AcCoNeg	-.74			.28
AcElNeg	-.72		-.22	-.26
BvRdPos		.92		
BvElPos		.84		
AcCoPos			.88	
AcIDPos			.81	-.34
BvElNeg				-.84
BvRdNeg		.5		-.58

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Confirmatory Factor Analysis

The resulting factor structures suggested by the Exploratory Factor Analysis were followed up in Confirmatory Factor Analysis to verify the validity of these factors using a more rigorous approach. Separate models were run testing a one-factor "support" model, a two-factor "academic-informative" model, and a two-factor "behavioral-informative" model (see Table 8). Maximum likelihood estimation was used, and the primary indicators of model fit considered were the CFI (.95 ideal), TLI (.95 ideal), RMSEA (less than .1 ideal) and SRMS (less than .08 ideal). While the Chi-Square Test of Model Fit, which tests the proposed model in relation to a perfectly fitting model, is one of the most commonly used indicators of model fit, it was determined to not be an ideal fit estimator for this particular study. This estimation assumes a very large sample size as well as normally distributed variables, and favors models with many parameters, all of which are not realities in the present study. These factors may make the chi-square test a less-than-

ideal measure of model fit for this particular set of analyses. The Standardized Root Mean Square Residual (SRMR) is a standardized calculation of the average difference between values in the observed and implied covariance matrices estimated by the model. The Comparative Fit Index and Tucker-Lewis Index are estimations that compare the proposed model to a hypothetical null model, with values close to .95 showing good fit. Finally, the Root Mean Squared Error of Approximation is an absolute measure of fit where zero indicates a perfectly fitting model and values less than .1 indicate well-fitting models.

Table 8

CFA Model Fit Indices for best-fitting feedback factor models.

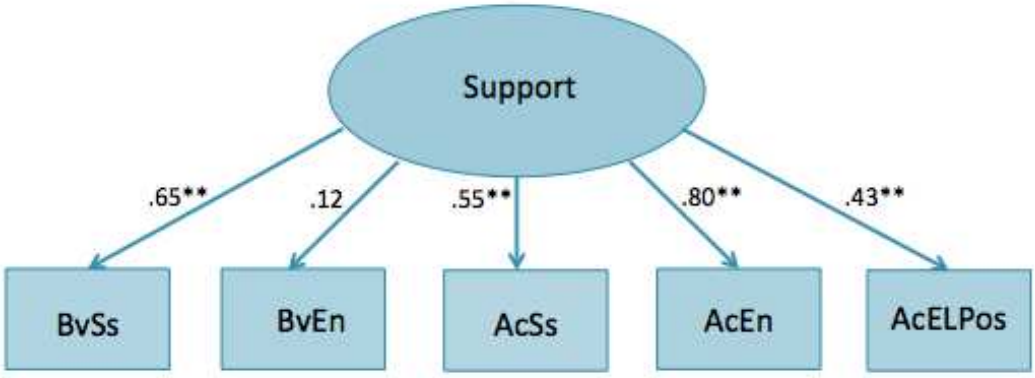
	RMSEA	p RMSEA <.05	CFI	TLI	SRMR
Support 1-Factor	.04	.44	.98	.97	.03
Academic-Info. 2-Factor	.06	.32	.98	.95	.03
Behavioral-Info. 2-Factor	.14	0	.78	.68	.08

Support Factor Model: The one-factor model run testing overall "support", comprised of both academic and behavioral encouragements and support statements and positive academic elaborations (see Figure 2), fit the data very well. The RMSEA was well under the .1 cutoff for good model fit at .045, with a probability statistic that RMSEA is less than .05 estimated at .436. CFI and TLI were both close to the ideal

.95 value, at .983 and .965 respectively. Lastly, the SRMR was well under the .08 cutoff for good fit at .031. These values indicate that this one-factor model incorporating the four "encouragement/support" variables and the positive academic elaboration variable is very accurately representing the data. Factor loadings were strong and significant in all cases (ranging from .43 to .80) except for behavioral encouragement feedback, which showed a weak and non-significant loading onto the 'Support factor'. This is consistent with what was previously shown in EFA (see Table 6), and suggests that the overall 'Support' factor is not indicating any significant variance in this particular type of feedback within the data. However, when this model was run without behavioral encouragements, model fit did not significantly improve.

Figure 2

Supportive Feedback 1-Factor Model.



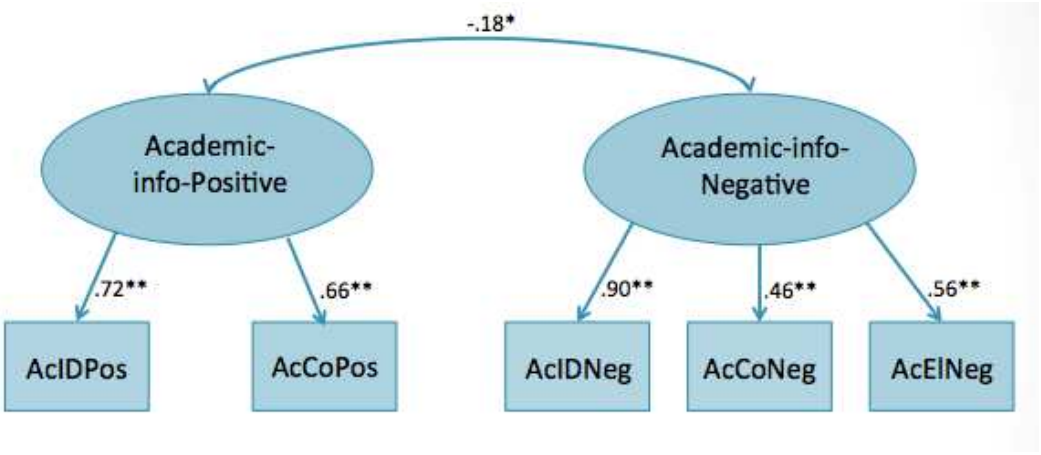
Note: Please refer to Feedback Shorthand Guide in List of Symbols/Nomenclature

Academic-Informative 2-Factor Model: A model was run testing the two academic-informative factors suggested in EFA: Academic-Informative-Positive and

Academic-Informative-Negative. The Academic-Informative-Positive factor indicated positive academic identifications and corrections while the Academic-Informative-Negative factor indicated negative identifications, corrections and elaborations (see Figure 3). These two factors were allowed to correlate with each other in the model, and showed a small but significant negative correlation (r = -.18, p = .026) . Model fit estimates indicated that this factor structure fit the data very well, with RMSEA = .06 (p RMSEA <= .05 = .323), CFI and TLI = .98 and .95 respectively, and SRMR = .027. Factor loadings of feedback variables onto their respective latent variables were all strong and significant, ranging from .46 to .90 across both latent variables. These values indicate that these two factors accurately represent the data.

Figure 3

Academic-Informative Feedback 2-Factor Model.



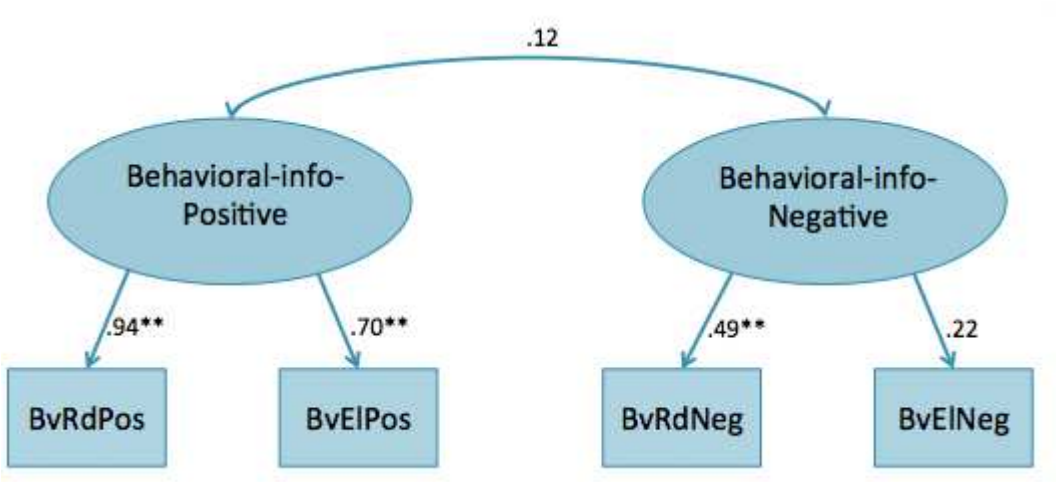
Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Behavioral-Informative 2-Factor Model: The final model run tested the fit of a 2-factor model that included the factors Behavioral-Informative-Positive and Behavioral-Informative-Negative. Behavioral-Informative-Positive indicated positive redirections

and elaborations, while Behavioral-Informative-Negative indicated negative redirections and elaborations (see Figure 4). These two factors were allowed to correlate with each other in the model, however model results indicated that these two latent variables were not significantly correlated with each other (r = .12, p = 11). This model showed poor fit to the data in comparison with the results of the other two models. The RMSEA = .1 (p RMSEA <= .05 = .000) was just at the cutoff for good fit, CFI and TLI = .78 and .68 respectively were less than the ideal .95 value, and SRMR = .08 was also right at the cutoff for a good fitting model. Factor loadings showed that feedback variables loaded strongly and positively onto the "behavioral informative-positive" factor, suggesting that this latent variable was indeed indicating variance in positive behavioral redirections and elaborations within the data. However, factor loadings for the "behavioral-informative negative" latent variable were more controversial, suggesting that the poor model fit may be due to this particular piece of the model. Alternatives to this model were run, for example a one-factor model proposing one "behavioral-informative" factor and a two-factor model that did not allow the two latent variables to correlate and excluded the feedback codes that did not load significantly onto the latent variables, but all models displayed very similar significantly worse fit than the estimates revealed for this two-factor model. While this was the weakest of the models run in CFA, this factor structure was used in subsequent predictive analyses as it was the strongest fitting structure revealed in CFA analyses.

Figure 4

Behavioral-Informative Feedback 2-Factor Model



Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Final Factors

After exploration and confirmation of the factor structure of the data collected using the Teacher Feedback Coding System, five factors were revealed that were used in subsequent analyses:

1. Support – comprised of both academic and behavioral encouragements and support statements, as well as positive academic elaborations.
2. Academic-Informative-Positive – comprised of positive academic identifications and corrections.
3. Academic-Informative-Negative – comprised of negative academic identifications, corrections and elaborations.
4. Behavioral-Informative-Positive – comprised of positive redirections and elaborations.

5. Behavioral-Informative-Negative: comprised of negative redirections and elaborations.

Preliminary Analyses

Descriptive Statistics

Descriptive Statistics were run prior to formal analyses for all variables of interest (see Table 9). In general, students made expected gains in math from fall to spring and the math achievement and behavior variables were normally distributed. Teachers did not report remarkably high average levels of depression, but there was considerable variability across teachers marked by the high standard deviation for the teacher depression measure, and this variable showed a slightly leptokurtic distribution. Non-normal distributions were detected for all of the teacher feedback factors, with high positive estimates of both skewness (ranging from 1.2 to 3.8) and kurtosis (ranging from 4.1 to 15.7). Overall, students received more negative feedback (both behavioral and academic) than positive, as well as high amounts of supportive feedback.

Table 9

Descriptive statistics for all primary variables.

	N	Min	Max	Mean	SD	Skew	SE	Kurt	SE
Teacher Depression	235	22	62	35.8	9.1	1.1	.16	1.1	.3
Q-CLE	295	-2.2	2.1	.1	.91	-.45	.14	.6	.3
Prob. Behaviors	288	85	142	96.8	13.6	1.02	.14	.1	.3
Fall Math	281	-3.1	2.5	.02	1.03	-.1	.15	.1	.3
Spring Math	278	-3.7	2.5	.01	1.02	-.3	.15	.4	.3
BvInfoPos	284	0	8	.5	1.41	3.8	.15	15.7	.3
BvInfoNeg	284	0	20	4.4	3.76	1.8	.15	4.1	.3
AcInfoPos	284	0	4	.3	.7	3.2	.15	11.4	.3
AcInfoNeg	284	0	16	2.3	3.1	2.3	.15	6.1	.3
Support	284	0	14	3.8	3.4	1.2	.15	0.7	.3

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Covariates

Four covariates were identified that may have impacted student-level outcomes above and beyond the target independent variables (feedback factors). As discussed previously, classrooms were assigned to participate in either a literacy-focused or math-focused instructional intervention within the longitudinal parent study. Effects of intervention on student outcome variables and teachers' feedback were tested using multilevel modeling and subsequently controlled for when appropriate. Classroom observations of Q-CLE took place during either literacy or mathematics instruction,

however when these observations took place was determined by assignment to intervention (i.e. classrooms exposed to the math intervention were observed for Q-CLE during math instruction). Therefore this difference in Q-CLE observation context was captured by controlling for overall assignment to intervention when applicable.

Additionally, teachers varied in the minutes of mathematics instruction they applied to students during video observations, with the shortest amount of math instruction at 27 minutes and the longest at 80 minutes (mean 60 minutes). Minutes of instruction was controlled for in all multilevel analyses. Finally, student-level gender and SES were controlled for in all multilevel models with student-level outcomes. Students were assigned a '0' in the data to indicate male gender, and a '1' to indicate female, and SES was noted using parent-reported level of enrollment in school-wide free and reduced lunch program. A zero score indicated no enrollment, a '1' indicated a denied application for enrollment, a '2' indicated enrollment in reduced lunch pricing, and a '3' indicated enrollment in free lunch.

