



Classroom quality as a predictor of first graders' time in non-instructional activities and literacy achievement



Leigh McLean ^{a,*}, Nicole Sparapani ^b, Jessica R. Toste ^c, Carol McDonald Connor ^d

^a Arizona State University, T. Denny Sanford School of Social and Family Dynamics, United States

^b Arizona State University, Institute for the Science of Teaching and Learning, United States

^c The University of Texas at Austin, College of Education, United States

^d University of California at Irvine, School of Education, United States

ARTICLE INFO

Article history:

Received 20 January 2015

Received in revised form 23 March 2016

Accepted 25 March 2016

Available online xxxx

Action Editor: Scott Graves

Keywords:

Classroom quality

Instruction

Literacy

Early elementary

First grade

ABSTRACT

This study investigated how quality of the classroom learning environment influenced first grade students' ($n = 533$) time spent in two non-instructional classroom activities (off-task and in transition) and their subsequent literacy outcomes. Hierarchical linear modeling revealed that higher classroom quality was related to higher student performance in reading comprehension and expressive vocabulary. Further, classroom quality predicted the amount of time students spent off-task and in transitions in the classroom, with slopes of change across the year particularly impacted. Mediation effects were detected in the case of expressive vocabulary such that the influence of classroom quality on students' achievement operated through students' time spent in these non-instructional activities. Results highlight the importance of overall classroom quality to how students navigate the classroom environment during learning opportunities, with subsequent literacy achievement impacted. Implications for policy and educational practices are discussed.

© 2016 Society for the Study of School Psychology. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Students' early literacy performance is one of the strongest indicators of success throughout formal schooling, with patterns of achievement established in early elementary school often lasting into adulthood (National Institute of Child Health and Human Development [NICHD], 2000). Students who are able to read proficiently by third grade are less likely to exhibit attention and behavior problems (Adams & Snowling, 2001; Maughan & Carroll, 2006) and are far more likely to experience overall long-term success (Reynolds & Ou, 2004; Spira, Bracken, & Fischel, 2005). In addition, multiple studies have documented relations between poor achievement patterns and negative life outcomes, such as higher rates of school dropout and incarceration (Duncan et al., 2007; NICHD, 2000). Despite these findings, recent large-scale evaluations of reading achievement in elementary students present a discouraging picture, with only 38% of fourth graders reading at or above a proficient reading level with rates even lower for students from low income families (National Association of Educational Progress [NAEP], 2011) even in the decade following major federal initiatives such as Reading First and Race to the Top. This staggering deficit in reading achievement illustrates the need for research investigating the reasons behind reading failure in early elementary students. The present study seeks to contribute to these efforts by clarifying exactly how the classroom

Funding for this project was provided by the U.S. Department of Education Institute of Education Sciences (R305H04013 and R305B070074) and the National Institute of Child Health and Human Development (R01HD48539, R21HD062834, and P50 HD052120)

* Corresponding author at: Arizona State University, PO Box 873701, 850 S. Cady Mall, Tempe, AZ. 85287-3701.

environment, a highly influential developmental context, impacts the way students spend their time during learning opportunities and what the implications of these factors are for early literacy skill growth.

1.1. Theoretical framework

First grade is considered a critical period for literacy development because it is the first time students are exposed to purposeful, formal literacy instruction (Spira et al., 2005). During this time, teachers have the capacity to set the course of student development through their contributions to the learning environment. The bio-ecological model of child development (Bronfenbrenner & Morris, 2006) and dynamic systems theory (Yoshikawa & Hsueh, 2001) are two views that illustrate how the classroom context can influence student development in these early years. The bio-ecological model highlights the significant systems in which child development takes place, paying close attention to the contexts in which students spend their time (e.g., microsystems). Following this, we view the classroom as a significant microsystem of influence that has strong potential to impact early development across multiple domains. Further, dynamic systems theories posit that within this microsystem, there are multiple elements of the classroom that interact synergistically to determine how students experience (and what they take away from) their time. That is, characteristics within a classroom have the potential to influence each other in a chain reaction of events that work together to impact student outcomes. Utilizing this framework, we hypothesize that classroom systems indicative of classroom quality, namely the instruction, management and organizational systems in place to promote efficient learning, may impact student outcomes indirectly by influencing how students spend their time in the classroom. Specifically, if a classroom is poorly organized with ineffective management and instruction, students may have difficulty navigating their environment, attending to learning opportunities, or staying on-task, which may negatively impact learning.

We recognize that multiple student characteristics bring unique contributions to the microsystem of the classroom (Bronfenbrenner & Morris, 2006), undoubtedly influencing classroom quality independently of the teacher; however, teacher-focused elements of classroom quality have been conceptualized as the “starting point” in the relations hypothesized within this study. This is because we posit that it is the teachers' responsibility to make purposeful adjustments to classroom elements that may be hindering student development. As such, we approach this study from the teachers' points of view, focusing on their contributions to the classroom environment and the impact they may have on student actions and subsequent literacy achievement (Guo, Connor, Tompkins, & Morrison, 2011; Merritt, Wanless, Rimm-Kaufman, & Peugh, 2012).

1.2. Defining and measuring classroom quality

The quality of the classroom-learning environment has been found to be a primary driving force behind student development (Bronfenbrenner & Morris, 2006; Cameron, Connor, & Morrison, 2005; Connor et al., 2009a, 2009b; Guthrie, Anderson, Aloa, & Rinehart, 1999; Rimm-Kaufman, La Paro, Downer, & Pianta, 2005). High quality classrooms support social-emotional development and academic competence (Connor, Son, Hindman, & Morrison, 2005; Domitrovich, Cortes, & Greenberg, 2007; Hamre & Pianta, 2007; Ponitz, McClelland, Matthews, & Morrison, 2009a) as well as buffer against academic and social delays associated with student risk characteristics, such as difficult temperament, low self-regulation, and low socioeconomic status (SES; Curby, Rudasill, Edwards, & Perez-Edgar, 2011; Hamre & Pianta, 2005; Rimm-Kaufman et al., 2005; Reyes, Brackett, Rivers, White, & Salovey, 2012). Many of the individual components of classroom quality, including organization, management, and instruction have been found to be independently predictive of student outcomes (Bohn, Roehrig, & Pressley, 2004). Additionally, there appears to be a large degree of relatedness between the individual classroom quality components. Pressley et al. (2001) investigated the impacts of multiple classroom features concurrently on first graders' 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 predictive of student achievement. Additionally, Mashburn et al. (2008) found that teachers' instructional and emotional interactions with students were predictive of both academic and social emotional outcomes. Such findings suggest that these classroom elements are well-conceptualized together as one comprehensive construct that has the capacity to impact student outcomes (Connor et al., 2011; Hamre & Pianta, 2007).

In the present study, we measure classroom quality using the Quality of the Classroom Learning Environment (Q-CLE; Connor et al., 2011, 2014) rubric, a novel observational measure developed by the authors and used multiple times within the longitudinal parent study described below (Connor et al., 2009a; b; Connor et al., 2014). Building directly on our theoretical framework, which views the classroom as a dynamic system of developmental influence, the Q-CLE rubric separately assesses the observable physical characteristics of the classroom environment, the purposeful instruction provided by the teacher, and the interactions that take place between teachers and their students—then considers these factors concurrently as a whole construct. Using this approach, we attempt to provide a comprehensive, multi-component view of the classroom as a dynamic microsystem that involves many elements interacting with each other in various ways to impact student experiences.

