

Effective Grade 1 Classroom Contexts for Reading and Language Growth: Evidence from Typical Children and Children at Risk of Attention Difficulties

Louise Deault and Robert Savage*

Associate Professor, Faculty of Education, McGill University, Montreal Quebec, H3A 1Y2, Canada

Abstract

The present study used a nested hierarchical design to assess different aspects of literacy teaching as predictors of change in students' reading and attention in first grade. Observations of literacy teaching were obtained for $n=18$ classrooms using the Classroom AIMS Instrument, which assesses different aspects of teaching quality (Classroom Atmosphere, Literacy Instruction, Management and Student Engagement). For students who started the year with strong reading skills, classroom management predicted higher rates of growth in reading comprehension whereas for students with weaker initial reading ability, student engagement predicted greater reading comprehension growth. For students at risk of attention difficulties, the overall quality of the teaching environment predicted growth in listening comprehension. These results are consistent with goodness-of-fit models of the influence of classroom practices on children's reading and that some classrooms can be a source of resilience for children at-risk of attention problems.

Keywords: Pedagogical; Phonics instruction; Meta-cognitive strategies

Introduction

Research on best practices in early literacy teaching is a practical and important focus for schools, as early reading difficulties can put children at risk for a trajectory of academic failure. Studies have also shown that students who start school with attention problems often experience difficulties with learning to read [1-5], together presenting a serious challenge for teachers in meeting the learning needs of these students. On the other hand, effective classrooms may provide clear 'value-added' educational advantages for all children [6,7]. Furthermore, different classroom contexts may be differentially effective from some children over other children [8,9]. The present study thus considers firstly the role(s) of teachers in diverse classroom contexts in shaping the development of typically developing first grade students and then secondly, of students who may be at risk of attention difficulties in these same classrooms.

Research on Effective Teaching

There exists a sizeable research literature considering teachers' effectiveness in promoting student achievement [7,10,11]. Until relatively recently the quality of this research literature could be viewed as modest at best. A formal systematic review of all of the literature by Hall and Harding identified 1,276 studies from an initial screening [12]. The selection committee identified only 12 studies worthy of in-depth review and just three studies which were judged to be high quality [6,13,14]. Hall and Harding criticize the majority of research carried out in this area for its lack of empirical evidence in defining "effective" literacy teaching, with the majority of studies using peer nominations to categorize teachers as "exemplary." In addition, Hall and Harding noted that very few studies actually measured student achievement. Across the better of the studies, some common characteristics of literacy teachers described as highly effective emerged [12]. These included: creating positive, motivating and supportive literacy environments; offering a balance of instructional elements and experiences with good quality literature; promoting student self-regulation through excellent classroom management skills and responsiveness to student needs; and explicit modeling and teaching of reading and writing strategies [6,13].

Among the key findings of Pressley et al. was that effective literacy teaching is a complex interaction of contexts and specific practices that was made up of 4 overarching components labeled AIMS: Atmosphere, Instruction, Management and Student Engagement [13]. First, effective

teachers created an atmosphere that was welcoming, democratic, emotionally supportive, and that promoted student diversity and cooperation. The teachers nominated as effective engaged in much more explicit strategy instruction in the areas of word recognition, self-monitoring, comprehension and writing. Students were immersed in reading and writing experiences with excellent literature and cross-curricular content. Students read and wrote alone, with peers, or with adults, and learned to write in a graduated series of steps (i.e., planning, drafting, revising and publishing of student writing). Teachers' instructional decisions appeared to be matched to student competence, such that students were provided with appropriately challenging tasks and supported through expert scaffolding to meet increasing expectations. Finally, effective teachers' classroom management promoted student self-regulation and independent problem-solving strategies, and in combination with excellent classroom management skills, created a context where there was visibly high student engagement and enthusiasm for learning [13].