Correlation Analyses

Correlations were run examining the baseline relations among all primary variables, as well as summed totals representing all academic feedback and all behavioral feedback (see Table 10). A negative relation was replicated between teachers' depression and the quality of the classroom environment ($r = -.357$, $p<.001$), consistent with findings from McLean & Connor, 2015. Further, these correlations revealed relations between teachers' depression and amounts of positive behavioral-informative feedback ($r = -.211$, $p = .002$), positive academic-informative feedback ($r = -.183$, $p = .007$), and the total

amount of behavioral feedback (r = -.155, p = .023). Classroom quality was related to students' observed problem behaviors (r = -.119, p = .048) spring math achievement (r = .166, p = .006), as well as to total amounts of both academic and behavioral feedback (r = .221.145, p<.001, =.017). Classroom quality was also related to positive behavioral-informative feedback (r = .29, p<.001), positive academic-informative feedback (r = .158, p = .009) and supportive feedback (r = .278, p<.001). No relations between feedback factors and students' mathematics or behavioral outcomes, but negative academic-informative, as well as total academic feedback were both related to students' fall math performance (r = -.139, -.145, p = .026,.02). This more general finding provides support for testing interactions involving students' fall math performance in multilevel modeling analyses.

As a follow-up to this more general correlation analysis, relations between teachers' depression, classroom quality, and affect displayed during feedback events were investigated (see Table 11). Total amounts of feedback characterized by each of the five affect categories were summed, disregarding feedback type and specificity, and these five affect variables were examined in their relations to teachers' self-reported depressive symptoms. Correlations revealed highly significant negative relations between teachers' depression and both enthusiastic/exuberant (r = -.19, p = .005) and content (r = -.207, p = .002) feedback, suggesting that teachers who were experiencing more depressive symptoms were less likely to have positive affect during feedback events. Classroom quality was positively related to both of these "positive" feedback types (r = .19, p = .002 for enthusiastic/exuberant, r = .45, p < .001 for content) and was also negatively related to neutral feedback (r = -.14, p = .023).

Finally, correlations were run investigating the relations between the four identified covariates and each of the primary variables involved in the present study (see table 13). Results revealed that minutes of instruction predicted students spring math outcomes ($r=.152$, $p<.05$), but was not related to any of the feedback factor variables. Student gender and SES were both related to student behavior, such that boys and low-SES students showed more behavior problems. Boys also received more behavioral-informative-negative feedback. An expected relation between SES and spring math existed, such that lower-SES students struggled more with mathematics performance ($r = -.212$, $p<.001$). Student SES was not related to any of the teacher feedback factor variable**s.**

Table 10

Correlations among primary variables, including total academic, behavioral and overall feedback amounts.

	1	2	3	4	5	6	7	8	9	10	11	12
1. TDQ	1											
2. Q-CLE	-.36**	1										
3. PB	.06	-.12*	1									
4. F Math	-.05	.10	-.08	1								
5. Sp Math	-.09	.17**	-.14*	.83**	1							
6. TotAcFB	-.08	.22**	.04	-.15*	-.08	1						
7. TotBvFB	-.16*	.15*	-.03	-.01	-.06	.17*	1					
8. BvInfPos	-.21**	.29**	-.05	.09	.11	.29**	.58**	1				
9. BvInfNeg	-.12	.01	.01	-.03	-.08	-.04	.91**	.25**	1			
10. AInfPos	-.18**	.16**	-.01	-.07	-.03	.29**	.25**	.07	.18**	1		
11. AInfNeg	-.04	-.10	.05	-.14*	-.12	.58**	.09	-.09	.10	-.16**	1	
12. Supp	.07	.28**	-.01	-.09	-.04	.75**	.21**	.47**	-.11	.36**	-.01	1

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

*Correlation is significant at the < .05 level
** Correlation is significant at the < .001 level

Table 11

Correlations among teacher depression, classroom quality, and total feedback affect scores.

	1	2	3	4	5	6	7
1. TDQ	**1**						
2. Q-CLE	-.36**	**1**					
3. TotFBEx	-.19**	.19**	**1**				
4. TotFBCon	-.21**	.45**	.31**	**1**			
5. TotFBNu	.06	-.14*	.01	-.12*	**1**		
6. TotFBAn	.08	.06	-.04	-.14*	.17**	**1**	
7. TotFBSd	-.09	-.04	.02	-.29**	-.29**	.32**	**1**

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

*Correlation is significant at the < .05 level
** Correlation is significant at the < .001 level

Table 12

Correlations among covariates (total minutes of math instruction, student gender and student SES) and all feedback variables, student behavior and spring math achievement.

	1	2	3	4	5	6	7	8	9	10
1. MinsInst	1									
2. Gender	.06	1								
3. SES	-.01	-.06	1							
4. Sp. Math	.15*	-.06	-.21**	1						
5. Behavior	.05	-.16*	.14*	-.14*	1					
6. BvInfPos	0	-.08	-.03	.01	-.02	1				
7. BvInfNeg	.01	-.14*	.08	.01	.17**	.25**	1			
8. AInfPos	.01	-.04	.08	.02	.02	.07	.17**	1		
9. AInfNeg	.02	-.09	-.04	-.13*	.07	.02	.18**	.09**	1	
10. Support	.01	.06	-.01	-.04	.05	.02	.12	.16**	.23**	1

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

*Correlation is significant at the < .05 level
** Correlation is significant at the < .001 level

Multilevel Modeling

MLM in SPSS was used to investigate relations between the five teacher feedback factors and student math and behavioral outcomes. This method was deemed most appropriate due to the nested structure of the data, with students (level-1) nested within classrooms (level-2).

Unconditional Models

Unconditional multilevel models were first run on all primary level-1 variables to provide information about the intercepts (means) of each variable, as well as the amount of variance at each level of the data (see Table 13). The total numbers of feedback events received by students for each feedback factor (a level-1 variable) were used in these models. These models were estimated in SPSS based on the following formula:

$$\mathrm{SpMath}_{ij} = \gamma_0 + \mu_{0j} + \varepsilon_{ij}$$

Where the spring math factor score for person i in group j (SpMath_{ij}) is equal to the grand mean of the sample (a fixed effect, γ_0) plus the level-2 mean differences (a random effect, μ_{0j}) plus level-1 within-person differences represented by a residual term (a random effect, ε_{ij}). This equation yields three important estimates: γ_0 the grand mean, θ^2_{uo} the level-2 intercept variance (the average variance of the cluster means from the grand mean) and θ^2_e the level-1 residual variance (the average variance of individual scores from their cluster means). These estimates will be used to calculate the intra-class correlation (ICC), which quantifies the proportion of variance in the data that exists at level 2. The equation for the ICC is as follows:

$$\mathrm{ICC} = \frac{\theta^2_{uo}}{(\theta^2_{uo} + \theta^2_e)}$$

ICC estimates range from 0, indicating no variance at level 2, to 1, indicating 100% of the variance at level 2. An ICC of .1 or higher indicates need for MLM (Wagner et al., 2013), but in practice, MLM can be used with ICCs as low as .05 (Dyer, Hanges & Hall 2005). Generally, ICCs ranging from .15 to .25 tend to be seen in educational data (Hedges & Hedberg, 2007). Therefore, multilevel modeling will be used if ICCs greater than .08 are detected in the data.

Supportive Feedback. The unconditional model for the total amount of supportive feedback received by students, as measured by the "support" factor determined in factor analysis, revealed a grand mean of 3.8 (interpreted as 3.8 instances of feedback received). The ICC calculated from the estimates given for level-1 and level-2 variance was .82, meaning that 82% of the variance of this variable was attributable to classroom-level differences.

Academic-Informative-Positive Feedback. The unconditional model for the total amount of positive informative academic feedback received by students revealed a grand mean of .29. The ICC calculated from the estimates given for level-1 and level-2 variance was .75.

Academic-Informative-Negative Feedback. The unconditional model for the total amount of negative informative academic feedback received by students revealed a grand mean of 2.3. The ICC calculated from the estimates given for level-1 and level-2 variance was .74.

Behavioral-Informative-Positive Feedback. The unconditional model for the total amount of positive informative behavioral feedback received by students revealed a grand mean

of .52. The ICC calculated from the estimates given for level-1 and level-2 variance was .96.

Behavioral-Informative-Negative Feedback. The unconditional model for the total amount of negative informative behavioral feedback received by students revealed a grand mean of 4.5. The ICC calculated from the estimates given for level-1 and level-2 variance was .78.

Student Behavior. The unconditional model for students' observed problem behaviors, as measured by standard scores on the SSRS Problem Behaviors subscale, revealed a grand mean of 96.6 for the child sample. The ICC calculated from the estimates given for level-1 and level-2 variance was .13, meaning that 13% of the variance of this variable was attributable to classroom-level differences.

Student Math Achievement. The unconditional model for students' spring math achievement, as measured using the factor score created from scores earned on the WJ math tasks, revealed a grand mean of -.004 for the child sample. This is to be expected as factor scores center around a grand mean of zero. The ICC calculated from the estimates given for level-1 and level-2 variance was .12, meaning that 12% of the variance of this variable was attributable to classroom-level differences.

Table 13

Unconditional model estimates for all child-level variables.

	Intercept	SE	L-1 Variance	L-2 Variance	ICC
Support	3.8	.58	2.18	10.11	.82
AcInfoPos	.29	.12	.14	.42	.75
AcInfoNeg	2.3	.5	2.63	7.4	.74
BvInfoPos	.52	.26	.09	2.14	.96
BvInfoNeg	4.5	.64	3.2	12.53	.78
Prob. Behaviors	96.6	1.18	161.82	24.13	.13
Spring Math	0	.09	0.9	.13	.12

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Testing for ISI Intervention Effects

As mentioned previously, teachers (classrooms) were randomly assigned to participate in either a math-focused or literacy-focused instructional intervention as part of the longitudinal parent study. While neither of these interventions focused on teachers' mental health characteristics or feedback patterns, effects of intervention and potential group differences between the reading and math intervention classrooms were tested. Past analyses with this data set revealed no significant differences between intervention groups on measures of teachers' depression and observed classroom quality.

Multilevel means-as-outcomes models with level-2 predictors (intervention assignment, in this case) were used to do this. This type of model uses a cluster-

level variable to predict between-cluster differences in a target outcome variable measured at level-1 (student math achievement). The classroom-level predictor variables (assignment to either of the two interventions) was used to predict between-classroom differences on all student-level variables (math, behavior, teacher feedback factors). The equation for each of this two-level model is as follows, with spring math achievement as the outcome of interest in this example:

$$\mathrm{SpMath}_{ij} = \gamma_0 + \gamma_1 \left(\mathrm{condition}_j\right) + \mu_{0j} + \varepsilon_{ij}$$

Here, the spring math performance of student i in classroom j can be predicted by the grand mean of the sample (γ_0) plus the explained mean differences dependent on classroom participation in the assigned intervention condition j ($\gamma_1 \left(\mathrm{feedback}_j\right)$) plus the residual mean differences of classrooms from the grand mean (μ_{0j}) plus a residual term representing within-person variation (ε_{ij}). This equation yields four important parameter estimates, the intercept of the outcome variable (γ_0), the regression coefficient representing the impact of the independent variable on the dependent variable (γ_1), the level-2 residual variance, or residual mean differences (θ^2_{uo}) and the level-1 residual variance, or within-person differences (θ^2_e). Important to note, this same multilevel modeling structure will be applied to all future analyses (with the exception of the final mediation model).

Intervention Effects on Student Outcomes

Dummy coding was used to create a variable indicating assignment to condition at the child level. If the child was in a classroom assigned to the mathematics intervention, they were given a code of '1'. If they were in a classroom assigned to the literacy

intervention, they were given a code of '0'. Multilevel predictor models (described above) were run with this variable as the predictor to test any differences in target child-level dependent variables based on group assignment (see Table 14). Results revealed no significant differences between groups on spring math outcomes (B = .007, p = .968) or the SSRS Problem Behaviors scale (B = .86, p = .722).

Intervention Effects on Feedback Factors

The same modeling strategy was used to assess between-group differences on each of the teacher feedback factors between the literacy and math intervention groups (see Table 14). It was revealed that students in the math intervention experienced fewer instances of supportive feedback (B = -2.2, p = .05) as well as fewer instances of negative behavioral-informative feedback (B = -2.8, p = .03). All other feedback factors were not significantly influenced by assignment to intervention group. As a result of these analyses, effects of intervention will be controlled for in all further analyses in which support feedback or negative behavioral-informational feedback is included as a predictor variable.

Table 14.

Effects of intervention assignment on child-level variables.

	Int./ Coeff.	SE	Sig.
Support	5		
MathInt	-2.25	1.1	.05
AcInfoPos	.42		
MathInt	-.24	.24	.32
AcInfoNeg	2.74		
MathInt	-.88	1	.39
BvInfoPos	.95		
MathInt	-.82	.51	.12
BvInfoNeg	5.9		
MathInt	-2.78	1.21	.03
SSRS - PB	96.15		
MathInt	.86	2.41	.72
Sp. Math	-.01		
MathInt	.01	.18	.97

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Teacher Feedback Predicting Student Outcomes (RQ 2)

Multilevel random intercept models with level-1 predictors were run to investigate whether the teacher feedback factors, each assessed individually, predicted students' behavioral and mathematics outcomes:

$$\mathrm{SpMath}_{ij} = \gamma_0 + \gamma_1 \left(\mathrm{support}_{ij}\right) + \mu_{0j} + \varepsilon_{ij}$$

In this equation, the spring math score for student i in classroom j is a sum of the grand mean of the sample (γ_0), the explained level-1 variation (γ_1), the residual mean differences (μ_{0j}) and the residual level-1 variation (ε_{ij}). This model yields four estimates: the intercept, the regression coefficient, and the level-1 (within-person) and

level-2 (cluster mean) residual variances. This model assumes a common slope across clusters and levels (as opposed to a random slope model), as well as normal distribution of residuals and consistency across clusters in level-1 variance.