High-quality classrooms are defined herein as learning environments that involve purposeful and teacher-implemented systems of organization, instruction, and management meant to aid in successful student learning. Organization includes both the physical characteristics of a classroom as well as the techniques employed by the teacher 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). Management refers to the teachers' use of intentional proactive and reactive actions to maintain an environment in which students can successfully learn. This includes both a teachers' level of warmth and responsiveness to students (analogous to ‘emotional support’ in other comparable measures, i.e., Pianta, La Paro, & Hamre, 2007) as well as classroom control and

discipline techniques. Instruction is conceptualized as the purposeful communication of academic information from the teacher to the students as well as the extent to which the teacher successfully individualizes instruction based on the unique needs of each student. The Q-CLE rubric attempts to capture all of these components simultaneously in an effort to reflect the hypothesized dynamic nature of the early elementary classroom.

There are a number of well-validated measures available in the field today that assess aspects of the classroom environment, each of which bring a unique perspective to how the classroom might influence student development. These include, among others, the Classroom Assessment Scoring System (CLASS; Pianta et al., 2007) and the NICHD-SECYD Classroom Observation System (NICHD-ECCRN, 2002) (Kane & Staiger, 2012). The CLASS views teacher-student interactions as the driving force behind student learning (and thus might best be conceptualized as a measure of teacher-student interaction quality) and assesses these interactions across three domains: emotional climate, management, and instructional support (Mashburn et al., 2008; Pianta et al., 2007). Similarly, the COS includes global ratings of control, emotional climate, classroom management, instruction, feedback, and child responsibility. The Q-CLE's quality indicators of 'instruction' and 'classroom management' mirror these approaches, considering the quality of the instruction offered by teachers as well as their responsiveness to students and their techniques for managing student behavior. In addition to this, the Q-CLE also considers the physical structure and function of the classroom and how these factors contribute to efficiency during learning opportunities. In this respect, this measure may best be conceptualized as a more global assessment of classroom quality.

The Q-CLE is unique in that it considers teachers' warmth/responsiveness and control/discipline as one dimension rather than two separate classroom characteristics. The grouping of these variables within the Q-CLE was heavily informed by parenting research, specifically Baumrind's (1971; 1978; 1991) work identifying parenting styles based on levels of parent responsiveness and demandingness. Within this framework, 'authoritative' parents display high amounts of responsiveness as well as high amounts of demandingness in their interactions with their children, and this combination of characteristics leads to the most optimal child outcomes. We took this framework and applied it to the classroom in our development of the Q-CLE, positing that the implications of these parenting styles on child development would be mirrored in the relationships that students share with their teachers. Thus, an authoritative teacher would then theoretically display high amounts of responsiveness ('warmth/responsiveness') and demandingness ('control/discipline'). We hypothesize that a teachers' level of success in maintaining high levels of both characteristics simultaneously is what contributes most to the overall construct of classroom quality.

Past work has strongly supported the dynamic interplay between classroom structure/function, instructional quality, and teachers' responsiveness and control techniques (Cameron et al., 2005; Eccles & Gootman, 2002); however there is still variability in the measurement of classroom quality across the field, with measures often ignoring one or more of these domains. Given that recent findings suggest even the most experienced observers of teachers (i.e., principals, other teachers) are relatively unable to distinguish effective from ineffective teachers based on in-person observations (Strong, Gargani, & Hacifazlıoğlu, 2011), the need for high-quality, valid observational tools for measuring classroom quality becomes clear. As such, we hope to provide further evidence within the present study that the Q-CLE rubric can successfully be used within educational research practice as a valid tool for classroom observation.

1.3. Non-instructional behavior in the classroom

Young students' time spent engaged in learning opportunities, more specifically the related behaviors of persisting at individual work, staying on-task for extended periods, and attending to learning goals, are strongly associated with academic success (Cameron et al., 2005; Cameron, Connor, Morrison, & Jewkes, 2008; DiPerna, Lei, & Reid, 2007; Fredricks, Blumenfeld, & Paris, 2004; Ladd, Birch, & Buhs, 1999). However, despite the clear importance of engagement in learning tasks, early foundational studies have found that students spend up to half of a day's allotted instructional time in activities that are not related to learning (Anderson, 1981; Fredrick, Walberg, & Rasher, 1979); in this study, we qualify these instances as "non-instructional" activities. A primary goal of teachers is to maximize student time spent engaged in meaningful learning opportunities by maintaining student attention and facilitating smooth, efficient transitions between activities (Gettinger & Seibert, 2002), but challenges to maintain success in these areas are likely to result in lost instructional time for students, and thus dampened academic growth. The present study posits that one reason students may spend less time engaged in learning opportunities is because they are spending more of their time in the classroom either off-task or in less inefficient transitions between activities. Classroom quality may contribute to these behaviors, with research suggesting that students in higher quality classrooms spend significantly less time off-task (Rimm-Kaufman, Curby, Grimm, Nathanson, & Brock, 2009).

This study is unique in its attempt to capture the contribution of classroom quality to the relation between students' engagement in learning opportunities (specifically, their time spent in non-instructional activities) and their academic outcomes. On the part of the teacher, successfully re-engaging students who become off-task and ensuring efficient transitions are important components of successful classroom management (Finn, Pannozzo, & Voelkl, 1995), and while these seem like fairly straightforward maneuvers, many steps spanning all three domains of classroom quality considered in this study contribute to these processes. A teacher must constantly monitor all students in the classroom, be able to identify off-task behavior, and successfully redirect students in a way that is both effective and that promotes a positive teacher-student relationship. All of this must happen in a context that has purposefully structured organization systems to support efficient movement throughout the classroom and minimize distractions. Given that these processes must happen repeatedly across a school day, the importance (and challenge) of teachers' ability to maintain success in all of these areas simultaneously becomes clear. Foundational to the present study, Connor et al.

(2010a,b) found that teachers' level of success in implementing effective classroom instruction, organization, and management (as measured by the Q-CLE rubric) was related to students' literacy gains in early elementary school.

1.4. Present study

The purpose of the present study was to examine the relations among classroom quality, student time spent in the two observed non-instructional activities (i.e., off-task and transitions), and their literacy outcomes. We sought to address the following research questions: Firstly, is there a relation between observed classroom quality and student literacy growth in first grade? We predict that students in higher quality classrooms will have higher literacy gains across the school year. Secondly, is there a relation between classroom quality and students' time spent off-task and in transitions? We hypothesize that students in higher quality classrooms will spend less time off-task and transitioning between activities overall, and that the amount of time they spend in these activities will decrease from fall to spring. Alternatively, we predict that students in low quality classrooms will spend more time in non-instructional activities overall, and will show an increase in the amount of time they spend in these activities by the end of the year. Lastly, does student participation in these non-instructional activities impact literacy skill growth, and does time in non-instructional activities mediate the association between classroom quality and student literacy achievement? We predict a negative association between student participation in non-instructional activities and their literacy outcomes as well as hypothesize that a mediation effect will exist, such that the relation between classroom quality and student literacy achievement will be explained by students' participation in non-instructional activities, keeping in mind that time spent in non-instruction largely reflects less time in meaningful learning opportunities.

2. Methods

2.1. Participants

This study utilized data that were collected as part of a longitudinal study which began in 2005/2006 (Connor et al., 2013) and examined the feasibility and effects of individualizing literacy instruction based on students' learning needs throughout elementary school. Participants of the current study were recruited in the 2006–2007 school year as first graders and include 533 students in 49 classrooms across 18 schools in a North Florida school district. Of these students, 46% were African American, 45% were Caucasian, and the remaining 9% were other ethnicities including Hispanic and Asian. In addition, 51% of students were female and 49% were male.