Recent quantitative studies that have looked at growth in learning over time suggest that there is a significant degree of between-class variability in terms of the experiences and activities to which children are exposed in both kindergarten and grade one [15,16]. Teachers who are warm and responsive, and spend more time engaging students in academic activities, tend to have students who demonstrate greater academic growth [17]. Stuhlman and Pianta identified four different typologies of first-grade classrooms that demonstrated varying levels of emotional and instructional support [18]. As expected, teachers who provide both a warm, responsive emotional climate in the classroom and stimulating teaching that is of high instructional quality demonstrate greater academic gains with their

*Corresponding author: Robert Savage, Associate Professor, Faculty of Education, McGill University, Montreal Quebec, H3A 1Y2, USA, E-mail: robert.savage@mcgill.ca

Received August 19, 2013; Accepted September 26, 2013; Published September 28, 2013

Citation: Deault L and Savage R (2013) Effective Grade 1 Classroom Contexts for Reading and Language Growth: Evidence from Typical Children and Children at Risk of Attention Difficulties. J Psychol Abnorm Child 2: 106. doi:10.4172/2329-9525.1000106

Copyright: © 2013 Deault L, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

students. The importance of this dual focus on emotional and academic needs is reinforced by evidence from several studies where teachers who as a group were responsive and sensitive to children, but were less successful in engaging students in learning and in providing appropriate academic support, had students who made less academic progress [15,19,20].

Finally, Downer and Pianta observed that classrooms that spent more time on literacy, language and math instruction were associated with higher reading achievement, phoneme knowledge and long-term retrieval, after controlling for children's academic and cognitive functioning, as well as family and childcare factors [9]. This suggests that there are both quantitative (amount) and qualitative (teaching quality) factors associated with the delivery of optimally effective instructional support. By contrast, although studies of the impact of teaching quality in the early elementary years tend to consistently demonstrate modest effects, program characteristics such as teacher's credentials, class size, child-teacher ratio, aide time, and additional services, are often reported to be unrelated to children's outcomes [15,17-19,21,22].

Goodness-of-Fit Models of Learner x Classroom Effects

"Goodness-of-fit" models emphasize that a child's success in school is influenced by the transaction between individual characteristics and the particular educational context they find themselves in, such that the student's profile of strengths and weaknesses interacts with social, physical, pedagogical and systemic aspects of the classroom [23]. Juel and Minden-Cupp conducted a year-long examination of four first grade classrooms that richly-described the literacy teaching practices that students experienced through the year [24]. Juel and Minden-Cupp identified that children with weak reading skills made the greatest gains when they received intensive phonics instruction at the beginning of the year [24]. In contrast, children who possessed typical reading skills at entry benefited from classrooms that offered more time reading and writing texts. Classrooms that emphasized small-group, differentiated instruction were more successful than classrooms that spent more time on whole-class instruction.

A series of studies by Connor and colleagues, replicate and extend these findings [8,25,26]. Connor et al. report that children who began the first grade year with strong vocabulary and decoding skills tended to fare well regardless of classroom instructional practices [26]. However, teacher's instructional practices had a greater impact on children who exhibited weaker vocabulary and decoding skills at entry. Children with weaker skills at the beginning of grade one made greater growth in their decoding skills when teachers spent more time in teacher-managed, explicit decoding instruction (e.g., teacher-directed alphabet, letter-sound and phonics activities) compared to classrooms characterized by more child-managed and implicit tasks (such as independent reading and writing activities). In contrast, children with high initial vocabulary scores demonstrated the greatest decoding and word recognition skill growth in classrooms that spent more time in child-managed activities.

Similarly, in a series of intervention studies where type of instruction was systematically manipulated to attempt to optimize learning for children with different literacy profiles, children made the greatest growth in reading when teachers were better able to tailor their instruction precisely based on individual students' needs [8,27]. Connor et al. argue the term "quality" literacy instruction needs to be re-conceptualized: what constitutes high-quality instruction for one child may be considered poor-quality for another [27].

In advancing knowledge here, arguably what remains to be established is a richer picture of the observable aspects of teaching that underpins effective teaching of reading in elementary classrooms and, in particular, of the student by classroom interactions identified in current goodness-of-

fit models. It thus could be important to explore some of the conceptually well-established but empirically weakly-evaluated pedagogical constructs identified by the best of the original research by Pressley et al.: Classroom Atmosphere, Instructional content, classroom Management and Student engagement and enthusiasm for learning (AIMS) [13]. The present study seeks, for the first time in the literature, to use a quantitative 'value-added' design to explore whether the AIMS tool predicts reading and related language growth of first grade students. For the first time, this will involve observing the literacy teaching of an unselected sample of regular teachers rather than case studies of nominated 'expert' teachers, to establish generalizable patterns. In addition, the present study explores the contribution of effective teaching on reading performance with a subgroup of students at risk of attention difficulties.