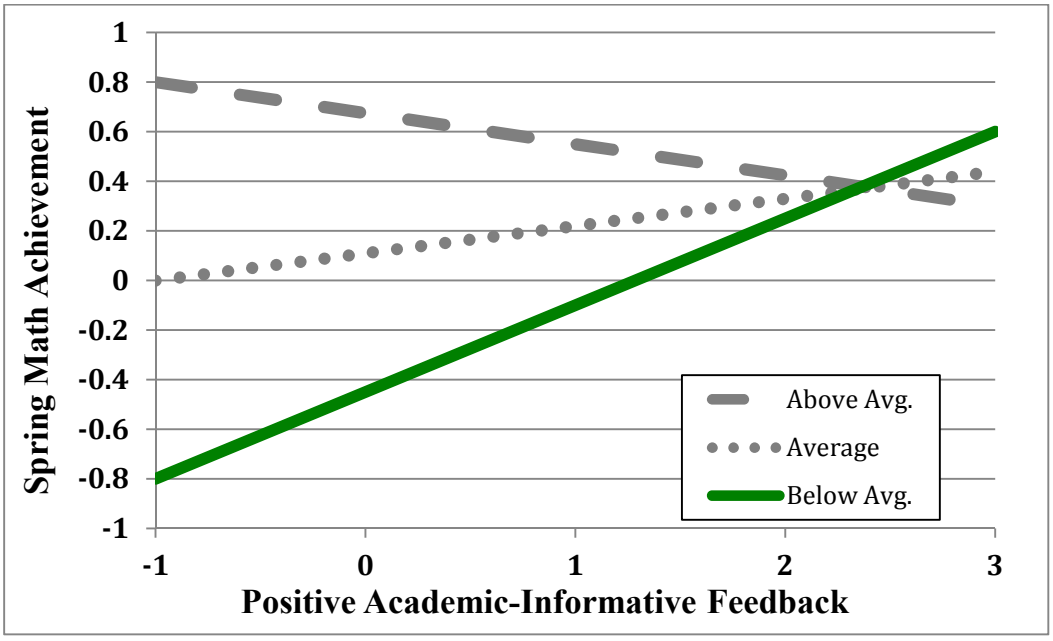
Student Math Achievement

Models were run with each of the five level-1 teacher feedback variables (child-level deviances from the classroom mean) predicting student's spring math achievement based on the factor variable created for math. Each of these models controlled for minutes of mathematics instruction provided by the teacher, student gender and SES, and assignment to intervention condition in the cases of supportive and negative behavioral-informative feedback. It was found that negative academic-informative feedback received by students was related to students' spring math achievement ($B = -.12$, $p = .002$) above the influence of the included covariates, such that the more of this type of feedback received by students, the lower their spring mathematics achievement scores. The estimated effect size of this predictive model was .23, meaning that this type of feedback, along with the included covariates, accounted for about 23% of the variance in students' math outcomes. Follow-up analyses were run testing interaction effects between students' fall math achievement and each of the feedback variables. A significant interaction effect was found between fall math and positive academic-informative feedback ($B = -.23$, $p = .05$), controlling for all covariates (see Figure 5). The estimated effect size of the model including this interaction was .74, meaning that about 74% of the variance in the outcome variable was attributable to the inclusion of the predictor variables. This is a very large effect size, however most of this is likely attributable to the inclusion of students' fall math scores in interaction included in the

model. The simple slopes of this interaction effect, or the relative effect of teachers' positive academic-informative feedback on students; spring math outcomes at high, average and low levels of students' fall math, were calculated. The tests of simple slopes revealed that students who began the year struggling in math (students who scored one standard deviation below the mean on fall math), but received more positive academic-informative feedback showed higher spring math scores than peers who received less feedback ($B = .35$, $p = .034$). Neither of the simple slopes depicting the effect of positive academic-informative feedback on spring math scores at high (+1 SD) and average (mean) levels of fall math achievement were significant, meaning the effect of teachers' positive academic informative feedback only predicted stronger outcomes for students who began the year struggling in math.

Figure 5

Fall Math X Positive Academic-Informative Feedback interaction effect. Students who began the year struggling in math and who received more positive academic-informative feedback showed higher spring math scores (in green).



Student Behavior

Models were run with each of the five level-1 teacher feedback variables (child-level deviances from the classroom mean) predicting student behavior measured by the SSRS. Total minutes of math instruction, student gender, and student SES were controlled for in each of these models, and assignment to ISI intervention was controlled for in the cases of supportive and negative behavioral-informative feedback. It was found that negative behavioral-informative feedback received by students was related to students' behavior problems ($B = 1.10$, $p = .025$), such that more of this type of feedback was related to more problematic behavior in students, above the effect of the included

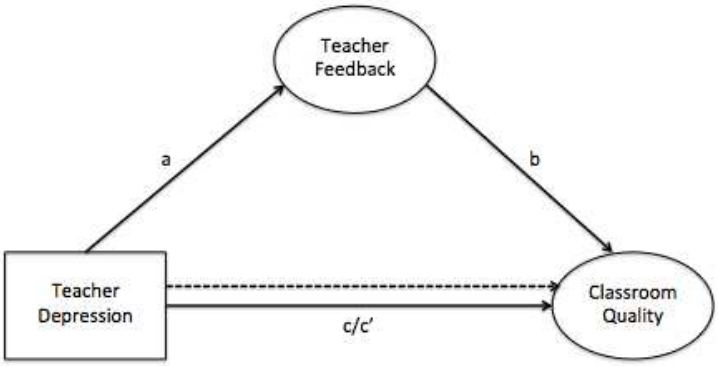
covariates. The estimated size of this effect was .19, meaning that the introduction of feedback and the included covariates accounted for about 19% of the variance in students' behavior patterns. Because both of these variables were measured in the winter of the academic year, no temporal precedence exists to suggest the directionality of this effect. Follow-up analyses testing interaction effects between students' fall math achievement and each of the feedback variables did not reveal any significant effects for any of the teacher feedback factors.

Mediation Modeling

The relations among teachers' depression, classroom quality, and each of the teacher feedback variables were first tested in separate pathways using basic linear regression analysis, and then together in the proposed mediation model (see Figure 6). Whereas teachers' feedback was previously represented as a level-1 variable (student-level deviations from the classroom mean) for each factor, it was now represented at level-2 with classroom-level means for each factor. Because each variable involved in the tested pathways of the proposed mediation model were all at level-2, multilevel modeling was no longer used.

Figure 6.

Proposed mediation model. Teacher feedback patterns mediate the relation between teachers' depression and classroom quality.



Teacher Depression and Classroom Quality (replication)

A basic linear regression model was run testing the relation between these two variables with teachers' depression as the predictor and classroom quality as the outcome. This relation represents the c pathway of the proposed mediation model. Results were consistent with the relation detected in McLean & Connor, 2015: teachers who reported more depressive symptoms tended to have lower-quality classrooms ($B = -.034$, $p < .001$, $R^2 = .13$). Therefore, the c pathway of the proposed mediation model was significant.

Teacher Depression and Teacher Feedback (RQ 3)

Further regression models were run investigating the relations between teachers' depression (always the predictor variable) and each of the five feedback factors (each a separate outcome). Significant relations between teachers' depression and positive behavioral-informative feedback, negative behavioral-informative feedback, and positive

academic-informative feedback were detected (see Table 15). Teachers who reported more depressive symptoms tended to provide less positive behavioral-informative feedback (B = -.033, p = .001), less negative behavioral-informative feedback (B = -.051, marginally significant at p = .051) and less positive academic-informative feedback (B = -.014, p = .002). Estimations of the proportions of reduction in variance in teacher feedback attributable to teachers' reported depressive symptoms were small, ranging from .02 to .05. Overall, results show that the a pathway of the proposed mediation model was significant in the cases of teachers' depression predicting both positive and negative behavioral-informative feedback, as well as positive academic-informative feedback.

Table 15.

Effects of teachers' depression on each of the teacher feedback factor variables.

	Int./ Coeff.	SE	Sig.	R-Squared
BvInfoPos	1.6			
TDQTot	-.03	.01	<.01	.05
BvInfoNeg	6.4			
TDQTot	-.05	.03	.05	.02
AcInfoPos	.78			
TDQTot	-.01	.01	<.01	.04
AcInfoNeg	2.05			
TDQTot	-.01	.01	.65	<.01
Support	2.9			
TDQTot	.02	.02	.38	<.01

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Teacher Feedback and Classroom Quality (RQ 4)

This same technique was used to investigate relations between each of the teacher feedback factors, this time as separate predictor variables, and classroom quality. Total minutes of math instruction were controlled for in each of these models, and assignment to ISI intervention was controlled for in the cases of supportive and negative behavioral-informative feedback (See Table 16). These models revealed that positive behavioral-informative feedback was positively related to classroom quality, such that more of this feedback type was indicative of higher quality classrooms ($B = .13$, $p < .001$). Positive academic-informative feedback was also positively related to classroom quality, again such that more of this type of feedback was indicative of higher quality classrooms ($B = .33$, $p < .001$). Additionally, negative academic-informative feedback was negatively related to classroom quality: the more of this type of feedback given by the teacher, the lower quality the classroom environment tended to be (B -.046, $p = .012$). Finally, supportive feedback was positively related to classroom quality, such that higher quality classrooms had more supportive feedback given by the teacher ($B = .065$, $p < .001$). Estimations of the proportions of reduction in variance in Q-CLE attributable to each of the feedback factors were larger in these cases, ranging from .28 to .33. The b pathway of the proposed mediation model was significant in the cases of positive behavioral-informative feedback, positive academic-informative feedback, negative academic-informative feedback, and supportive feedback.

Table 16.

Effects of each of the teacher feedback factor variables on classroom quality, controlling for total minutes of instruction and assignment to ISI intervention.

	Constant/ Coeff.	SE	Sig.	R-Squared
Q-CLE	-.01			
BvInfoPos	.13	.03	<.01	.30
Q-CLE	.08			
BvInfoNeg	-.03	.01	.07	.29
Q-CLE	.03			
AcInfoPos	.33	.07	<.01	.32
Q-CLE	.20			
AcInfoNeg	-.05	.02	.01	.28
Q-CLE	-.23			
Support	.07	.02	<.01	.33

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

Because teacher affect was a consideration in the measurement of both teacher feedback patterns and Q-CLE (represented in the "warmth, responsiveness, control & discipline" component of the Q-CLE rubric) correlations were run investigating the relations among these four feedback factors and the three sub-components of the Q-CLE rubric (see Table 17). These analyses revealed that, except in the case of negative academic-informative feedback, relations between teacher feedback and Q-CLE were stronger when Q-CLE was considered as a comprehensive factor including all three indices of classroom quality. Both positive academic and behavioral-informative feedback factors showed no significant relations to the sub-components of the Q-CLE rubric but instead were only significantly related to the overall Q-CLE factor. The factor representing supportive feedback showed small but significant relations to the Q-

CLE sub-components of ‘implementation of instruction’ and ‘warmth, responsiveness, control, discipline’, however the relation of this feedback factor to the overall Q-CLE factor was much stronger. Negative academic-informative feedback was significantly related to the sub-component of ‘warmth, responsiveness, control, discipline’ but not to either of the other two Q-CLE indices. This feedback factor was also related to the overall Q-CLE factor, and these two relations were relatively the same size. It could be that in this case, the relations detected in predictive models between negative academic-informative feedback and Q-CLE were mostly attributable to this factor’s unique relation to ‘warmth, responsiveness, control, discipline’ in the measurement of Q-CLE.

Table 17.

Correlations among feedback factors and sub-components of Q-CLE factor.

	1	2	3	4	5	6	7	8
1. AcInfPos	**1**							
2. AcInfNeg	-.25**	**1**						
3. BvInfPos	.08	-.11	**1**					
4. Support	.4**	-.07	.53**	**1**				
5. IndImp	-.08	-.01	.01	-.12*	**1**			
6. OrientOrg	-.02	-.01	0	-.06	.65**	**1**		
7. W/R/C/D	-.01	-.14*	-.09	-.16**	.40**	.32**	**1**	
8. Q-CLE	.17**	-.13*	.30**	.31**	-.13*	-.14*	-.03	**1**

Note: Please refer to Feedback Shorthand guide in List of Symbols/Nomenclature

IndImp: Individualized implementation of instruction
OrientOrg: Orientation & organization
W/R/C/D: Warmth, responsiveness, control & discipline

*Correlation is significant at the < .05 level
** Correlation is significant at the < .001 level

Mediation Model (RQ 5)

In reviewing the significance of the individual mediation model pathways, it becomes clear that a mediation effect of teachers' feedback on the relation between teachers' depression and classroom quality is possible in the cases of positive behavioral-informative and positive academic-informative feedback, as these variables showed significant relations across all three pathways. To test for mediation, a multiple regression model was run with both predictors, teachers' depression and feedback, predicting classroom quality, controlling for total minutes of instruction. A mediation effect was determined by the potential change in significance of the original relation between teachers' depression and classroom quality once variation of each of the potential feedback mediators was accounted for.

The model testing positive behavioral-informative feedback as a mediator revealed no mediation effect. The effect of teachers' depression on classroom quality remained stable with regard to its coefficient size and significance ($B = -.038$, $p < .001$), whereas the effect of this type of feedback was no longer significant ($B = .015$, $p = .656$). The model testing positive academic-informative feedback as a mediator revealed a partial mediation effect: the original effect of teachers' depression remained significant ($B = -.033$, $p < .001$), however the effect of positive academic-informative feedback over and above the teacher depression effect was also significant ($B = .347$, $p < .001$). This model's effect size estimate was .44, suggesting that teachers' depressive symptoms and positive academic-informative feedback, along with the included covariates, accounted for about 44% of the variance in Q-CLE. This suggests that a significant portion of the

variance in the relation between teachers' depression and classroom quality can be accounted for by teachers' level of provision of positive academic-informative feedback.