Schools represented a wide range of community SES, determined by school-wide percentage of student enrollment in a Free and Reduced Lunch (FARL) program, a widely used indicator of community poverty level. FARL enrollment ranged from 4% (affluent) to 96% (high-poverty) across schools. All teachers met state-mandated credential requirements and had attained at least a bachelor's degree in education. On average, teachers had about 12 years ($M = 11.73$ years) of teaching experience. Additionally, 46% had a master's degree or higher and were mostly female (two male teachers). The majority were Caucasian (73%), with 24% African American or other ethnicities (3%) based on self-report.

2.2. Measures

2.2.1. Literacy skills

The Letter-Word Identification (LWID), Passage Comprehension (PC), and Picture Vocabulary (PV) tests of the Woodcock-Johnson III Tests of Achievement (WJ III; Woodcock, McGrew, & Mather, 2001) were used to assess students' word reading, reading comprehension, and vocabulary skills respectively. Scores from these tests have demonstrated internal consistency reliability estimates of .98 for LWID, .96 for PC and .71 for PV for students in first grade (ages 6–7 years). LWID asks students to recognize and identify letters and then read increasingly more difficult and complex words. PC asks students to fill in the word missing from increasingly difficult passages, students must be able to decode the passages while comprehending what they read well enough to select the appropriate word (Keenan, Betjemann, & Olson, 2008). PV is a measure of expressive vocabulary that asks students to name pictures of increasingly more complex objects. Students received either Form A or Form B, at random, in the fall and the opposite form in the spring; there were no form effects when means were compared.

Scores of '0' (incorrect) or '1' (correct) are assigned to each item of a WJ test based on student answers to item prompts. The raw score is used to calculate each student's age and grade-equivalent score, standard score, percentile rank within the sample, and W score (which was used in the analyses in the present study and described in detail in the analytic strategy section). Raw scores on the WJ range from 0 to 76 for LWID, 0–44 for PV, and 0–47 for PC; with '0' indicating skills below preschool-level and higher scores indicating more advanced skills. Fall raw scores for the present sample ranged from 4 to 64 for LWID, 2–28 for PV, and 1–30 for PC. Spring scores on each measure within this sample ranged from 10 to 70 for LWID, 6–30 for PV, and 5–33 for PC.

2.2.2. Classroom video observations

Three video recordings of classroom instruction were captured for each of the 72 classrooms across the 2006–2007 academic year, one each in the fall, winter, and spring. During classroom observations, two video cameras were used to record all classroom activities taking place, with one camera focused on the majority of students from a wider perspective and the other focusing

primarily on the teacher and the students he/she was directly interacting with. Videographers wrote physical descriptions of all students present and took detailed notes of classroom activities. The designated literacy instruction blocks of these observations were coded using the Individualizing Student Instruction coding system (ISI; Connor et al., 2009a,b). Eight to twelve students per classroom were randomly selected for coding from strata based on their fall literacy achievement. A team of trained coders achieved adequate inter-rater reliability of .76 (Cohen's kappa) on a randomly selected 12% of video observations.

Coding within the ISI system consists of documenting the frequency and duration of all classroom activities experienced by each target student across three domains: content area, context/management, and instructional activity. *Content area* refers to the overall subject being taught during the video observation, for example, Language Arts or Science. *Context/Management* refers to the grouping of students within the classroom in relation to the teacher and includes teacher-managed or child/peer-managed. For example, if a student is working in a small group that is led by the teacher, that student's context/management would be coded as small-group, teacher-managed. *Instruction* refers to the specific learning activity that the target student is engaged in. Types of instruction identified within the ISI system range from phonological awareness to reading comprehension to writing. Within each instruction type, there are multiple activities that a student can be engaged in. For example, a student would be coded as participating in reading comprehension instruction, with an "add-on" code specifying that the specific type of reading comprehension instruction taking place is a compare/contrast activity. Coding takes place at the level of the student, and any activity lasting 15 s or longer is coded.

The focus in the present study is on the portion of coded data that represents a separate category, time spent in *non-instructional* activities. Whereas purposeful instruction (e.g., reading comprehension as outlined above) involves direct and observable participation by the student in a learning activity, non-instructional activities are defined as any activity in which the target student is not engaging in any type of purposeful academic instruction. Within the ISI system, non-instructional activities include students' time spent off-task, transitioning between instructional activities, being exposed to planning/organizing, waiting for the teacher to prepare for/start an activity, waiting for the teacher while she/he is handling a disruption, and going to the bathroom. This study investigates the total amount of time, summed in minutes, students spent engaging in off-task behavior and in transitions between activities.

2.2.3. Quality of the classroom-learning environment

Classroom quality was assessed using the Quality of the Classroom Learning Environment (Q-CLE) rubric (Connor et al., 2011, 2014), which is an observational measure that captures the contribution of teacher practices to the quality of the classroom learning environment by concurrently considering the interactions between teachers and their students, the developmental appropriateness of teachers' instruction, and the structural features of the classroom environment. In addition to the CLASS (discussed above), development of this measure was heavily informed by the NICHD-SECYD Classroom Observation System (NICHD-ECCRN, 2002). In addition to time-sampled coding, the COS includes global ratings of classroom quality across multiple dimensions such as control, emotional climate, classroom management, instruction, feedback, and child responsibility. The Q-CLE has predicted student literacy achievement in past studies (Connor et al., 2014; Connor et al., 2013; Connor et al., 2009a,b), consistently indicating that higher classroom quality, as conceptualized by the Q-CLE, is associated with higher student achievement outcomes.

Raters, all of whom were experienced ISI coders, were trained on the Q-CLE rubric and achieved an inter-rater reliability kappa coefficient of .89 on a random selection of 15% of the video recordings. Important to note, raters for the present study had experience with the ISI system but did not code the same videos using ISI prior to assessing them for classroom quality. Raters provided one rating on a scale of 1 to 6 for each of the Q-CLE rubric's three quality indicators. A score of 6 indicated absolute success in an area and a score of 1 indicated poor teacher contributions. Following are the evaluation parameters for each classroom quality indicator:

Classroom Implementation of Individualized Instruction: refers to the overall quality of instruction provided by the teacher as well as the degree to which the teacher is taking into account the individual needs of the students in her instructional applications.

Classroom Orientation and Organization: refers to the extent to which observable systems are in place to promote efficiency of learning in the classroom. These systems can be either physical structures or can be instructional tactics employed by the teacher.

Classroom Management (Warmth and Responsiveness, Control, Discipline): refers to the teachers' ability to secure and maintain student attention, respond appropriately to students' needs, and address/redirect student behavior when needed.

2.3. Procedures

Institutional Review Board (IRB) approval was gained for the parent longitudinal study, and all parents, teachers, and school administrators gave informed consent for participation in that and subsequent studies using these data. Students were given all WJ-III literacy subtests three times throughout the year, once each in the fall (September/October), the winter (January/February), and the spring (April/May). Students were pulled out of class individually to complete the tests by a trained administrator. Video recordings of classroom instruction also followed this schedule, with three full days of instruction captured per classroom, once each in the fall, winter, and spring. The literacy portions of all videos, three per classroom spanning three seasons, were coded

for literacy instruction using the ISI coding system, and the non-instruction variables of interest in the present study were calculated from these coded data. The winter video observation taken for each classroom was assessed for classroom quality using the Q-CLE. This was done in an attempt to capture the most stable time-point of the year for classrooms (Hamre & Pianta, 2007).