The Effects of Teaching Quality on Children with Attention Difficulties

What role might different classrooms have on the development of children with or at-risk of attention difficulties? Traditional clinical descriptions of ADHD specify both inattentive and hyperactive symptoms as central to a diagnosis [28]. Indeed, learning to listen, sustain attention, and organize materials are all skills that children must master to be successful at school. In this respect, ADHD has been aptly described as "a disorder of conduct in the classroom", since this is the context where children's symptoms are typically identified and expressed [29]. Although it is recognized that the core behaviors of ADHD are manifested within school contexts, arguably, the contribution of the environment has not been fully integrated into clinical models of ADHD.

Some researchers have speculated that unique environmental factors may play a larger role in helping us to map developmental trajectories and to trace the risk and protective factors that contribute to different outcomes [30-32]. Reviews of family factors associated with ADHD, have highlighted high levels of family conflict, parenting stress, and parental psychopathology as co-occurring factors [33,34]. However, few studies have considered the impact of different environmental factors on the development of functional impairments in ADHD, such as academic underachievement or evaluated the natural variation within schools as a means of understanding their contribution to the academic outcomes of children at risk of attention difficulties [33,35]. Zentall reports that children with ADHD have difficulty sustaining attention to long, repetitive or passive tasks and instead prefer learning that involves social engagement, movement and stimulation [36]. Students tend to display more inattentive and off-task behavior in particular contexts: during passive tasks that provide less structure; when there is less teacher direction, such as during centres, seatwork and free time; when adults are unclear in communicating instructions; and when there is a mismatch between the teacher's expectations and student's abilities [36-39].

While experts in the field emphasize school-based interventions as a critical component of the treatment plan for children with ADHD, recent reviews evaluating the evidence base on academic interventions emphasize that there are very few studies documenting their effectiveness in improving academic achievement [40-43]. Raggi and Chronis report that approaches associated with increased on-task behavior and academic outcomes include: (a) class wide peer tutoring; (b) instructional and task modifications; (c) self-monitoring and reinforcement; (d) strategy training, including study and organizational skills; and (e) homework-focused interventions that involve parents [42]. These strategies share an emphasis on promoting active student engagement, collaboration among students, parents, and teachers, as well as facilitating the development of meta-cognitive strategies. Similarly, Tannock and Martinussen speak to the need for multi-pronged interventions integrating both child-focused (e.g., self-monitoring of behavior, academic skills training) and contextual strategies

(e.g., instructional supports to maximize student engagement) to promote academic success [44].

In spite of the recognition that multi-faceted classroom interventions are important, intervention designs often instead isolate specific strategies using a small-group approach and tightly controlled conditions. Although such studies with clinical populations are valuable, they are often conducted with small samples, which can limit their generalizability and their relevance to a broader group of children with attention problems in inclusive classrooms. In addition, it is not clear how specific teaching strategies identified through intervention studies fit into the rest of the classroom context or to other elements of the teacher's literacy instruction. Alternatively, the benefit of using naturalistic observation is that it allows for variation in teaching practices in real classrooms to be documented in order to better identify strategies that teachers use effectively with students amidst the demands of a real classroom.

There are relatively few ecologically valid studies that consider teacher effectiveness with respect to students' acquisition of reading for children with low attention skills. Existing studies do suggest that children whose classrooms offer more emotional support have better social skills and fewer problem behaviors, including both internalizing and externalizing symptoms [45,46]. Perry et al. found that in classrooms where more supportive teaching practices were observed, children demonstrated more positive interpersonal behavior (i.e., ability to socialize with peers), and better behavioral adjustment (defined here as symptoms of depressed mood or anxiety) [47]. Similarly, Wilson et al. reported that children in classrooms marked by high levels of emotional support and evaluative feedback displayed higher engagement and positive peer interactions, and fewer instances of negative or disruptive behaviors, relative to classrooms with fewer supports [20].