CHAPTER 4. DISCUSSION

Overview of Findings

The purpose of this study was to apply and test a newly developed classroom observation system that examines the nature of teachers' feedback to students, and to use these data to investigate the extent to which teachers' feedback was associated with student math and behavior outcomes. Additionally, this study sought to investigate whether teachers' feedback differed as a function of their self-reported depressive symptoms, and how these variables might be associated with the quality of the classroom learning environment. Factor analyses suggested that patterns of feedback observed were best characterized by positive and negative affect, and to a lesser extent by the nature of the information being communicated (informative vs. supportive). Overall, students tended to receive more negative feedback than positive in the cases of both academic and behavioral informative feedback, and many teachers provided very little or none of these feedback types. Students received comparatively more supportive feedback, and this type was the most stable in terms of distribution in the data. Finally, estimates of the proportion of variance at each level of the data suggested that the vast majority of variability was attributable to between-classroom (or teacher) differences in feedback patterns, rather than individual differences in the students receiving feedback. Important to note for the interpretation of the discussed findings, student participants made expected gains from fall to spring on math outcomes.

Direct relations were revealed between teachers' negative academic-informative and negative behavioral-informative feedback and student math and behavior outcomes. Students who received more negative academic-informative feedback generally showed

weaker performance in mathematics, and students who received more negative behavioral-informative feedback had more behavior problems. An interaction effect was also detected such that students who began the year with weaker math skills exhibited stronger spring math achievement when they received more positive academic-informative feedback.

Investigating relations among teachers' self-reported depressive symptoms, their feedback patterns, and the observed quality of the classroom learning environment (Q-CLE) revealed that teachers who reported more depressive symptoms provided less positive behavioral-informative, negative behavioral-informative, and positive academic informative feedback to students. Teachers' feedback was also related to Q-CLE; higher quality classrooms were characterized by less positive behavioral-informative feedback, more positive academic-informative feedback and more supportive feedback. Finally, a partial mediation effect was detected in the case of positive academic-informative feedback: teachers who reported more depressive symptoms provided less of this type of feedback, and this effect was partially responsible for a decrease in Q-CLE.

Limitations

There are limitations to this study that should be considered when interpreting these findings. First and foremost, the Teacher Feedback Coding System central to this study is the first iteration of a novel measure that has not undergone empirically based testing and refinement. Though applying this measure in order to inform such refinement was a primary aim of this study, it should be recognized that observed feedback within the classroom may not have been represented with as much validity in this study as it will

be in future work involving improved versions of this coding system. This is especially true in the cases of positive and negative behavioral-informative feedback, which showed poor fit to the data in factor analyses. Codes involving these factors will be revisited and changed and/or clarified using the information gained in the present study with the goal of more accurately capturing the nature of teachers' feedback in the classroom.

Secondly, teachers' observed affect while interacting with students was contributed to the measurement of their feedback patterns as well as to the measurement of Q-CLE. Because both of these measures included affect judgments, it is possible that these two variables are conflated. In order to investigate this, correlations were run examining the relations between teachers' feedback and each of the three sub-components of the Q-CLE rubric, as well as the relations of these factors to the overall Q-CLE factor (see Table 17). Results of this analyses revealed that, in the case of negative academic-informative feedback, conflation may have existed. The direct relation detected between this feedback type and Q-CLE should be interpreted with this potential limitation in mind. However, we are confident that the patterns observed in these correlations are evidence that this is the only case where this might be an issue. The other feedback factors that were found to directly contribute to Q-CLE were either not significantly correlated to the sub-components of Q-CLE (positive academic and behavioral-informative feedback), or showed relations to multiple Q-CLE sub-components but stronger relation to the overall Q-CLE factor (supportive feedback).

Additionally, the adapted version of the CES-D that was used to measure teachers' self-reported depressive symptoms was altered slightly from its original version. The wording of some questions was changed and two questions were removed

due to principal concerns about their highly sensitive nature. Further, these questions were mixed into a larger self-efficacy survey whose questions could have influenced the way teacher participants answered the CES-D target questions. While this measure displayed high reliability within this sample, future research efforts would benefit from using the original versions of clinical screening tools to measure these teacher characteristics.

Further, while this study involved an adequate number of student participants (N=310), the number of teacher participants was small (N=31). Thus, this study's level-2 (classroom) analyses were underpowered, especially in the case of the final mediation model. Low power at this level means that variation in level-2 data may have been underestimated. Given this, it is encouraging that significant results were detected, however future work will seek to verify these findings within a sample that includes more teachers/classrooms. Finally, classrooms were exposed to instructional interventions. While the direct influence of these interventions was controlled for in analyses, there could have been indirect influences of these interventions on aspects of focus within the present study. These limitations should be considered when interpreting results, and conclusions drawn should be done so with caution.

Teacher Feedback and Student Outcomes

It was hypothesized that feedback characterized as more elaborative and positive would be predictive of stronger academic and behavioral performance in students, while more simplistic and negative feedback would predict poorer outcomes. Factor analyses did reveal factors that differentiated between positive and negative affect, however level

of specificity in feedback was not differentiated to the extent that was anticipated. The factors resulting from these analyses became the independent variables of interest in investigation into how teachers' feedback relates to student mathematics and behavior outcomes.

It was found that students who received more negative academic-informative feedback tended to have lower math achievement scores in the spring. This relation may suggest that this type of feedback is less effective as an instructional tactic, resulting in lower math achievement for students when they receive more of it. Alternatively, it could also be that when teachers are faced with having to help a student who does not understand a math concept, they are more likely to respond with negative affect when communicating feedback. Some temporal precedence does exist in this relation since feedback in the winter was predicting spring outcomes, however it is important to recognize the possibility of reciprocal effects.

Additionally, students who began the year struggling in math improved in their math skills when they were provided more positive academic-informative feedback (interaction effect). Together, these findings suggest that academic-informative feedback provided by the teacher has different impacts on students based on the type of affect displayed by teachers during these feedback events, as well as the incoming skills of students experiencing them. While the information being conveyed may be the same, the differences in affect appear to predict different student achievement outcomes. Specifically, these results suggest that negative academic-informative feedback may not be the most effective way to support students in their math learning, whereas positive academic-informative feedback might be a useful tool to build students' math skills,

especially in the cases of students who already struggle with mathematics.

This finding is especially relevant to the foundational study upon which this study expands (McLean & Connor, 2015), which found that the impact of teachers' self-reported depressive symptoms on student math outcomes operated in the same type of interaction effect as was detected with positive academic-informative feedback. Students who began the year struggling in math appeared to be especially sensitive to the negative influence of teachers' depressive symptoms, and this same subgroup of students was also found to experience more positive growth in mathematics when exposed to more positive academic-informative feedback. When considered together, these findings may suggest that one reason academically at-risk students whose teachers are struggling with more depressive symptoms show decreased academic growth across the year is because these teachers are less likely to provide a type of feedback that has been found to be especially helpful to these students.

Adding to these findings, it was revealed that more negative behavioral-informative feedback from teachers predicted more problem behaviors in students. Because both of these variables were measured in the winter of the school year, no temporal precedence exists and directionality of this relation can only be assumed and may be reciprocal. It is possible that when teachers are faced with students who exhibit more problem behaviors, they tend to become more negative in their behavioral feedback because of the more difficult and frustrating nature of interacting with these students. It could also be that students who receive more negative behavioral-informational feedback are somehow less inclined to attempt to improve their behavior. Future work will attempt to better capture the directionality of this relation in order to inform how these variables

are interacting with each other. Overall, the effect sizes detected in the cases of teachers' feedback directly predicting student outcomes were small, but fell within the normal range of effects commonly detected within educational data.

Depressive Symptoms, Teacher Feedback, and Classroom Quality

Hypotheses involving these variables predicted that teachers struggling with more depressive symptoms would be more simplistic and negative in their feedback to students, rather than elaborative/supportive and positive. Further, it was predicted that high quality classrooms would be characterized by more elaborative/supportive and positive feedback patterns, whereas low-quality classrooms would be characterized by more simplistic and negative feedback. Finally, it was hypothesized that these feedback patterns would mediate an already-established relation between depressive symptoms and the Q-CLE. While factor analyses did not reveal factors accounting for feedback specificity to the extent that was hypothesized, results from correlation analyses and the proposed mediation model suggest some interesting relations among these variables.

Overall, teachers who were struggling with more depressive symptoms showed less 'enthusiastic/exuberant' and 'content' affect in their feedback interactions with students, yet no such relations existed between depressive symptoms and observed negative affect types ("neutral", "angry/frustrated", "sad/depressed"). These findings suggest that a teacher who is struggling with depression may not necessarily display more negative affect within the classroom, as was hypothesized. However, she may alternatively be less likely to display positive affect in her interactions with students. This conclusion is generally supported in clinical depression literature, which

characterizes depression more as a dampening of positive emotions rather than an over-abundance of negativity (American Psychiatric Association, 2000).

Investigation into the hypothesized mediation model revealed that teachers who reported more depressive symptoms provided less of both positive and negative behavioral-informative feedback to their students. While these relations were significant, the effect sizes estimated in these relations were very small. Teachers' depression only accounted for between 2% and 5% of the variability in their feedback patterns. These small effects, though significant, could signify that teachers' depression may operate more strongly in other ways within the classroom, in addition to surfacing in their feedback patterns. There are multiple ways that these relations could be operating, but because no temporal precedence was established in measurement (all variables were measured at the same time point), interpretation is speculative. It might be that teachers who are struggling with common depressive symptoms such as depleted energy levels and adverse reactions to daily stressors are not making as many attempts to manage student behavior when compared to their colleagues who are not struggling with these symptoms. It is important to acknowledge, however, that a portion of these relations could also be accounted for by student characteristics such as the proportion of problem-behavior students in the classroom (Skibbe, Phillips, Day, Brophy-Herb & Connor, 2012)

Teachers struggling with more depressive symptoms also tended to provide less positive academic-informative feedback. This is a particularly important finding when considered along with the finding that this type of feedback was found to predict stronger math achievement for students who started the school year with weaker math skills. Considered concurrently, it appears that one of the reasons low-performing students may

continue to struggle when paired with teachers experiencing more depressive symptoms is because these teachers provide less positive academic-informative feedback. This finding speaks to the specific mechanisms through which teachers' depression might be operating to impact student outcomes, in this case their likelihood of providing effective feedback instruction to students, especially those who need it most.

Examining the relations between the teacher feedback factors and Q-CLE also revealed some interesting trends in terms of how "high-quality" classrooms might be characterized when considering feedback. While teachers' depressive symptoms were negatively related to positive academic and behavioral-informative feedback, classroom quality showed an opposite pattern: higher quality classrooms tended to have more of these types of feedback. Additionally, higher quality classrooms were characterized by more supportive feedback. This same pattern was detected when comparing relations between teachers' depressive symptoms and the five affect categories and classroom quality's relations to these types of affect. Whereas teachers who reported more depressive symptoms seemed to have less 'enthusiastic/exuberant' and 'content' affect, teachers whose classrooms were rated as higher-quality used more of these types of affect. Q-CLE was also negatively related to neutral affect, a category that was grouped with overall "negative" affect in factor analyses.

Looking further, higher quality classrooms were indeed characterized by increased positive behavioral and academic informative feedback, as well as by more supportive feedback. Alternately, teachers with higher Q-CLE were less likely to provide negative academic-informative feedback. Effect sizes estimated for these relations were larger than what were detected for teachers' depressive symptoms predicting their

feedback. Feedback patterns accounted for between 28% and 33% of the variability observed in Q-CLE. The differences in effect size between these paths of the proposed mediation model suggest that the strongest relations between variables of interest are those that were revealed between feedback and Q-CLE. Interpreting these findings, these patterns suggest that the nature of the interactions that take place between teachers and their students not only impact student outcomes, but also contribute to the larger, dynamic microsystem of the classroom environment (Hamre & Pianta, 2005; Pianta, La Paro, Payne, Cox & Bradley, 2002). More specifically it appears that affect, or emotional tone, is an important feature of instruction that should be considered by teachers in the interest of optimizing classroom quality.

Results of the final mediation model suggested that teachers struggling with more depressive symptoms tended to utilize less positive academic-informative feedback, and that this was partially responsible for a decrease in classroom quality. What is especially important to consider is how these relations might impact students' learning experiences. Results of multilevel modeling revealed that positive academic-informative feedback was especially effective in improving math achievement for low-achieving students. Further, the interaction effect detected in McLean & Connor (2015) suggested that teachers' depressive symptoms negatively impacted the math achievement of this same subset of struggling students. It appears, based on these findings, that teachers struggling with depression may be less likely to provide the type of feedback that is the most important for struggling students, and that this may be why low-achieving students of more depressed teachers are particularly impacted.

Informing Revision of the Teacher Feedback Coding System

It was predicted that factor analysis would reveal factors accounting for feedback type (academic vs. behavioral), specificity (level of elaboration), and affect (positive vs. neutral vs. negative). These hypotheses were supported in the data to some extent, but differentiation between factors did not occur exactly as was predicted. In terms of feedback specificity, or how much teachers elaborated in their feedback to students, no significant differences were detected between more simplistic feedback (i.e. 'identifications' and 'corrections') and feedback where the teacher provided more information ('elaborations'). Instead, differentiation took place at the level of feedback type only (academic vs. behavioral), and all feedback that involved some type of communication of information was observed to group together into one larger category of "informative" feedback, which was then further characterized by either positive or negative affect. While the hypothesis that academic and behavioral feedback would be found to be different from each other in these cases was supported, findings regarding the differentiation of feedback specificity within these categories did not align with what was predicted.