2.4. Analytic strategy

2.4.1. Variables

W scores, a Rasch-based score which provide a common scale of comparison that represents task difficulty and student growth, were computed in the WJ-III Compuscore software program and used to represent student achievement on the WJ-III tests. These scores are centered at 500 ($SD = 15$), which is interpreted as the approximate average performance of a 10-year-old child. Additionally, principal components analysis (PCA) was used to investigate the factor structure of the Q-CLE rubric's three quality indicators and to create a factor score representing a composite of these indicators. PCA provided additional verification of our assumption that the Q-CLE dimensions represent one construct for this sample of children based on past investigations of this measure (Connor et al., 2011, 2014; McLean & Connor, 2015). While studies using this rubric have revealed consistent relations among the three dimensions across multiple grades and have all suggested a single common factor, this rubric is still relatively new within the field and thus exploratory, rather than confirmatory, factor analyses were deemed most appropriate. This analysis revealed that scores on the three quality indicators loaded strongly onto one factor which accounted for 82% of the variance of classroom quality within the data, with loadings ranging from .85 to .94. Direct Oblimin rotation was initially used as each of the three dimensions of Q-CLE were highly correlated; however as only one common factor was detected rotation was unnecessary. Lastly, total time spent by each student in each of the target non-instructional activities was calculated using data from the ISI coding system that was applied to video observations. The total amount of time each individual student was observed to be off-task or in transition was summed and converted to minutes for each activity, per child, per season.

2.5. Preliminary analyses

Descriptive statistics (see Table 1) revealed that students' performance on all WJ-III literacy tasks fell within the expected range for first grade. Additionally, students made expected gains across the year on these measures. These analyses also revealed distributions within normal ranges of skewness and kurtosis on all student-level variables. Bivariate correlations run at the student level (i.e., the same teacher-level variable scores were assigned to each student within each classroom) among all primary variables revealed small, positive associations between Q-CLE and students' fall and spring Letter-Word Identification (LWID), Picture Vocabulary (PV), and spring Passage Comprehension (PC) scores (see Table 2). Correlations also revealed a positive relation between Q-CLE and students' time in transitions in the fall, with the nature of this relation changing in the spring. Furthermore, increased time spent in transitions in the fall was positively related to students' fall and spring PV scores. Finally, students' time off-task in the fall was negatively related to their spring PC scores.

2.5.1. Missing data

Small amounts of missing data on literacy measures and coded video recorded data were present and appeared to be due to unforeseen student absences on classroom video recording and assessment days or to family relocations during the study. The presence of missing data for each variable was correlated with all variables used in analyses in order to assess the likelihood that data were missing at random. Missing data on the fall WJ tests were not significantly related to students' time off-task, in transition or to Q-CLE. Small relations existed between missing data on spring WJ tasks and students' time off-task ranging in size from $r = .1$ to $r = .13$ ($p < .05$). No significant relations were detected between missing data on the spring WJ tests and Q-CLE, nor between missing data for student time off-task or in transition and Q-CLE. While it is likely that data were missing at random, all analyses used the robust maximum likelihood (ML) estimator to account for those small relations that were detected. Due to a higher proportion of missing data on the fall measurement of PC, statistical models involving this measure controlled for fall LWID as these two variables were highly correlated in the fall season ($r = .84$, $p < .001$).

Table 1

Descriptive statistics for fall & spring student-level variables.

	Fall					Spring				
	N	Min	Max	M	SD	N	Min	Max	M	SD
Letter-Word ID	525	314	536	411.86	31.71	520	345	556	459.97	26.82
Picture Vocab	525	398	513	479.69	10.78	522	430	521	485.35	11.00
Passage Comp	380	358	500	443.54	24.84	520	405	509	469.45	15.87
Transition (Mins)	512	0	106.87	16.74	14.86	477	0	83.70	12.64	10.60
Off Task (Mins)	512	0	66.83	4.70	8.35	477	0	86.83	4.93	8.23

Note: WJ-3 tests are reported as W-scores, non-instruction variables are reported in minutes.

Table 2

Correlations among Q-CLE, fall & spring non-instruction and fall & spring literacy achievement.

	1	2	3	4	5	6	7	8	9	10	11
1. Q-CLE	1										
2. Fall Trans.	.26**	1									
3. Sp. Trans.	-.11**	.23**	1								
4. Fall Off Task	-.09	.16**	.06	1							
5. Sp. Off Task	-.08	-.01	.16**	.17**	1						
6. Fall LWID	.08*	.07	-.01	-.08	-.09	1					
7. Fall PV	.13**	.13**	.04	-.05	-.03	.40**	1				
8. Fall PC	.07	.03	-.01	-.08	.01	.83**	.39**	1			
9. Sp. LWID	.11**	.04	-.05	-.09	-.03	.78**	.37**	.73**	1		
10. Sp. PV	.19**	.18**	.05	-.06	-.04	.42**	.77	.45**	.42**	1	
11. Sp. PC	.16**	.06	-.01	-.14**	-.03	.70**	.44**	.70**	.83**	.51**	1

Note: Fall season (Fall); Spring season (Sp.); Active Transition (Act. Trans.); Letter-Word Identification (LWID); Picture Vocabulary (PV); Passage Comprehension (PC); LWID, PV, and PC are subtests on the WJ-III.

* Correlation is significant at $p < .05$.

** Correlation is significant at $p < .001$.

2.5.2. Hierarchical linear modeling

In the present study, students from the same classroom may have had scores more similar to one another due to the between-teacher variance expected on ratings' of classroom observations and quality (Q-CLE). Failing to take this dependency into account can result in biased statistical results (Raudenbush & Bryk, 2002). As such, Hierarchical Linear Modeling (HLM 7; Raudenbush & Bryk, 2002) was performed, which accounted for the nested structure of this data and reduced such potential bias. HLM was used for all statistical analyses for a data structure where students (level 1) were nested within teachers (level 2).

Unconditional models, or models with no predictor variables, were first run on all student-level variables of interest to investigate grand-mean intercepts and fall-to-spring slopes of change of non-instruction variables. Additionally, these models allowed the calculation of intra-class correlation coefficients (ICC), which represent the proportion of variance in the data at each level of nesting for all student-level variables. Classroom-level ICC estimates for student performance on the WJ tasks across fall and spring were also small, ranging from .11 to .16. Classroom-level ICC estimates were large for both non-instruction variables at .36 for off-task and .48 for transitions. These high coefficients suggested that much of the variability in student behavior was attributable to differences between classrooms, justifying the use of multilevel modeling for further analysis (Dyer, Hanges, & Hall, 2005). ICCs for non-instruction variables at the school level were comparatively smaller, ranging from .14 to .16. However, non-zero ICC estimates alone do not necessarily indicate a need to account for a given cluster level, as that level's power is also an important factor (Peugh, 2010). As power at the school level in this study was low with only 18 schools, design effects (Peugh, 2010), or estimates of the multiplier necessary to apply to standard errors in order to correct for negative bias introduced by nested data, were calculated. A design effect above 2 paired with a positive ICC indicates a need to include a given level in multileveled analysis. Level-3 design effects within this study fell well below this cutoff, ranging from .3 to .43. Thus, two-level models were used that accounted for variance at the student and classroom levels.