Of the few studies specifically exploring attention, the NICHD data explored teaching quality as a predictor of cognitive measures of attention [48]. The study identified that both instructional and emotional supports modestly predicted first grade students' performance on tests of sustained attention, impulsivity and memory, after controlling for the quality of the family environment. Rudasill et al. conducted a longitudinal study that examined the relationships between children's temperament at preschool (specifically their attention and activity levels) and academic outcomes in third grade, relative to the degree of emotional support children received in their third grade classrooms [49]. They found that classroom emotional support in third grade moderated the relation between children's early attention and later reading and mathematics achievement. Specifically, inattention was associated with lower academic achievement for children who were in classrooms that were less emotionally supportive. The authors interpreted these results to suggest that highly supportive classroom climates may buffer children from the risk that poor attention poses for academic difficulties. While significant, the interaction between attention and classroom emotional support was modest and explained less than 1% of the additional variance in reading achievement. However, as temperament was measured at age 4.5 years using maternal report, this may have underestimated the underlying relationship.

Aims of the Present Study

This study examines variation in students' literacy and attention skills and observed classroom factors for both typically developing students and those who may be at risk for academic difficulty due to attention problems. In addition, this study seeks to add to the literature by exploring the effects of different natural classroom contexts (i.e., highly effective vs. less effective in overall observed AIMS quality) on children's reading development in grade 1 for students considered at-risk for attention problems. The following two research questions are addressed:

- Are there cross-level (student x classroom) interactions between observed classroom-level literacy teaching factors (Atmosphere, Instruction, Management and Engagement) and student-level variation in students' reading and attention skills in grade one, controlling for initial skill levels?
- Do students at risk of attention problems who experience contrasting classroom environments (i.e., those rated as highly effective vs. less effective based on observation) show different outcomes with respect to their reading and attention skills?

Method

Participants

This classroom-based research project involved both students and teachers in bilingual (English/ French) grade one classrooms in Quebec, Canada. The total sample consisted of 284 grade one students (50% male, 50% female) from 18 classrooms in 11 schools. The mean age of students was 77 months (6 years, 5 months) with an age range of 5 years 7 months (67 months) to 6 years, 11 months (83 months). Participating teachers ($n=18$) were all female (the norm most without exception in classrooms in this region in this student age group), with varying years of teaching experience ($M=15.39$, $SD=11.66$). Classrooms varied in size, ranging from 15 to 22 students ($M=19.74$, $SD=2.14$) with the majority of instruction taking place in English and smaller proportions of French. All English or Bilingual first grade classrooms at each school participated, but French immersion classes were excluded from the study. There was also considerable variation in the style of language arts instruction provided; only 19% of teachers were using a board-mandated reading program with the rest using an instructional style of their own choosing.

Parents were asked to complete a background questionnaire that included information on parental language background, languages spoken at home, maternal education level, and the frequency of home reading. All students were eligible to participate (provided parental consent was obtained) with the result that 65% of students across classrooms participated. Information on language background revealed that 29% of families report speaking only English at home, 7% only French, and 2% speak a third language. The majority of the sample (60%) described their families as bilingual or trilingual. Maternal education was also obtained according to a 7-point ordinal scale: (1) elementary school only (1.1%); (2) did not receive high school graduation diploma (5.5%); (3) received high school graduation diploma (14.5%); (4) technical training; (5) college/ CEGEP (39.6%); (6) undergraduate degree (19.6%); (7) graduate degree (7.6%). Maternal education was normally distributed in the sample ($M=3.73$, $SD=1.34$).

Measures

Reading assessments: Reading skills were assessed using a standardized reading test: The Group Reading Assessment and Diagnostic Evaluation [50]. Students were evaluated on word reading, word meaning and listening comprehension at pre-test, as well as sentence and passage comprehension at post-test. Administration time is 60-90 minutes and was completed in a whole-class format. The examiner read the instructions to the group with students marking their answers individually in a student response booklet. Split-half reliability coefficients, corrected by the Spearman-Brown formula, are high for this measure ($r=0.95$), indicating that there is a high degree of homogeneity among items in the first-grade form of the GRADE. Test-retest reliability is also high ($r=.96$) for the first grade version of the test.