These patterns are consistent with past findings (Mandernach, 2005; Pridemore & Klein, 1995; Smits, Boon, Sluijsmans, & van Gog, 2008) that detected similar grouping patterns in teachers' feedback types, although such studies have not attempted to conceptualize feedback as specifically as is done in this coding system. It could be perhaps that the approach taken by this coding system did not adequately capture the varying techniques that teachers employ during feedback opportunities, or it could also indicate that level of specificity given by the teacher is not a particularly important aspect

of feedback (discussed below). Future iterations of this coding system might benefit from trying to capture and asses more specific teacher moves during feedback events, in an attempt to provide more detailed information about what strategies teachers are using during feedback. For example, including codes that indicate whether a teacher is using the same or different instructional tactics to explain a missed concept to a student during a feedback event (i.e. are they 're-explaining' or are they finding a new way to present the material?) could offer more insight into the different types of elaborations that teachers use and whether those differentiate from each other.

Supportive feedback, characterized by both academic and behavioral encouragements and support statements, as well as positive academic elaborations, was identified as a different type of feedback than informative feedback, however this type of feedback did not differentiate between academic and behavioral. This finding suggests that all moves by the teacher to catch and reinforce positive student actions, both behavioral and academic, indicate a unique type of instructional tactic that is characteristically different from a teachers' communication information. The fact that both academic and behavioral codes grouped together to indicate an overall 'support' factor was contrary to what was initially hypothesized, however the hypothesis that supportive feedback would be found to be characteristically different from the communication of information to students was supported. A particularly interesting finding was that positive academic elaborations were included in this 'supportive' feedback, as opposed to being grouped with academic-informative feedback. Coders noted during video observations that the majority of positive academic elaborations were in the form of purposeful, guided questioning by the teacher with the goal of helping the

student infer the correct answer on their own, instead of direct communication of information. This could suggest that when teachers are assisting a student using such techniques, and allowing them to "find" the answer more independently through the use of these guided questions, they are actually providing support rather than direct instruction.

Follow-up factor analysis using CFA provided a more rigorous and in-depth look at the validity of the patterns revealed in EFA. The 1-factor model representing supportive feedback was shown to fit the data very well, as indicated by the model fit indices. This model was compared to a 2-factor model with academic and behavioral support modeled separately, and fit for this second model was comparatively much poorer. Despite the very good fit of the 1-factor model, closer examination of the factor loadings of each included feedback code showed that behavioral encouragements did not load strongly or significantly onto the 'support' factor. This is consistent with what was revealed in EFA analyses (see table 6), and suggests that behavioral encouragements by the teacher are somehow characteristically different from the other types of support that were found to load strongly onto the 'support' factor. This feedback code was removed from the model but its removal did not result in any improvement of model fit, indices remained the same between the two models.

Academic-informative feedback was modeled in CFA using a 2-factor model, with positive and negative academic-informative feedback each representing a latent factor. Positive identifications and corrections were indicated by an "academic-informative-positive" factor, and negative identifications, corrections and elaborations were indicated by the "academic-informative-negative" factor, and the two latent factors

were allowed to correlate. This model fit the data very well, and all feedback codes loaded significantly onto their respective factors. This model was compared to a 1-factor model in which all academic-informative feedback loaded onto a single factor, disregarding affect categorization, and model fit decreased significantly. Thus, these analyses suggest that the 2-factor model for academic-informative feedback is a stronger representation of the data.

The final model run in CFA was a 2-factor model representing behavioral-informative feedback. This model included latent variables indicating positive behavioral-informative and negative behavioral-informative feedback. The positive behavioral-informative latent factor was comprised of positive redirections and elaborations, while the negative behavioral-informative feedback was comprised of negative redirections and elaborations. These two latent variables were allowed to correlate with each other, however no significant correlation was detected. Further, this model fit the data poorly, as indicated by all model fit indices. In looking more closely at the factor loadings for each latent variable, it appears that the 'behavioral-informative-positive' factor does indeed hold significant influence over it's related codes, with strong and significant loadings in both cases. However, this is not the case with the 'behavioral-informative-negative' latent factor. The significant factor loading for negative redirections was smaller than is ideal, and the factor loading for negative elaborations was small and non-significant. It is possible that the poor fit of the model comes from this latent variable.

Considered together, results of these analyses provide guidance for further refinement of the Teacher Feedback Coding System. Firstly, the finding that positive

academic elaborations were found to be better indicated by 'supportive' feedback rather than 'academic-informative' feedback may suggest a need to redefine how academic elaborations are conceptualized within this coding system. It might also be prudent to identify a new type of supportive feedback that better captures the guided questioning that was coded in this iteration of the system as academic elaborations. In doing this, coders might be able to differentiate between these two feedback techniques and provide the system with more specified information that better captures both supportive and informative feedback.

The finding that behavioral elaborations were not well-indicated by the 'support' feedback factor (or any other factor in EFA analyses) may suggest that this type of feedback might be relatively unimportant or unrelated to the other types of feedback represented by the coding system. Upon reflection of actual coding, these moves were usually made very casually and quickly by the teacher (i.e. more of a "filler" instead of a purposeful statement), and students often did not show any observable reactions to these moves. In contrast, teachers' attempts to provide academic encouragement to students were more pointed, and students' reactions to these moves were more easily observable. Future iterations of this coding system will likely either remove behavioral encouragements, or redefine their conceptualization to capture more purposeful statements by the teacher that are meant to encourage good behavior, but that are not pointed/enthusiastic enough to warrant coding a support statement.

Finally, the lack of model fit for factors indicated by behavioral-informative feedback provide evidence that this portion of the coding system needs revision. While positive behavioral-informative feedback seemed valid as a separate factor, the negative

behavioral-informative feedback factor did not represent the data well. Specifically, the negative behavioral elaboration codes were where much of the fault lay with this factor. A thorough revisit of how these codes are conceptualized is needed in this case, as these codes are not consistently representing their intended feedback styles. It could be that some codes are too broad and need to be split into separate feedback codes, or it could be that some need to be removed entirely from the system. This particular aspect of the coding system will be carefully addressed in further iterations, and all results involving negative behavioral-informative feedback should be interpreted with caution.

Implications for Policy and Practice

The patterns revealed in factor analyses provided some practical insight into how feedback operates within the classroom that may be useful when considering policy and practice. There was no indication that attempts made by the teacher to elaborate in their feedback to students was characteristically different from simply pointing out a mistake or providing an alternative answer. One possible conclusion that could be drawn from this finding is that teacher attempts to convey more 'elaborative' information may instead be more effective when used in other instructional contexts, rather than in feedback events. This approach could be taken with either academic or behavioral instruction within the classroom. It would stand to reason that the students of a teacher who is more elaborative in her initial instruction will better understand the material from the beginning, and will therefore need less feedback later on. If a teacher does not provide enough information while introducing a new topic, however, she may find herself needing to elaborate more in later feedback events when some students show a lack of

understanding. This applies to behavior management as well; if a teacher provides elaborative behavioral instruction before misbehavior happens, such as taking time to review classroom ground rules for behavior, students may benefit more in terms of positive behavior than they would from behavioral elaborations that follow student actions (feedback events).

In terms of the grouping of affect codes, it was surprising that 'neutral' feedback tended to group more strongly with negative affect ('sad/depressed' and 'angry/frustrated'), instead of standing on its own as was hypothesized. This particular finding may suggest that, while 'neutral' affect can in some contexts be interpreted as just that (i.e. simply a lack of extreme emotion of any kind), neutral emotionality within the context of the classroom may be perceived as more of a negative emotional state by students and/or outside observers. During the actual coding of video data, coders noted that the majority of affect assignments were either 'neutral' or 'content', and that more extreme emotions (both positive and negative) were comparatively less common. It was also noted during coding that individual teachers tended to be very consistent in their affect across video observations, establishing strong patterns of observable affect (confirmed in ICC calculations) that mainly included one or two primary affect codes. These patterns warrant future investigation, but could suggest that interventions aimed at improving teachers' feedback may need to include training in identifying and potentially adjusting affect patterns within the classroom. Such training could focus on purposefully maintaining positive affect during instruction, as well as recognizing one's own tendency to display a certain type of affect and making adjustments accordingly.

More generally, results of this study can be used to inform instructional

interventions and recommendations for best teaching practices within elementary classrooms. Overall, findings suggest that when it comes to teachers' provision of feedback to students, it is not only what is said that is important, it is how it is said. This is directly supported in results regarding positive and negative academic informative feedback, where negative feedback was negatively related to student outcomes and positive feedback positively impacted students who were already struggling. These patterns were mirrored in the relation of feedback types to the more global construct of classroom quality: feedback characterized as positive indicated higher quality classrooms, whereas negative feedback indicated the opposite. Overall, it appears that teacher affect is an important contributor to both individual student outcomes and to larger classroom-level constructs such as Q-CLE.

As discussed previously, a vast majority of the variability observed in teachers' feedback was attributable to teacher-level differences, suggesting that teachers were consistent in their patterns of feedback, rather than feedback style being dependent on the student involved in the interaction. This suggests that intervention at the teacher level might be an appropriate approach to improving feedback within classrooms. Teacher training that focuses identifying one's own patterns of feedback (and more specifically, affect during feedback) and making adjustments to these could lead to an improvement in feedback quality within the classroom. Taking this a step further, teachers can use results from this study to inform how best to approach giving feedback to students who are struggling academically. Past research has indicated that efforts made by teachers to individualize instruction based on student needs are effective in supporting student learning (Connor et al., 2010; 2011). Because positive academic informative feedback

was found to be helpful for students who began the year struggling in math, interventions that aim to individualize instruction could include a focus on providing more of this type of feedback to the students who need it most.

Additionally, professionals in the field can use the information herein regarding depressive symptoms, feedback, and classroom quality to better understand how teacher characteristics may be influencing their contributions to the classroom environment. By revealing some of the ways that teachers struggling with these depressive symptoms differ in their instructional tactics, the field can better identify, understand and subsequently address how practitioners' affective characteristics are influencing their students. Results suggested that teachers who are struggling with depressive symptoms are not only less likely to provide a type of feedback found to be helpful to struggling students, but were also more generally less likely to display positive affect within the classroom. Implications of these patterns surfaced both in terms of student academic growth, and the overall quality of the classroom environment. Future efforts, either through continued research or intervention, can benefit from this information because it provides a more clear, detailed picture of what a teacher struggling with depression might look like, how these characteristics operate to influence student experiences and outcomes, and what steps might be taken to intervene in the hopes of improving the classroom experiences of both teachers and their students.

Lastly, the findings regarding the contributions of teachers' feedback to observed classroom quality could serve to better inform the systematic observations of classrooms that take place in almost all educational settings. Since, as previously discussed, recent research has shown limited consistency or precision in the evaluation of teachers and

classrooms based on expert observations (Strong, Gargani & Hacifazlioğlu, 2011), there is an abundant need for research-proven recommendations of what to look for when assessing a classroom using observation, especially given the extreme implications of such assessments of teacher performance. Such recommendations regarding specific teacher moves and observable characteristics that are indicative of higher-quality classrooms have the potential to lead to better accuracy in classroom observation, and thus in the assessment of teachers. Specifically, this study indicates that those assessing classrooms for quality should look for feedback that is positive and supportive as evidence of a high-quality environment, and feedback that is negative as evidence of the latter.

Implications for Theory and Future Research

The theoretical framework utilized for this study focused on Dynamic Systems theories (Yoshikawa & Hsueh, 2001) and the Bio-Ecological Model of development (Bronfenbrenner & Morris, 2006) to conceptualized the classroom as a complex microsystem in which multiple developmental systems interact synergistically to influence child development. Applied to this study, it was hypothesized that the characteristics brought into the classroom by teachers, namely their depressive symptoms, would impact the nature of the instructional interactions they have with their students on a regular basis. These things would, in turn, influence both student development at the individual level, as well as contribute to the larger systems that impact the student experience within the classroom, in this case Q-CLE. The results supported this theoretical approach, as patterns and relations were detected among

variables of interest that went beyond simple direct predictive relations. Teachers' depressive symptoms were found to impact both the overall Q-CLE, as well as the nature of the smaller, repeated interactions that took place between teachers and their students during feedback. Further, feedback types were found to relate to the larger construct of Q-CLE, as well as to student outcomes both directly and through interaction effects. Importantly, types of feedback that were not found to be predictive of student outcomes were in fact found to contribute to overall Q-CLE, suggesting that these types of feedback may have a more indirect impact on student experiences and outcomes.

One important aspect of the foundational theory that was not captured by this study was the possibility of reciprocal effects between primary variables. A key element of the Bio-Ecological model is that no developmental process is ever conceptualized to be purely unidirectional. Rather, variables interact dynamically with each other, potentially forming reciprocal loops. For example, within the context of this study it could very well be that some teachers come into the classroom struggling with depressive symptoms, but due to classroom factors such as a high proportion of at-risk students these depressive symptoms become more aggravated, causing teachers' instructional quality to suffer, thus putting students even more at risk for failure, thus further aggravating depressive symptoms, etc. It is important to recognize these potential relational patterns when considering how the variables of interest in this study might be relating to each other in ways that were not directly investigated.