These predictive models assessed the influence of classroom quality on the intercepts of student achievement variables (i.e. means-as-outcomes models), the influence of classroom quality on both the intercepts and slopes of student time in each non-instruction variable, the influence of these non-instruction intercepts and slopes on student achievement variables, and mediational effects of non-instructional time (and change over time) on the original relation between classroom quality and student achievement. Effects of Q-CLE on the intercept of each outcome variable are presented as a fixed effect (analogous to a regression coefficient) whereas the effect of Q-CLE on the slopes of change across the year are modeled in terms of standard deviations from

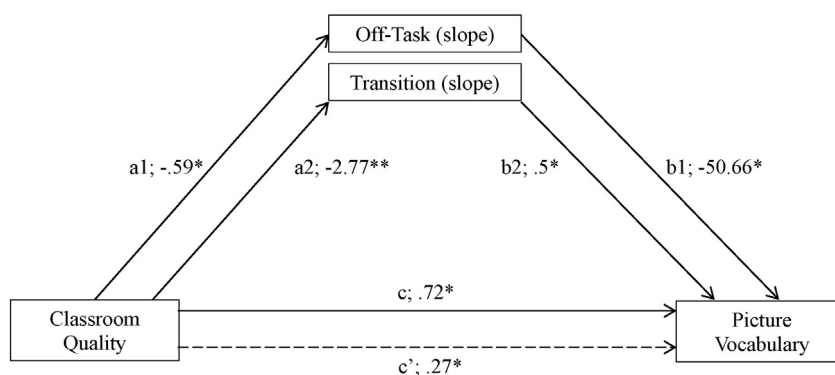


Fig. 1. Estimated fixed effects for each pathway of the mediation model. Note: SE for path $c = .34$, for path $a1 = .28$, for path $a2 = .42$, for path $b1 = 24.41$, for path $b2 = .23$, for path $c' = 34$.

the mean. Higher quality classrooms fell +1 SD above the mean and lower quality classrooms fell –1 SD below the mean (applicable when interpreting Figs. 1 and 2). Interpretation of model results change slightly based on the relative levels of the independent and dependent variables, and interpretations for each of these models are outlined in the following sections. Further, all slopes were assumed to be linear as there were not enough time points to reliably fit quadratic estimations, and were modeled as such in the figures. Effect sizes were calculated and are reported as r -squared, or the proportion of reduction in variance observed between the baseline (unconditional) models and full (predictive) models.

3. Results

3.1. Relations between classroom quality and student literacy achievement

Separate models were run with classroom-level Q-CLE predicting students' literacy achievement on each of the WJ tasks in order to test whether Q-CLE differentially affected student literacy skills. Literacy outcomes were centered around the grand mean and the model quantified student achievement outcomes at the intercept-level (means-as-outcomes), thus results of these models can be interpreted as classroom-level differences in Q-CLE predicting classroom-level deviations from the grand mean of the overall sample in literacy outcomes. These models revealed that Q-CLE predicted student achievement in PV and PC (see Fig. 1, path *c*). Students achieved a grand mean W score of 486 on the PV task with a slope of growth of .77 from fall to spring. Classroom quality had a significant effect on the intercept of this variable (i.e., spring fitted mean) such that lower quality classrooms displayed lower classroom-level average performance on this task (fixed effect = .72, $p < .05$, $r^2 = .15$). The same effect was found for student performance PC—students achieved a grand mean W score of 470, and in lower quality classrooms had lower average scores on this task (fixed effect = 1.55, $p < .05$, $r^2 = .40$).

3.2. Relations between classroom quality and non-instruction

Further models were run investigating the relations between Q-CLE and minutes spent by students in the two non-instruction activities (see Fig. 1, paths *a1* & *a2*). Q-CLE did not have a significant impact on the intercept of student time spent in either of these activities but did have a significant effect on the slopes of change across the year for each of the two variables. Students in higher quality classrooms decreased in the amount of time they spent off-task from fall to spring (fitted slope = .11, Q-CLE slope effect = $-.59$, $p < .05$, $r^2 = .05$), whereas students in lower quality classrooms increased, and students in average classrooms remained relatively consistent (see Fig. 2). Furthermore, students in high quality classrooms decreased the most in terms of their time spent in transitions from fall to spring (fitted slope = -2.02 , Q-CLE slope effect = -2.77 , $p < .001$, $r^2 = .11$), followed by average classrooms; low quality classrooms decreased the least, remaining almost consistent across the year in transition time (see Fig. 3).

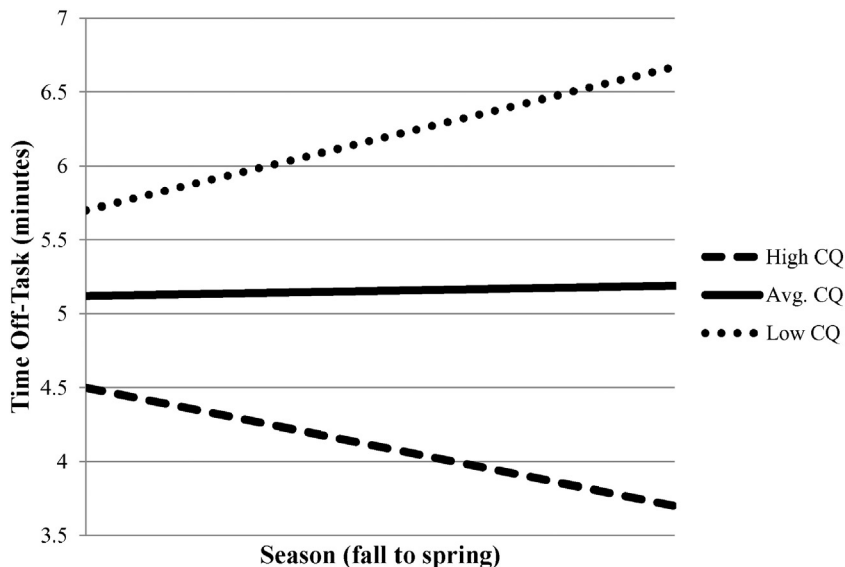


Fig. 2. Change in students' time off-task from fall to spring as a function of classroom quality; high, average and low-quality classrooms compared.

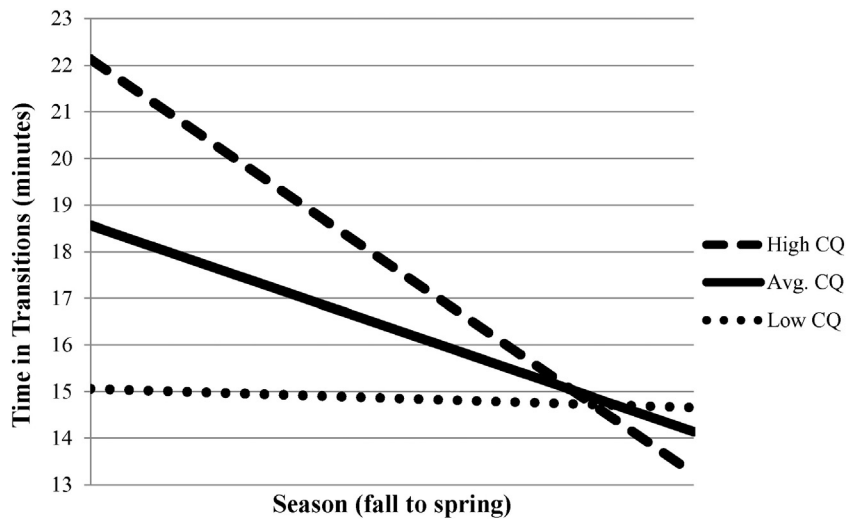


Fig. 3. Change in students' time in transitions from fall to spring as a function of classroom quality; high, average and low-quality classrooms compared.

3.3. Relations between non-instruction and literacy achievement

Models were run with students' time off-task and in transition (now a level-1 independent variable) predicting student literacy achievement, also at level 1. Results of this model can be interpreted as individual student differences in the fall to spring change in time spent participating in non-instruction (off-task and transitions tested separately) predicting individual student differences in spring literacy outcomes (see Fig. 1, paths *b1* & *b2*). Although both variables are at the student level, multilevel modeling was still used to parse out the contribution of classroom-level differences in these relations as ICC estimates for these variables did show some classroom influence. These models were run with both PV and PC as separate outcomes; however time off-task and in transition were modeled together as co-predictors to capture the independent impact of each of these as well as their collective effect on literacy outcomes. While no significant results were found for PC, both time spent off-task and in transition predicted students' spring achievement in PV. Specifically, students in classrooms who showed smaller decreases in time (i.e., less slope) off-task from fall to spring scored lower (slope effect = -50.66 , $p < .05$) compared to students in classrooms with greater decreases in time. At the same time, students who were in classrooms with greater decreases (i.e., slopes) in time from fall to spring in transitions scored higher on this measure compared to students in classrooms with lesser decreases (slope effect = $.5$, $p < .05$). The effect size for these relations was $r^2 = .21$.

3.4. Time in non-instruction as a mediator

All estimates of random effects from models testing each pathway of the proposed mediation are presented in Table 3. Q-CLE predicted student achievement on the PV and PC tasks as well as slopes predicted changes in the amounts of time students spent off-task and in transition. Further, these changing slopes predicted student achievement in PV only, thus mediation is possible in the case of PV (Baron & Kenny, 1986). In order to investigate this potential mediation relation, the slope effects of both non-instruction variables as well as the original Q-CLE variables were all modeled as co-predictors of students' PV skills to see if the slope effects of student participation in non-instruction predicted students' PV outcomes over and above the originally detected effect of Q-CLE (see Fig. 1, path *c'*). Once the slope effects of students' time off-task and in transition was introduced into the model, the original direct relation between Q-CLE and PV decreased in size and became insignificant (fixed effect = $.27$, $p > .05$), with time in non-instruction activities instead significantly predicting student performance (off-task slope effect = -50.25 , $p < .05$; act. trans. slope effect = $.45$, $p = .05$, $r^2 = .1$). These patterns indicate mediation effects (Baron & Kenny, 1986; Judd & Kenny, 1981; Zhang, Zyphur, & Preacher, 2009). The total mediation effect (Krull & Mackinnon, 2001) of the slopes of student time in both non-instructional activities was calculated by subtracting the fixed effect of the initial variable (classroom quality) in the mediation model from the original fixed effect of the initial variable (classroom quality) on the outcome (PV). This resulted in a total mediated effect of $.45$, and this effect was found to be significant at $p = .05$ (Zhang et al., 2009).

4. Discussion

The aim of this study was to investigate the associations among classroom quality, student time spent in two non-instructional activities (i.e., off-task and transitions), and students' early literacy skills. These relations were tested in separate pathways and then together in a mediation model with change in students' time in non-instructional activities mediating a potential relation

Table 3

Estimates of random effects for all HLM models in the proposed mediation model.

Estimates of random effects																							
Path c						Path a						Path b						Path c'					
Model	SD	Variance component	df	X ²	p	Model	SD	Variance component	df	X ²	p	Model	SD	Variance component	df	X ²	p	Model	SD	Variance component	DF	X ²	P
PV on Q-CLE						Off-task slope on Q-CLE						PV on off-task slope and transition slope						Mediation Model					
Intercept r0	1.54	2.38	47	83.11	.001	Intercept r0	.26	.07	331	314.19	>.5	Intercept r0	.14	.02	48	38.62	>.5	Intercept r0	.15	.02	47	38.19	>.5
Level 1 e	6.65	44.29				Slope r1	.34	.11	379	375.20	>.5	Level 1 e	6.35	40.32				Level 1 e	6.35	40.36			
PC on Q-CLE						Transition slope on Q-CLE																	
Intercept r0	3.36	11.27	47	110.72	<.001	Intercept 1/2 u00	5.37	28.88	47	<.001													
Level 1 e	10.95	119.98				Transition slope on Q-CLE																	
						Intercept r0	2.03	4.14	331	368.74	.08												
						Slope r1	4.65	21.63	391	617	<.001												
						Level 1 e	8.77	76.92															
						Intercept1/2 u00	10.15	103.1	47	1126.75	<.001												

Note. The model with PC regressed on Q-CLE is included here but was not a part of the final mediation model tested with PV as the outcome.

between the quality of the classroom learning environment and student literacy achievement. As hypothesized, positive relations were found between Q-CLE and student literacy achievement, with students in higher quality classrooms performing better on measures of expressive vocabulary and reading comprehension. Q-CLE was also associated with the amount of time, specifically in terms of change across the year, that students spent off-task and in transitions during classroom instruction. Higher quality classrooms were characterized by a decreasing slope of student time off-task as the year progressed, while average and low quality classrooms were characterized by stagnant or even increasing student time off-task. Furthermore, while all classrooms decreased in the time spent in transitions from fall to spring, students in the highest quality classrooms decreased the most drastically. These changes in amounts of non-instructional time were related to students' end-of-year expressive vocabulary skills, such that higher performance on this task was predicted by decreasing time in both off-task and in transition from fall to spring. Finally, a full mediation effect of these two non-instructional activity slopes on the relation between classroom quality and student's expressive vocabulary skills was observed.

4.1. Conclusions

These results point to some interesting findings. We hypothesized that students in high quality classrooms would consistently spend less time off-task and in transition. This was true for off-task behavior, but we found that students in high quality classrooms began the year spending much *more* time in transitions than did their peers in other classrooms and ended the year spending the *least* amount of time transitioning compared to classrooms of average or low quality. While this may at first appear counter-intuitive, one possible interpretation of this finding is that a mark of higher-quality classrooms is that students become more efficient in their transitions as the year progresses. Thus, students may need more time in the fall to adjust to the classroom environment but will show steady rates of improvement in these transitions throughout the year and becoming most efficient by spring. A teacher leading such a class may recognize the need for more transitional time in the fall but may also purposefully promote skills within the classroom that improve these transitions.

While not tested directly, the patterns revealed in this study regarding classroom quality and students' time in non-instruction speaks to the relation between classroom quality and the development of self-regulation in young students. We conjecture that, within this study, the decreasing slope of students' transition time across the year in high quality classrooms were evidence of improvement in students' self-regulation skills brought about by teachers' purposeful contributions to classroom quality (see also Connor et al., 2010a, 2010b). This conclusion is supported in related research: Bodrova and Leong (2005) found that students with better self-regulation skills transitioned between tasks more efficiently, and Connor et al. (2010a, 2010b) found that transitions between instructional activities were generally more efficient in classrooms where teacher expectations for students' self-regulated behavior were higher. Furthermore, Rimm-Kaufman et al. (2009) found that higher levels of classroom organization were significantly linked with stronger gains in student self-regulation. Altogether, these findings suggest that in a classroom in which the teacher is purposeful about the positive promotion of classroom elements such as organization and behavior management, student self-regulation improves in a way that can be directly observed through patterns in off-task and transitional behavior. For example, teachers who are more successful at creating high quality classroom environments may spend more time at the beginning of the year establishing and reinforcing classroom routines, such as cleanup schedules or efficient procedures for lining up (indicated both by the 'planning/organizing' and 'classroom management' domains of the Q-CLE), in the interest of solidifying these systems early on, even though it may take some additional 'non-instructional' time (in this case, in transition). As the year progresses, teachers in high quality classrooms may push for more efficiency within the classroom as students become accustomed to routines and traditions, resulting in transitions becoming shorter by the spring. In addition, the teachers' use of successful classroom management could also be indicative of whether students who are off-task are redirected quickly and effectively by the teacher, possibly contributing to the patterns revealed in the associations among classroom quality and student time off-task.

Finally, we found that the student level observation variable (changes in time in non-instruction from fall to spring) fully explained the association between the classroom quality and students' vocabulary gains. Whereas the relation between classroom quality and student outcomes has been firmly established, this study adds insight into *how* this relation operates within the classroom. Keeping in mind that time in non-instruction is necessarily time lost to meaningful instruction, it makes sense that less wasted time should predict stronger achievement. What is notable is that it was decreasing time in non-instruction from fall to spring—not overall or mean time, which was held constant—that fully mediated the relation between classroom quality and student vocabulary gains. Moreover, because non-instruction was measured at the student level, this finding speaks to the possibility that students who do not decrease in their off-task and transitional behaviors, possibly because of their own self-regulation skills along with the teachers' classroom management, may miss out on important instructional opportunities and thus experience less academic growth.