Future research endeavors following this study will seek to further refine the Teacher Feedback Coding System with the goal of making it a reliable observational tool that can be used to contribute important knowledge to the field. Future iterations of this

system will attempt to better capture behavioral feedback, as well as expand it's coding of feedback in ways that will provide more information about the strategies teachers use to communicate information to students. Additionally, this system will be applied to different age groups of students in order to ascertain whether teachers' feedback impacts students differentially based on age (for example, younger students may be more sensitive to their teachers' feedback patterns). Following this vein, the relations of other student characteristics, such as gender, to teachers' feedback can be investigated in order to provide more in-depth information about whether or not teachers apply certain types of feedback to certain types of students, and what the potential educational implications of these patterns may be.

Further, this study expanded on foundational, exploratory work examining the impacts of teachers' depressive symptoms on multiple classroom and student processes, but there are still many steps to be taken to fully define these relations. Future work will seek to measure teachers' affective characteristics more precisely, and will relate these characteristics to a wider variety of outcomes such as teacher-student relationship quality, instructional quality (proactive instruction as opposed to reactive feedback), and various student characteristics. One particular question that future work will seek to answer is that of the directionality of the relations that have been observed so far. Do teachers come into the classroom struggling with these symptoms? Or do student characteristics somehow contribute to a teachers' vulnerability to chronic stress, burnout, etc.? In conducting more purposeful investigations into issues such as these, this line of research could provide valuable insight into the struggles experienced by elementary teachers, as

well as inform how best to assess, support, and train early childhood educators with the goal of improving the educational experiences of the children they teach.

REFERENCES

Achenbach. T. M. (1986). The Direct Observation Form of the Child Behavior Checklist (rev. ed.). Burlington. VT: University of Vermont. Department of Psychiatry.

Achenbach. T M.. & Rescorla. L. A. (2001). Manual for The ASEBA School-Age Forms & Profiles. Burlington. VT: Research Center for Children, Youth, and Families.

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington, D.C.

Atwell, N. (1998). In the middle: New understandings about writing, reading, and learning. Portsmouth, NH: Heinemann.

Baker, J. A. (2009). Contributions of teacher-child relationships to positive school adjustment during elementary school. Journal of School Psychology, 44(3), 211–229.

Barnett, W. S., Epstein, D. J., Friedman, A. H., Boyd, J. S., & Hustedt, J. T. (2008). The state of preschool 2008: State preschool yearbook. New Brunswick, NJ: National Institute for Early Education Research, Rutgers University.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy, & Practice, 5, 7–74.

Bohn, C. M., Roehrig, A. D., & Pressley, M. (2004). The first days of school in the classrooms of two more effective and four less effective primary-grades teachers. Elementary School Journal, 104(4), 269-287.

Borg, W. R. (1979). Teacher coverage of academic context and pupil achievement. Journal of Educational Psychology, **71**, 635–645.

Borko, H., & Niles, J. (1987). Descriptions of teacher planning: Ideas for teachers and research. In V. Richardson-Koehler (Ed.), Educators' handbook: A research perspective (pp. 167-187). New York, NY: Longman.

Boyd, D., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. Future of Children, 17, 45–68.

Bratcher, S. (with Ryan, L.) (2004). Evaluating children's writing: a handbook for grading choices for classroom teachers. 2nd edition. Mahwah, NJ: Lawrence Erlbaum.

Briner, R., & Dewberry, C. (2007). Staff well-being is key to school success. Retrieved from http://www.scribd.com/doc/34205400/Staff-wellbeingis-key-to-school-success-Full-Report

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), Handbook of child psychology: Theoretical models of human development (6th ed., Vol. 1, pp. 793-828). Hoboken, NJ: John Wiley & Sons.

Brophy, J.,& Evertson, C. (1976). Learning from teaching: A developmental perspective. Boston: Allyn & Bacon.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), Handbook of research on teaching (pp. 328–375). New York: Macmillan.

Burchinal, M., Howes, C., Pianta, R. C., Bryant, D., Early, D. M., Clifford, R. M., et al. (2008). Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher–child interactions and instruction. Applied Developmental Science, 12, 140–153.

Burnett, P.C. (2002). Teacher Praise and Feedback and Students' Perceptions of the Classroom Environment. Educational Psychology, 22:1, 5-16.

Butler, A. C., Karpicke, J. D., & Roediger, H. L., III. (2008). Correcting a meta-cognitive error: Feedback enhances retention of low confidence correct responses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34, 918–928.

Cameron C.E., Connor C.M., Morrison F.J., Jewkes A.M. (2008). Effects of classroom organization on Letter–Word reading in first grade. Journal of School Psychology. 6, 173–192.

Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. Journal of School Psychology, 43, 61–85.

Carlisle, J F., Kelcey, B., Berebitsky, D., & Phelps, G. (2011). Embracing the complexity of instruction: A study of the effects of teachers' instruction on students' reading comprehension. Scientific Studies of Reading, 15(5), 409-439.

Caspary, K. (2002). California's pioneering training and retention initiative for early childhood educators. Berkeley, CA: PACE.

Chang, M. (2009). An Appraisal perspective of teacher burnout: Examining the emotional work of teachers. Educational Psychology Review, 21, 193–218.

Chen, Y., Thompson, M.S., Kromrey, J.D., Chang, G.H., (2011). Relations of student perceptions of teacher oral feedback with teacher expectancies and student self-concept. The Journal of Experimental Education, 79, pp. 452-477.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. Economics of Education Review, 26, 673–682.

Cochran-Smith, M., & Lytle, S. L. (1990). Research on teaching and teacher research: The issues that divide. Educational Researcher, 19(2), 2-11.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37-46.

Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., & Morrison, F. J. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. Journal of Educational Psychology, 106, 762-778. DOI: 10.1037/a0035921.

Connor, C.M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized control study on the accumulating effects of individualized literacy instruction on students' reading from 1st through 3rd grade. Psychological Science, 24(8), 1408-1419.

Connor, C.M. (2012). Literacy instruction in third grade: Amount and quality of instruction predicts students' outcomes. Scientific Studies of Reading (in press).

Connor, C.M., Morrison, F.J., Fishman, B., Giuliani, S., Luck, M., Underwood, P., & Schatschneider, C. (2011). Classroom instruction, child X instruction interactions and the impact of differentiating student instruction on third graders' reading comprehension. Reading Research Quarterly, 46(3), 189-221.

Connor, C. M., Ponitz, C. E. C., Phillips, B., Travis, Q. M., Day, S. G., & Morrison, F. J. (2010). First graders' literacy and self-regulation Gains: The effect of individualizing instruction. Journal of School Psychology, 48, 433-455.

Connor, C.M., Morrison, F.J., Fishman, B., Ponitz, C.C., Glasney, S., Underwood, P., & Schatschneider, C. (2009). The ISI classroom observation system: Examining the literacy instruction provided to individual students. Educational Researcher, 38(2), 85-99.

Connor, C. M., Son, S. H., Hindman, A. H., and Morrison, F. J. (2005). Teacher qualifications, classroom practices, family characteristics, and preschool experience: complex effects on first graders' vocabulary and early reading outcomes. Journal of School Psychology, 43, 343–375.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), Handbook of human-computer interaction (2nd ed., pp. 849–874). New York, NY: Elsevier.

Craven, R.G., Marsh, H.W., & Debus, R.L. (1991). Effects of internally focused feedback and attributional feedback on enhancement of academic self-concept. Journal of Educational Psychology, 83, 17–27.

Curbow, B. (1990). Job stress in child care workers: A framework for research. Child and Youth Care Quarterly, 19, 215–231.

Curby, T.W., Rudasill, K.M., Edwards, T., Perez-Edgar, K. (2011). The role of classroom quality in ameliorating the academic and social risks associated with difficult temperament. School Psychology Quarterly 26(2): 175-188.

Darr, W., & Johns, G. (2008). Work strain, health and absenteeism: A meta-analysis. Journal of Occupational Health Psychology, 13, 293–318.

Dimsdale, J. E. (2008). Psychological stress and cardiovascular disease. Journal of the American College of Cardiology, 51, 1237–1246.

Dohrn, E., & Bryan, T. (1994). Attributional instruction. Teaching Exceptional Children, 26, 61–63.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., & Klebanov, P. (2007). School readiness and later achievement. Developmental Psychology, 43, 1428-1446.

DuPaul. G. J.. Volpe. R. J., Jilendra. A. K.. Lutz. J. G.. Lorah. K. S.. & Grubner. R. (2004). Elementary school students with attention-deficit/hyperactivity disorder: Predictors of academic achievement. Journal of School Psychology. 42.

Early, D. M., Maxwell, K. L., Burchinal, M., Alva, S., Bender, R. H., Bryant, D., et al. (2007). Teachers' education, classroom quality, and young children's academic skills: Results from seven studies of preschool programs. Child Development, 78, 558–580.

Eccles, J. & Gootman, J.A. (2002). Community programs to promote youth development. Washington, DC: National Academic Press.

Edmonds, M., & Briggs, K. L. (2003). The Instructional Content Emphasis Instrument: Observations of reading instruction. In S. Vaughn & K. L. Briggs (Eds.), Reading in the classroom: Systems for the observation of teaching and learning (pp. 31–52). Baltimore: Brookes.

Epstein, M., Atkins, M., Cullinan, D., Kutash, K., & Weaver, R. (2008). Reducing Behavior Problems in the Elementary Classroom. National Center for Education Evaluation and Regional Assistance. U. S Department of Education.

Forness, S.R. & Esveldt, K.C. (1975). Classroom observation of children with learning and behavior problems. Journal of Learning Disabilities, Vol.8, No.6, pp. 382-385.

Fraser, B.J. (1998). Classroom environment instruments: Development, validity, and applications. Learning Environments Research, 1, 7–33.

Gage, N. L., & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. Elementary School Journal, 89, 253-300.

Good, T.L. & Brophy, J.E., (1970). Teacher-child dyadic interactions: A new method of classroom observation. Journal of School Psychology, Vol. 8 No. 2.

Good, T. L., Grouws, D. A., & Beckerman, T. (1978). Curriculum pacing: Some empirical data in mathematics. Journal of Curriculum Studies, 10, 75–81.

Greenberg, M. T. (2004). Current and future challenges in school-based prevention: The researcher perspective. Prevention Science, 5(1), 5–13.

Gresham, F. M., & Elliot, S. N. (1990). Social Skills Rating System manual. Circle Pines, MN: American Guidance Service.

Gudmundsdottir, S. (1991). Ways of seeing are ways of knowing: The pedagogical content knowledge of an expert teacher. Journal of Curriculum Studies, 23, 409-421.

Gunnar, M., & Quevedo, K. (2007). The neurobiology of stress and development. Annual Review of Psychology, 58, 145–173.

Hamre, B. K., & Pianta, R. C. (2004). Self-reported depression in nonfamilial caregivers: Prevalence and associations with caregiver behavior in child-care settings. Early Childhood Research Quarterly, 19, 297–318.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? Child Development, 76(5), 949-967.

Hamre, B. K., and Pianta, R. C. (2007). Learning opportunities in pre-school and early elementary classrooms. School Readiness and the Transition to Kindergarten in the Era of Accountability, Baltimore: Brookes, 49–84.

Hamre, B. K., Pianta, R. C., Downer, J. T., & Mashburn, A. J. (2008). Teachers' perceptions of conflict with young students: Looking beyond problem behaviors. Social Development, 17, 115–136.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. London and New York: Routledge.

Hattie J. (1998) Assessment and classroom learning: A deductive approach. Assessment in Education 5(1), 111–122.

Hill H. C., Charalambous C. Y., Kraft M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. Educational Researcher, 41, 56-64.

Hillerich, R. L. (1985). Teaching children to write, K-8: a complete guide to developing writing skills. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., et al. (2008). Ready to learn? Children's pre-academic achievement in pre-Kindergarten programs. Early Childhood Research Quarterly, 23(1), 27–50.

Hoyenga. K. B., & Hoyenga, K. T (1993). Gender-related differences: Origins and outcomes. Boston, MA: Allyn and Bacon.

Jacob, B. A., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance.

Jennings, P. A., & Greenberg, M. (2009). The prosocial classroom: Teacher social and emotional competence in relation to child and classroom outcomes. Review of Educational Research, 79, 491–525.

Jepsen, C. (2005). Teacher characteristics and student achievement: Evidence from teacher surveys. Journal of Urban Economics, 57, 302–319.

Johnson, S., Cooper, C., Cartwright, S., Donald, I., Taylor, P., & Millet, C. (2005). The experience of work-related stress across occupations. Journal of Managerial Psychology, 20, 178 –187.

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, 613–631.

Kao, C. (2013). Effects of focused feedback on the acquisition of two English articles. The Electronic Journal for English as a Second Language. Vol. 17, 1.

Kontos, S., & Keyes, L. (1999). An eco-behavioral analysis of early childhood classrooms. Early Childhood Research Quarterly, 14, 35–50.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. Educational Psychology Review, 1(4), 279-308.

Li-Grining, C., Raver, C., Champion, K., Sardin, L., Metzger, M., & Jones, S. (2010) Understanding and improving classroom emotional climate and behavior management in the "real world": The role of Head Start teachers' psychosocial stressors. Early Education and Development, Vol 21:1 pp. 65-94.

Mandernach, B. J. (2005). Relative effectiveness of computer-based and human feedback for enhancing student learning. The Journal of Educators Online, 2, 1–17.

Marsh, H.W. (1990). The influences of internal and external factors of reference on the formation of English and math self-concepts. Journal of Educational Psychology, 82, 107–116.

Mashburn, A. J., Hamre, B. K., Downer, J. T., & Pianta, R. C. (2006). Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. Journal of Psychoeducational Assessment, 24(4), 367–380.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O., Bryant, D., et al. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development, 79, 732–749.

Maslach, C., Schaufeli, W. B., & Lieter, M. P. (2001). Job burnout. Annual Review of Psychology, 52, 397–422.