It is important to acknowledge that the relations between classroom quality and student characteristics are likely reciprocal. That is to say, high quality classrooms positively influence students' behavior, but the presence of multiple students with significant behavioral or self-regulatory difficulties may, in turn, impact the quality of the classroom over and above attempts made by the teacher to bolster quality that were investigated in this study (Yoshikawa & Hsueh, 2001). In fact, our foundational theory (Bronfenbrenner & Morris, 2006) explicitly states that no developmental processes are unidirectional—a viewpoint that we take quite seriously when considering our results. The effects of students' family SES, developmental delays or disorders, behavioral patterns, self-regulation, etc. all have the potential to influence the quality of the classroom environment and thus the educational experiences of students. However, we conjecture that it is teachers' *reactions* to these reciprocal effects that separate high from

low quality classrooms in terms of sustained quality across the year. A teacher who is more adept at creating a successful learning environment may be better able to make changes to the classroom based on student needs (e.g., a higher ratio of poorly regulated students than the previous year may indicate that a teacher needs to use different organizational techniques), whereas a less skilled teacher may continue on without attempting to change these patterns. We argue that, regardless of the outside influences coming into the classroom, it is ultimately the teachers' responsibility to make adjustments in the interest of optimizing classroom quality and student experiences while recognizing that having a high proportion of students with poor self-regulation (Skibbe, Phillips, Day, Brophy-Herb, & Connor, 2012) or language skills (Justice, Petscher, Schatschneider, & Mashburn, 2011) will make this more challenging.

4.2. Limitations

This study was correlational in design and thus no definitive causal claims can be made. The reasons behind the associations revealed are purely informed conjectures and should be interpreted with caution. Although large for an observational study, the overall sample size was limited by the number of classrooms observed, hence small but important effects might have been undetected. Finally, these data were collected during the days of the Reading First initiative. Thus, while the research cited previously supports the validity of our primary constructs of interest, our findings may not generalize as precisely to today's classrooms, which are likely to be influenced by more current policy, including the Common Core State Standards (Common Core State Standards Initiative, 2010). Nevertheless, this study included a diverse sample of students, including a high proportion of students from low SES families. The measures used in this study have been widely used across studies, and the primary constructs of interest (most specifically our conceptualization of classroom quality) have held as predictive in past research across multiple early elementary curricula.

4.3. Implications and future directions

The foremost implication of this study is its potential relevance in informing professional development interventions aimed at improving and maintaining the quality of the classroom-learning environment. These results speak to the identifiable "symptoms" of low (or high) classroom quality that a teacher might look for in order to identify a problem, specifically in terms of how much time students are spending off-task and participating in transitions, and how these amounts of time change throughout the year. A primary aim of this line of research is to shed light on the things taking place inside the classroom that have the potential to impact students' educational outcomes. Whereas the more dynamic influences of the classroom environment were limited to students' expressive vocabulary skills within this study, it is an important first step in providing evidence that the dynamic system of the classroom has the potential to impact student academics through multiple pathways, both direct and indirect, and that aspects of this microsystem can be altered to optimize student performance. Given the disconcerting state of early literacy achievement currently seen in schools across the nation, the importance of improving the learning context in which our young students participate becomes clear. This study can contribute to policy and professional development programs by providing specific, practical recommendations to practitioners that are framed in terms of how to identify a need for change in the classroom.

Another practical implication of this study is its ability to inform the systematic observation of classrooms in both research and policy settings. Classroom observation is one of the most common practices used to inform teacher evaluation, yet recent research has revealed that experienced practitioners show almost no accuracy or consistency in their ability to distinguish high from low quality classrooms in direct observation (Strong et al., 2011). This finding is alarming, especially given the implications of teacher performance evaluations that are based on such observations. In research settings, variability exists across conceptualizations and measurement of classroom quality, and information from this study can be used to better inform the accurate portrayal of this very important construct in empirical research. The use of observation to assess classroom quality provides a potentially more concrete and detailed assessment of the classroom as a developmental context, and the use of this method in the present study yielded results that, again, speak directly to what a 'high quality' classroom looks like.

Future work following this vein of study will examine the relations among these variables in different age groups of students. Specifically, we are interested to know whether classroom quality and student time in these non-instructional activities differentially impact students' literacy skills in middle, as opposed to early elementary school. We hypothesize that the mediation effects detected here for expressive vocabulary will be present for reading comprehension in later grades. Our aim in this line of research is to contribute information that can be used to better inform both research and practice, all in the hopes of improving the educational experiences of young students.

References

- Adams, J. W., & Snowling, M. J. (2001). Executive function and reading impairments in children reported by their teachers as 'hyperactive'. *British Journal of Developmental Psychology*, *19*, 293–306. <http://dx.doi.org/10.1348/026151001166083>.
- Anderson, L. W. (1981). Instruction and time-on-task: a review. *Journal of Curriculum Studies*, *13*(4), 289–303. <http://dx.doi.org/10.1080/0022027810130402>.
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173. <http://dx.doi.org/10.1037/0022-3514.51.6.1173>.
- Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology Monograph*, *4*(1, Pt. 2). <http://dx.doi.org/10.1037/h0030372>.
- Baumrind, D. (1978). Parental disciplinary patterns and social competence in children. *Youth and Society*, *9*, 239–276. <http://dx.doi.org/10.1177/0044118X7800900302>.

- Baumrind, D. (1991). Parenting styles and adolescent development. In J. Brooks-Gunn, R. Lerner, & A. C. Petersen (Eds.), *The Encyclopedia of Adolescence* (pp. 746–758). New York: Garland.
- Bodrova, E., & Leong, D. J. (2005). Self-regulation as a key to school readiness: How can early childhood teachers promote this critical competence? In M. Zaslow, & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 223–270). Baltimore: Brookes.
- 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. *The Elementary School Journal*, 104(4), 269–287. <http://dx.doi.org/10.1086/499753>.
- Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. *Handbook of child psychology*.
- Cameron, C. E., Connor, C. M., & Morrison, F. J. (2005). Effects of variation in teacher organization on classroom functioning. *Journal of School Psychology*, 43(1), 61–85. <http://dx.doi.org/10.1016/j.jsp.2004.12.002>.
- 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*, 46(2), 173–192. <http://dx.doi.org/10.1016/j.jsp.2007.03.002>.
- Common Core State Standards Initiative (2010). *Common Core State Standards for Mathematics*.
- Connor, C. M., Son, S., Hindman, A. H., & 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. <http://dx.doi.org/10.1016/j.jsp.2005.06.001>.
- Connor, C. M., Morrison, F. J., Fishman, B., Ponitz, C. C., Glasney, S., Underwood, P., et al. (2009a). The ISI classroom observation system: Examining the literacy instruction provided to individual students. *Educational Researcher*, 38(2), 85–99. <http://dx.doi.org/10.3102/0013189X09323273>.
- Connor, C. M., Piasta, S. B., Fishman, B., Glasney, S., Schatschneider, C., Crowe, E., et al. (2009b). Individualizing student instruction precisely: Effects of child by instruction interactions on first graders' literacy development. *Child Development*, 80(1), 77–100. <http://dx.doi.org/10.1111/j.1467-8624.2008.01247.x>.
- Connor, C. M., Ponitz, C. E. C., Phillips, B., Travis, Q. M., Day, S. G., & Morrison, F. J. (2010a). First graders' literacy and self-regulation gains: The effect of individualizing instruction. *Journal of School Psychology*, 48, 433–455. <http://dx.doi.org/10.1016/j.jsp.2010.06.003>.
- Connor, C. M., Ponitz, C. C., Phillips, B. M., Travis, Q. M., Glasney, S., & Morrison, F. J. (2010b). First graders' literacy and self-regulation gains: The effect of individualizing student instruction. *Journal of School Psychology*, 48(5), 433–455. <http://dx.doi.org/10.1016/j.jsp.2010.06.003>.
- 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 child instruction on third graders' reading comprehension. *Reading Research Quarterly*, 46(3), 189–221. <http://dx.doi.org/10.1598/RRQ.46.3.1>.
- Connor, C. M., Fishman, B., Crowe, E., Underwood, P., Schatschneider, C., & Morrison, F. J. (2013). Teachers' use of Assessment to Instruction (A2i) software and third graders' reading comprehension gains. *In Technology as a support for literacy achievements for children at risk* (pp. 123–139). Netherlands: Springer.
- 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. <http://dx.doi.org/10.1037/a0035921>.
- 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. <http://dx.doi.org/10.1037/a0023042>.
- DiPerna, J. C., Lei, P., & Reid, E. E. (2007). Kindergarten predictors of mathematical growth in the primary grades: An investigation using the early childhood longitudinal study-kindergarten cohort. *Journal of Educational Psychology*, 99(2), 369–379. <http://dx.doi.org/10.1037/0022-0663.99.2.369>.
- Domitrovich, C. E., Cortes, R. C., & Greenberg, M. T. (2007). Improving young children's social and emotional competence: A randomized trial of the preschool "PATHS" curriculum. *The Journal of Primary Prevention*, 28(2), 67–91. <http://dx.doi.org/10.1007/s10935-007-0081-0>.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., et al. (2007). School readiness and later achievement. *Developmental Psychology*, 43, 1428–1446. <http://dx.doi.org/10.1037/0012-1649.43.6.1428>.
- Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, 16(1), 149–167.
- Eccles, J., & Gootman, J. A. (2002). *Community programs to promote youth development*. Washington, DC: National Academic Press.
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <http://dx.doi.org/10.3102/00346543074001059>.
- Fredrick, W. C., Walberg, H. J., & Rasher, S. P. (1979). Time, teacher comments, and achievement in urban high schools. *The Journal of Educational Research*, 73(2), 63–65. <http://dx.doi.org/10.1080/00220671.1979.10885209>.
- Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive-withdrawn behavior and achievement among fourth graders. *The Elementary School Journal*, 421–434. <http://dx.doi.org/10.1086/461853>.
- Gettinger, M., & Seibert, J. K. (2002). Best practices in increasing academic learning time. *Best practices in school psychology IV*, 1, 773–787.
- Guo, Y., Connor, C. M., Tompkins, V., & Morrison, F. J. (2011). Classroom quality and student engagement: contributions to third-grade reading skills. *Frontiers in Psychology*, 2, 157. <http://dx.doi.org/10.3389/fpsyg.2011.00157>.
- Guthrie, J. T., Anderson, E., Alos, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *The Elementary School Journal*, 99, 343–366. <http://dx.doi.org/10.1086/461929>.
- 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, 949–967. <http://dx.doi.org/10.1111/j.1467-8624.2005.00889.x>.
- Hamre, B. K., & 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* (pp. 49–84). Baltimore: Brookes.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis estimating mediation in treatment evaluations. *Evaluation Review*, 5(5), 602–619. <http://dx.doi.org/10.1177/0193841X8100500502>.
- Justice, L. M., Petscher, Y., Schatschneider, C., & Mashburn, A. (2011). Peer effects in preschool classrooms: Is children's language growth associated with their classmates' skills. *Child Development*, 82(6), 1768–1777. <http://dx.doi.org/10.1111/j.1467-8624.2011.01665.x>.
- Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. *Research paper. MET Project. Bill & Melinda Gates Foundation*.
- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281–300. <http://dx.doi.org/10.1080/10888430802132279>.
- Krull, J. L., & MacKinnon, D. P. (2001). Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research*, 36(2), 249–277.
- Ladd, G. W., Birch, S. H., & Buhs, E. S. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Development*, 70(6), 1373–1400. <http://dx.doi.org/10.1111/1467-8624.00101>.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., ... Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732–749. <http://dx.doi.org/10.1111/j.1467-8624.2008.01154.x>.
- Maughan, B., & Carroll, J. (2006). Literacy and mental disorders. *Current Opinion in Psychiatry*, 19(3), 350–354. <http://dx.doi.org/10.1097/01.yco.0000228752.79990.41>.
- McLean, L., & Connor, C. M. (2015). Depressive symptoms in third-grade teachers: Relations to classroom quality and student achievement. *Child Development*, 86, 945–954. <http://dx.doi.org/10.1111/cdev.12344>.
- Merritt, E. G., Wanless, S. B., Rimm-Kaufman, S. E., & Peugh, J. L. (2012). The contribution of teachers' emotional support to children's social behaviors and self-regulatory skills in first grade. *School Psychology Review*, 41(2), 141–159.
- National Assessment of Education Progress (2011). *The nation's report card*. Washington, DC: National Center for Education Statistics.
- 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: U.S. Government Printing Office.
- NICHD Early Child Care Research Network (2002). The relation of global first-grade classroom environment to structural classroom features and teacher and student behaviors. *The Elementary School Journal*, 102, 367–387. Retrieved from <http://www.jstor.org/stable/1002181>
- Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, 48(1), 85–112. doi: <http://dx.doi.org/10.1016/j.jsp.2009.09.002>.

- Pianta, R. C., La Paro, K., & Hamre, B. (2007). *Classroom Assessment Scoring System [CLASS] manual, K–3*. Baltimore: Brookes.
- Ponitz, Claire Cameron, McClelland, Megan M., Matthews, J. S., & Morrison, Frederick J. (2009a). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Developmental Psychology*, 45(3), 605. <http://dx.doi.org/10.1037/a0015365>.
- 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. http://dx.doi.org/10.1207/S1532799XSSR0501_2.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage Publications.
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104(3), 700. <http://dx.doi.org/10.1037/a0027268>.
- Reynolds, A. J., & Ou, S.-R. (2004). Alterable predictors of child well-being in the Chicago longitudinal study. *Children and Youth Services Review*, 26, 1–14. <http://dx.doi.org/10.1016/j.childyouth.2003.11.005>.
- Rimm-Kaufman, S. E., La Paro, K. M., Downer, J. T., & Pianta, R. C. (2005). The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal*, 105(4), 377–394. <http://dx.doi.org/10.1086/429948>.
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., & Brock, L. L. (2009). The contribution of children's self-regulation and classroom quality to children's adaptive behaviors in the kindergarten classroom. *Developmental Psychology*, 45, 958–972. <http://dx.doi.org/10.1037/a0015861>.
- 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. <http://dx.doi.org/10.1037/a0029153>.
- Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, 41(1), 225–234. <http://dx.doi.org/10.1037/0012-1649.41.1.225>.
- Strong, M., Gargani, J., & Hacifazlıoğlu, Ö. (2011). Do we know a successful teacher when we see one? Experiments in the identification of effective teachers. *Journal of Teacher Education*, 62(4), 367–382. <http://dx.doi.org/10.1177/0022487110390221>.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*.
- Yoshikawa, Hirokazu, & Hsueh, JoAnn (2001). Child development and public policy: Toward a dynamic systems perspective. *Child Development*, 72(6), 1887–1903. <http://dx.doi.org/10.1111/1467-8624.00384>.
- Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models problems and solutions. *Organizational Research Methods*, 12(4), 695–719. <http://dx.doi.org/10.1177/1094428108327450>.