McConaughy. S. H.. Kay. P. J.. & Fitzgerald. M. (1998). Preventing SED though parent-teacher action research and social skills instruction: First-year outcomes. Journal of Emotional and Behavioral Disorders. 6. 81 -93.

McEwen, B. S. (2008). Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. European Journal of Pharmacology, 583, 174–185.

McGaw, B., Wardrop, J.L., & Bunda, M.A. (1972). Classroom observation schemes: Where are the errors? American Education Research Journal. Vol. 9, No.1, pp.13-27.

McLean, L.E. & Connor, C.M (Under Review). Differences in Classroom Instruction Provided to First Graders with Problem Behaviors vs. Their Typically/Exceptionally-Behaved Peers. Elementary School Journal.

McLean, L.E., Sparapani, N., Toste, J. & Connor, C.M. (Under Review). Classroom quality as a predictor of first-graders' time in non-instructional activities and literacy achievement. Journal of School Psychology.

McLean, L.E. & Connor, C.M. (2015). Depressive symptoms in 3rd grade teachers: Relations to classroom quality and student achievement. Child Development. Volume 86 (3), pp. 945-9.

Meisels, S. J., & Atkins-Burnett, S. (2006). Evaluating early childhood assessments: A differential analysis.

Meyer, J.P., Cash, A.H., & Mashburn, A. (2011). Occasions and the reliability of classroom observations: Alternative conceptualization and methods of analysis. Educational Assessment, Vol. 16, No. 4, pp. 227-243.

Miller, R. T., Murnane, R. J., & Willett, J. B. (2007). Do teacher absences impact student achievement? Longitudinal evidence from one urban school district (Working Paper 13356). Cambridge, MA: National Bureau of Economic Research.

Morgan, C. (2001). The effects of negative managerial feedback on student motivation: implications for gender differences in teacher-student relations. Sex Roles, Vol. 44, pp. 514-535.

Mueller, C.M. & Dweck, C.S. (1998). Praise for intelligence can undermine children's motivation and performance. Journal of Personality and Social Psychology. Vol 75:1 pp. 35-52.

Murray L, Cooper P.J. (2007) Effects of postnatal depression on infant development. Arch Dis Child 1997; **77**:99-101.

NAEP. (2011). National Assessment of Educational Progress: The nation's report card. http://nces.ed.gov/nationsreportcard/

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00-4769). Washington, DC: National Institutes of Health, U.S. Department of Health and Human Services.

NICHD Early Child Care Research Network. (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. Elementary School Journal, 102, 367–387.

NICHD Early Child Care Research Network. (2007). Age of entry to kindergarten and children's academic and socioemotional development. Early Education & Development, 18(2), 337-368.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31, 199–218.

Neisworth, T., & Bagnato, S. J. (2004). The mismeasure of young children. Infants & Young Children, 17, 198–213.

Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B. Schieffelin (Eds.), Developmental pragmatics New York: Academic Press, 43-72.

O'Connor, E., & McCartney, K. (2007). Teacher-child relationships and achievement as part of an ecological model of development. American Education Research Journal, 44, 340–369.

Ota. K. R.. & DuPaul. G. J. (2002). Task engagement and mathematics performance in children with attention deficit hyperactivity disorder: Effects of supplemental computer instruction. School Psychology Quarterly, 17, 242-257.

Palermo, F., Hanish, L.D., Martin, C.L., Fabes, R.A., Reiser, M. (2007) Preschoolers' academic readiness: What role does the teacher–child relationship play? Early Childhood Research Quarterly, 22 (4), 407-422.

Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 3–8.

Pianta, R. C. (2003). Standardized observations from pre-K to 3rd grade: A mechanism for improving access to high-quality classroom experiences and practices during the p-3 years (Working paper). New York: Foundation for Child Development.

Pianta, R. C., La Paro, K. M., Payne, C., Cox, M. J., & Bradley, R. (2002). The relation of kindergarten classroom environment to teacher, family, and school characteristics and child outcomes. The Elementary School Journal, 102, 225–238.

Pianta, R. C, Belsky, J., Houts, R, Morrison, E, & the National Institute of Child Health and Human Development Early Child Care Research Network. (2007). Opportunities to learn in America's elementary classrooms. Science, 315, 1795-1796.

Pianta, R. C., Mashburn, A. J., Downer, J. T., Hamre, B. K., & Justice, L. (2008). Effects of web-mediated professional development resources on teacher–child interactions in pre-kindergarten classrooms. Early Childhood Research Quarterly, 23, 431–451.

Pianta, R.C. & Hamre, B.K. (2009). Conceptualization, measurement and improvement of classroom processes: Standardized observation can leverage capacity. Educational Researcher, 38, No.2, pp. 109-119.

Phye, G. D., & Sanders, C. E. (1994). Advice and feedback: Elements of practice for problem solving. Contemporary Educational Psychology, 19, 286–301.

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten classroom quality, behavioral engagement, and reading achievement. School Psychology Review, 38, 102–120.

Pressley, M., Wharton-McDonald, R., Allington, R., Block, C. C., Morrow, L., Tracey, D., et al. (2001). A study of effective first-grade literacy instruction. Scientific Studies of Reading, 5(1), 35-58.

Pridemore, D. R., & Klein, J. D. (1995). Control of practice and level of feedback in computer-assisted instruction. Contemporary Educational Psychology, 20, 444–450.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Zhai, F., Metzger, M., & Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. Journal of Consulting and Clinical Psychology, 77, 302–316.

Reed. M. L.. & Edelbrock. C. (1983). Reliability and validity of the Direct Observation Form of the Child Behavior Checklist. Journal of Abnormal Child Psychology. 521-530.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. Econometrica, 73, 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94, 247–252.

Roeser, R. W., Skinner, E., Beers, J., & Jennings, P. A. (2012). Mindfulness training and teachers' professional development: An emerging area of research and practice. Child Development Perspectives, 6, 167–173.

Shapiro. E. S.. & Heick. P (2004). School psychologist assessment practices in the evaluation of students refined for behavioral/emotional problems, Psvchology in the Schools. 41. 551 -561.

Skibbe, L. E., Phillips, B. M., Day, S., Brophy-Herb, H. E., & Connor, C. M. (2012). Children's early literacy growth in relation to classmates' self-regulation. Journal of Educational Psychology, 104(3), 451-553.

Smith, E. (1989). It doesn't bother me, but sometimes it's discouraging. 8, 253-265.

Smith, P. L, Waller, M. I., Waller, S. P. (1982). Generalizable observation of the teaching process. Educational and Psychological Measurement, 42, 467-478.

Smits, M. H. S. B., Boon, J., Sluijsmans, D. M. A., & van Gog, T. (2008). Content and timing of feedback in a web-based learning environment: Effects on learning as a function of prior knowledge. Interactive Learning Environments, 16, 183–193.

Snow, Catherine E. (2001). Reading for understanding. Santa Monica, CA: RAND Education and the Science and Technology Policy Institute.

Strong, M., Gargani, J., & Hacifazlioğlu, Ö. (2011). Do We Know a Successful Teacher When We See One? Experiments in the Identification of Effective Teachers. Journal of Teacher Education.

Supplee, L.H., Shaw, D.S., Hailstones, K. and Hartman, K. "Family and child influences on early academic and emotion regulatory behaviors". Journal of School Psychology 2004 vol. 42 221-242.

Tani, N., & Connor, C.M. (under review). The association of teachers' perceptions with second graders' behavior and academic achievement: Examining race and gender differences.

Travers, C. J. (2001). Stress in teaching: Past, present, and future. In J. Dunham (Ed.), Stress in the workplace: Past, present, and future (pp. 130–163). Philadelphia, PA.

Van de Ridder, J. M. M., Stokking, K. M., McGaghie, W. C., & Ten Cate, O. Th. J., (2008). What is feedback in clinical education? Medical Education, 42, 189–197.

Vasadey, P.F., & Sanders, E.A., (2008). Repeated reading intervention: outcomes and interactions with readers' skills and classroom instruction. Journal of Educational Psychology. Vol 100:2, pp. 272-290.

Volpe, R. J., DiPerna, J. C., Hintze, J. M., & Shapiro, E. S. (2005). Observing students in classroom settings: A review of seven coding schemes. School Psychology Review, 34, 454–474.

Whitaker, R.C., Becker, B.D., Herman, A.N., & Gooze, R.A. (2013). The physical and mental health of Head Start staff: The Pennsylvania Head Start staff wellness survey. Preventing Chronic Disease, 10, 130-171. DOI: 10.5888/pcd10.130171.

Whyte, M. M., Karolick, D. M., Neilsen, M. C., Elder, G. D., & Hawley, W. T. (1995). Cognitive styles and feedback in computer assisted instruction. Journal of Educational Computing Research, 12, 195–203.

Woodcock, R. W., McGrew, K.S. , Mather N. (2001) Woodcock-Johnson III Tests of Achievement, Riverside Publishing, Itasca, IL.

Yoshikawa, Hirokazu, & Hsueh, JoAnn. (2001). Child development and public policy: Toward a dynamic systems perspective. Child Development, 72(6), 1887-1903.

APPENDIX A

TEACHER FEEDBACK CODING SYSTEM – CODE KEY

Academic Feedback Codes

Academic Mistake Identification Codes:

- ACN: Mistake ID, Content
- AEX: Mistake ID, Enthusiastic/Exuberant
- ANU: Mistake ID, Neutral
- ASD: Mistake ID, Sad/Depressed
- AAN: Mistake ID, Angry/Frustrated

Academic Correction.

- BCN: Correction, Content
- BEX: Correction, Enthusiastic/Exuberant
- BNU: Correction, Neutral
- BSD: Correction, Sad/Depressed
- BAN: Correction, Angry/Frustrated

Academic Elaboration.

- CCN: Elaboration, Content
- CEX: Elaboration, Enthusiastic/Exuberant
- CNU: Elaboration, Neutral
- CSD: Elaboration, Sad/Depressed
- CAN: Elaboration, Angry/Frustrated

Academic Encouragement.

- DCN: Encouragement, Content
- DEX: Encouragement, Enthusiastic/Exuberant
- DNU: Encouragement, Neutral
- DSD: Encouragement, Sad/Depressed
- DAN: Encouragement, Angry/Frustrated

Academic Support Statement.

- ECN: Support Statement, Content
- EEX: Support Statement, Enthusiastic/Exuberant
- ENU: Support Statement, Neutral
- ESD: Support Statement, Sad/Depressed
- EAN: Support Statement, Angry/Frustrated

Inaccurate Academic Feedback.

- FCN: Inaccurate, Content
- FEX: Inaccurate, Enthusiastic/Exuberant
- FNU: Inaccurate, Neutral
- FSD: Inaccurate, Sad/Depressed
- FAN: Inaccurate, Angry/Frustrated

Behavior Feedback Codes

Behavioral Identification.

- 1CN: Behavior ID, Content
- 1EX: Behavior ID, Enthusiastic/Exuberant
- 1NU: Behavior ID, Neutral
- 1SD: Behavior ID, Sad/Depressed
- 1AN: Behavior ID, Angry/Frustrated

Behavioral Correction.

- 2CN: Correction, Content
- 2EX: Correction, Enthusiastic/Exuberant
- 2NU: Correction, Neutral
- 2SD: Correction, Sad/Depressed
- 2AN: Correction, Angry/Frustrated

Behavioral Elaboration.

- 3CN: Elaboration, Content
- 3EX: Elaboration, Enthusiastic/Exuberant
- 3NU: Elaboration, Neutral
- 3SD: Elaboration, Sad/Depressed
- 3AN: Elaboration, Angry/Frustrated

Behavioral Encouragement.

- 4CN: Encouragement, Content
- 4EX: Encouragement, Enthusiastic/Exuberant
- 4NU: Encouragement, Neutral
- 4SD: Encouragement, Sad/Depressed
- 4AN: Encouragement, Angry/Frustrated

Behavioral Redirection.

- 5CN: Redirection, Content
- 5EX: Redirection, Enthusiastic/Exuberant
- 5NU: Redirection, Neutral
- 5SD: Redirection, Sad/Depressed
- 5AN: Redirection, Angry/Frustrated

Behavioral Support Statement.

- 6CN: Support Statement, Content
- 6EX: Support Statement, Enthusiastic/Exuberant
- 6NU: Support Statement, Neutral
- 6SD: Support Statement, Sad/Depressed
- 6AN: Support Statement, Angry/Frustrated

Inaccurate Behavioral Feedback.

- 7CN: Redirection, Content
- 7EX: Redirection, Enthusiastic/Exuberant
- 7NU: Redirection, Neutral
- 7SD: Redirection, Sad/Depressed
- 7AN: Redirection, Angry/Frustrated

Affect Descriptions

CN: Content. The teacher is relaxed, happy, engaged with the student(s) but tones/movements are not exaggerated in any way.

EX: Enthusiastic/Exuberant. The teacher is excited, highly engaged, and providing exaggerated vocal tones, words and/or body movements.

NU: Neutral. The teacher is displaying neither positive nor negative characteristics. She is engaged with students but is doing so without any discernable emotion.

SD: Sad/Depressed. The teacher is showing signs of frustration, irritation, sadness in regard to student(s) or the material being taught. Affect may also be interpreted as flat and/or disengaged.

AN: Angry/Frustrated. The teacher is visibly and audibly upset with the student(s) and uses harsh tones, wording and/or body posturing in her attempts at feedback.

APPENDIX B

TEACHER DISPOSITION SURVEY

Teacher Disposition Survey

Please read each of the following statements carefully and select the one best answer that you feel represents you the most accurately.

1. I make my expectations clear about student behavior.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

2. Lately I have been bothered by things that don't usually bother me.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

3. I am able to respond well to defiant students.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

4. My appetite is significantly different than it used to be.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

5. I am able to assist families in helping their children do well in school.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

6. I feel that I cannot shake off the blues even with the help of my family, friends or students.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

7. I feel that I am as competent as the other teachers at my school.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

8. I have trouble keeping my mind on what I am doing.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

9. I am able to provide alternate explanations or examples when my students are confused.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

10. I feel depressed.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

11. I feel everything I do is an effort.

- Always true of myself
- Often true of myself

- Sometimes true of myself
- Rarely true of myself
- Never true of myself

12. I am able to get my students to follow classroom rules.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

13. I feel hopeful about the future.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

14. I am able to provide appropriate challenges for very capable students.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

15. I am still content with my choice of profession.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

16. I feel fearful.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

17. I am able to use a variety of assessment strategies.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

18. My sleep is restless.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

10 19. I am happy.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

20. I am able to calm students when they become disruptive or noisy.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

21. Lately, I have been talking less than usual.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

22. I feel lonely.

- Always true of myself
- Often true of myself

- Sometimes true of myself
- Rarely true of myself
- Never true of myself

23. I feel that people are unfriendly.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

24. I am able to use effective strategies to motivate students who show low interest in schoolwork.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

25. I enjoy life.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

26. I have crying spells.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

27. I feel sad.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

28. I am able to establish an effective classroom management system with each new

group of students.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

29. I feel that my students dislike me.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

30. It is difficult to “get going" at the beginning of the day.

- Always true of myself
- Often true of myself
- Sometimes true of myself
- Rarely true of myself
- Never true of myself

*Note: CES-D depressive symptom target questions in red.

APPENDIX C

QUALITY OF THE CLASSROOM LEARNING ENVIRONMENT RUBRIC

Quality of the Classroom Learning Environment (Q-CLE) Rubric

Classroom Implementation of Individualized Instruction	Classroom/Orientation Organization& Planning	Warmth and Responsiveness/Control/Discipline
Teacher Rating:	Teacher Rating:	Teacher Rating:
Fidelity Rating 1	Fidelity Rating 1	Fidelity Rating 1
The teacher is not differentiating instruction.	The classroom is not organized. Transitions are long and instructional delivery is unclear and confusing. The general feeling is of chaos.	Teacher appears as the authority figure in the class, but always punitive. Teacher does not select or incorporate students' responses, ideas, examples, and experiences into the lesson.
Fidelity Rating 1 Indicators	Fidelity Rating 1 Indicators	Fidelity Rating 1 Indicators
All of the instruction is whole class without attention to individual needs of students. Instructional delivery is not appropriately paced for children of varying skill levels.	There is no evidence of classroom organization. Teacher frequently does not have materials ready or enough materials for all children. Classroom is frequently chaotic and very little time is spent on meaningful instruction. There is no observable system in place to facilitate students' transition from one station or location to another.	There is no evidence that the teacher redirects in respectful ways, nor is there evidence that the teacher emphasizes student change in behavior through praise. There is no evidence of the teacher communicating what students did correctly or how they can improve. There is no evidence of students treating each other with respect. Whenever discipline is imposed, it is ineffective.

Classroom Implementation of Individualized	Classroom/Orientation Organization& Planning	Warmth and Responsiveness/Control/Discipline

Instruction		
Fidelity Rating 2	Fidelity Rating 2	Fidelity Rating 2
The teacher is using primarily whole class instruction. When small groups are used, they are not always focused on literacy/math. Instructional delivery is inconsistently paced for children with varying skill levels.	The classroom has inconsistent organization. Transitions are of long to reasonable duration and are inefficient. Limited instructional clarity (e.g., Teachers' instructions to children regarding how to complete activities, for example, are not always easy for students to understand).	Teacher appears as the authority figure in the class, but often punitive. Teacher rarely selects and incorporates students' responses, ideas, examples, and experiences into the lesson. Whenever discipline is imposed, it is inconsistent and only occasionally effective.
Fidelity Rating 2 Indicators	Fidelity Rating 2 Indicators	Fidelity Rating 2 Indicators
Instruction and activities are the same for all small groups.	There may be an observable, but inefficient or ineffectual system (e.g., center chart, daily schedule) in place for organizing students into groups. Teacher has inconsistent use of a lesson plan. Instructional clarity is inconsistent and children don't always understand expectations.	Teacher frequently redirects in disrespectful ways. Teacher talk is neither encouraging nor respectful. Rarely connects students' personal experiences to lesson content. Use of directive rather than open-ended behavior management (e.g., sit down rather than everybody please go to their reading corner).

Classroom Implementation of Individualized Instruction	Classroom/Orientation Organization& Planning	Warmth and Responsiveness/Control/Discipline

Fidelity Rating 3	Fidelity Rating 3	Fidelity Rating 3
Clear evidence of differentiation. The teacher is using small groups; however, the children in the small groups are generally receiving highly similar amounts and types of instruction.	The classroom is reasonably organized, instructional clarity is evident and transitions are fairly efficient.	Teacher appears as the authority figure in the class, and is occasionally punitive. Teacher is minimally effective at selecting and incorporating students' responses, ideas, examples, and experiences into the lesson. There is a behavior management system in place, but it is used inconsistently.
Fidelity Rating 3 Indicators	Fidelity Rating 3 Indicators	Fidelity Rating 3 Indicators
At least one of the following is evident: Teacher managed small groups, seatwork, or child managed literacy/math activities. The teacher may work with a group of students while other students are participating in child-managed activities. However, generally, child-managed activities are not differentiated.	Transitions are of reasonable length, but not consistently efficient (not all children). There is an observable, but not always efficient or working system (e.g., center chart, daily schedule) in place for organizing students into groups. The teacher may use a daily lesson plan (e.g., group activity planner print-out or similar written plan).	Teacher inconsistently redirects in respectful ways and inconsistently emphasizes student change in behavior through praise. Teacher talk is inconsistently encouraging and respectful and inconsistently connects students' personal experiences to lesson content. Inconsistently communicates clearly what students did correctly or how they can improve. Students inconsistently treat each other with respect.

Classroom Implementation of Individualized Instruction	Classroom/Orientation Organization& Planning	Warmth and Responsiveness/Control/Discipline

Fidelity Rating 4	Fidelity Rating 4	Fidelity Rating 4
Clear evidence of differentiation. The teacher is using small groups and there is evidence that instruction is individualized. Most of the language arts/math block is spent in meaningful literacy/math activities.	The classroom is fairly well organized and there is adequate but not excellent instructional clarity. Instruction is usually planned in advance.	Teacher appears as the authority figure in the class, and is seldom punitive. Partially effective at selecting and incorporating students' responses, ideas, examples, and experiences into the lesson. There is a behavior management system in place.
Fidelity Rating 4 Indicators	Fidelity Rating 4 Indicators	Fidelity Rating 4 Indicators
The teacher may work with a group of students while other students are participating in child managed activities. At least two of the following are evident: Teacher managed small groups, seatwork, or child managed literacy/math activities.	There is an observable, efficient, and working system (e.g., center chart, daily schedule) in place for organizing students into groups. There is an observable system in place supporting students' transition from one station or location to another, and there is evidence that the system is working at least part of the time. The teacher consistently follows a daily lesson plan (e.g., group activity planner print-out or similar plan).	Teacher fairly frequently redirects in respectful ways and emphasizes student change in behavior through praise. Teacher talk is fairly encouraging and respectful. Fairly frequently connects students' personal experiences to lesson content. Usually communicates clearly what students did correctly or how they can improve. Students fairly consistently treat each other with respect. There is a behavior management system in place that is used fairly consistently and effectively.

	most of the time.	
Classroom Implementation of Individualized Instruction	Classroom/Orientation Organization	Warmth and Responsiveness/Control/Discipline
Fidelity Rating 6	Fidelity Rating 6	Fidelity Rating 6
Teachers who fully implement multiple and flexible student grouping configurations and regrouping of students based on formal or informal assessment data. The content of literacy/math instruction is differentiated. The entire language arts/math block is spent in meaningful literacy/math activities.	The classroom is well organized and instruction is well organized. Classroom routine is evident. Transitions are efficient	Teacher appears the authority figure in the class, never punitive. Effectively selects and incorporates students' responses, ideas, examples, and experiences into the lesson. Classroom consistently offers a positive learning environment with clear expectations for students' behavior as a member of the learning community.
Fidelity Rating 6 Indicators	Fidelity Rating 6 Indicators	Fidelity Rating 6 Indicators
Teacher managed small groups, seatwork, and child managed literacy/math activities are all evident and differentiated. Students work independently at stations or are engaged in activities designed for their skill level and that reinforce developing skills	There is an observable, efficient, and working system (e.g., center chart, daily schedule) in place for organizing students into groups and the teacher consistently follows a daily lesson plan (e.g., group activity planner print-out or similar written	Teacher consistently redirects in respectful ways. Consistently uses friendly or unobtrusive, respectful gestures to redirect behavior. Consistently emphasizes student change in behavior through praise. Teacher talk is consistently encouraging and respectful. Consistently communicates clearly

while the teacher or other adult works with small homogeneous groups of students.	plan). Activities run smoothly with few disruptions. Materials are readily available to students. Disruptions are handled quickly and effectively. Students know what to do and what is expected of them. There is an observable, efficient, and working system in place for students to transition from one station or location to another and there is evidence that the system is working well.	what students did correctly or how they can improve. Consistently connects students' personal experiences to lesson content. Consistently elicits responses from all students, including students having difficulty with the task at hand. Consistently calls on a range of students. Consistently uses specific praise for behavior. Students consistently treat each other with respect. Teacher frequently encourages peer support in the learning process. The behavior management system is used consistently and effectively.

Note: For ratings of 1, the teacher is consistently weak in this area. For ratings of 3, the teacher shows the characteristic but is inconsistent; for ratings of 5, the teacher is consistently strong in this area, for ratings of 6, he or she is exemplary.

APPENDIX D

IRB APPROVAL

Original IRB approval for longitudinal parent study

To:	Carol Connor
From:	Mark Roosa, Chair Soc Beh IRB
Date:	08/17/2012
Committee Action:	**Expedited Approval**
Approval Date:	08/17/2012
Review Type:	Expedited F7
IRB Protocol #:	1208008102
Study Title:	Observation of Effective Classroom Teaching
Expiration Date:	08/16/2013

The above-referenced protocol was approved following expedited review by the Institutional Review Board.

It is the Principal Investigator's responsibility to obtain review and continued approval before the expiration date. You may not continue any research activity beyond the expiration date without approval by the Institutional Review Board.

Adverse Reactions: If any untoward incidents or severe reactions should develop as a result of this study, you are required to notify the Soc Beh IRB immediately. If necessary a member of the IRB will be assigned to look into the matter. If the problem is serious, approval may be withdrawn pending IRB review.

Amendments: If you wish to change any aspect of this study, such as the procedures, the consent forms, or the investigators, please communicate your requested changes to the Soc Beh IRB. The new procedure is not to be initiated until the IRB approval has been given.

Please retain a copy of this letter with your approved protocol.

Continued IRB Approval for longitudinal parent study

ASU Knowledge Enterprise Development

APPROVAL:CONTINUATION

Carol Connor
Learning Sciences Institute (LSI)
-
Carol.Connor@asu.edu

Dear Carol Connor:

On 3/19/2015 the ASU IRB reviewed the following protocol:

Type of Review:	Continuing Review
Title:	Response to Intervention: Following students who were in the Individualizing Student Instruction (ISI) Observation of Effective Classroom Teaching Studies
Investigator:	Carol Connor
IRB ID:	1304009064
Category of review:	(7)(b) Social science methods, (7)(a) Behavioral research
Funding:	Name: NICHD: National Institute of Child Health and Human Development , Funding Source ID: HHS-NIH-NICHD-National Institute of Child Health & Human Dev
Grant Title:	None
Grant ID:	None
Documents Reviewed:	

The IRB approved the protocol from 3/19/2015 to 4/14/2016 inclusive. Three weeks before 4/14/2016 you are to submit a completed Continuing Review application and required attachments to request continuing approval or closure.

If continuing review approval is not granted before the expiration date of 4/14/2016 approval of this protocol expires on that date. When consent is appropriate, you must use final, watermarked versions available under the "Documents" tab in ERA-IRB.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

IRB section showing that Leigh McLean is an approved member of the research team for secondary analysis of data collected in the longitudinal parent study.

View: SF: Study Team Members

Study Team Members

4/30/15, 9:40 AM

1. **Identify each additional person involved in the design, conduct, or reporting of the research:**

Name	Roles	Financial Interest	Involved in Consent	E-mail	Phone
Stephanie Day	Co-Investigator			Stephanie.Lynn.Day@asu.edu	-
Sarah Ingebrand	Graduate Student	no	no	Sarah.Ingebrand@asu.edu	-
Leigh McLean	Graduate Student	no	no	Leigh.McLean@asu.edu	-
Nicole Sparapani	Post Doctoral Scholar	no	no	Nicole.Sparapani@asu.edu	-

BIOGRAPHICAL SKETCH

Leigh Ellen McLean was born in Portland, Oregon on May 27th, 1987. She attended Oregon State University from 2006 to 2010, earning her bachelor's degree in Human Development and Family Sciences, with a specialty in Early Childhood Development and Education. Upon completion of this degree, she began a graduate program in Developmental Psychology at Florida State University, where she also was a member of the IES Predoctoral Interdisciplinary Research Training (PIRT) fellowship program. She earned her master's degree from FSU and then transitioned to Arizona State University to pursue her doctorate in Developmental Psychology. Leigh's research interests include the contributions of teacher and student characteristics to overall classroom quality in early elementary settings. More specifically, she studies teachers' depressive symptoms and other affective characteristics and their impact on teacher performance and student development. Upon completion of her doctoral degree, Leigh will pursue a career as an independent researcher in the field of developmental psychology.