Word reading: The word reading subtest assesses children's recognition of sight words and ability to decode regularly spelled words. For each of

the 20 items on this subtest, the examiner reads a target word, reads it in a sentence and then repeats the word. Students must identify the target word from a list of four or five choices and select it by drawing an “x” in the multiple-choice item. The Spearman-Brown odd-even internal consistency correlation was high ($r_s=0.75, p<0.05$).

Word meaning: This subtest assesses students’ word recognition, as well as their understanding of early reading vocabulary. Students complete this 27-item task independently after the examiner has modeled two examples. Students are asked to silently read a target word and then choose its matching picture from four choices. Raw scores on the Word Reading and Word Meaning subtests are combined to obtain a composite Vocabulary score. The Spearman-Brown odd-even internal consistency correlation in this sample was high ($r_s=0.78, p<0.05$).

Listening comprehension: This subtest assesses students’ understanding of spoken language. Students listen to a sentence and then choose a picture from four choices that best illustrates the meaning of the sentence. Scoring of this test is based on correct identification of the picture for each item, for a total of 17 items. The Spearman-Brown odd-even internal consistency correlation was moderate ($r_s=0.68, p<0.05$) in this sample.

Sentence comprehension: The sentence comprehension task is a cloze task consisting of 19 items. After the examiner modeled two examples, students were asked to independently read a series of sentences that have a missing word represented by a blank space. Students must read each sentence and choose a word from four single-word choices that best fits into the context of the sentence. Students were asked to complete the sentence comprehension subtest at post-test only, as the reading demands of this task may exceed the capabilities of some students at the beginning of grade one.

Passage comprehension: For this subtest, students were asked to read a passage and to respond to multiple-choice questions, drawing on different elements of reading comprehension (e.g., questioning, clarifying, summarizing and predicting). This task provides challenging content for first grade students and thus was administered at post-test only. The test consists of 24 items. Raw scores on the Sentence and Passage Comprehension subtests were combined to yield a Comprehension Composite score [51].

Teacher ratings of attention and behavior: Teachers were asked to rate children on their attention and behavior using the *Conners’ Global Index* (CGI-T). The CGI-T is a short (10-item) and efficient scale that has been found to be effective at discriminating between children with ADHD and non-clinical samples [52]. Correlations between the CGI-T ratings on parent and teacher versions of the long form range from 0.28 to 0.50 [51], suggesting that the parent and teacher measures provide distinct but related descriptions of children’s behavior in different contexts. The Spearman-Brown internal consistency of this measure was high ($r_s=0.80, p<0.05$) within the present sample.

Classroom observation scale: Classrooms were observed using a structured observation tool, the *Classroom AIMS Instrument*, in order to try to operationalize the natural variation in literacy teaching [1]. The original items were all drawn from the characteristics of teachers nominated as effective across numerous studies and were inductively categorized to create four categories (i.e., Atmosphere, Instruction/Content, Management and Student Engagement). As defined by the AIMS, Atmosphere refers to the physical and interpersonal environment of the classroom, encompassing attributes such as a sense of community, interest, focus on student effort rather than performance, opportunity for student choice, emphasis on the value of learning, high expectations for all students and use of informative feedback. Instruction/Content refers to the lessons and activities included

in the literacy program, as well as the teacher’s instructional style. Subcategories included in this domain reflect elements such as the degree to which the literacy content and activities are engaging for students, the density of instruction, cross-curricular connections, modeling and explicit teaching of thinking processes, scaffolding and setting an appropriate level of challenge for individual students, provision of academic monitoring and encouragement of academic self-regulation. Management refers generally to the organization, rules, routines and procedures that guide the running of the classroom environment, including teachers’ use of monitoring for on-task behavior and promotion of behavioral self-regulation. Student Engagement is defined as observable indicators of student engagement, including participation, excitement, and staying on task. In its initial validation study, the reliability of the AIMS instrument was strong for each category: Atmosphere ($\alpha=0.87$); Instruction ($\alpha=0.90$); Management ($\alpha=0.74$); and Student Engagement ($\alpha=0.79$); suggesting that the items within each category reflect meaningful and consistent interpretations of different elements of teaching (Roehrig and Christesen, 2010) [52]. Similarly, in the present sample, the AIMS demonstrated excellent reliability across each of the four constructs, using Cronbach’s Alpha: Atmosphere ($\alpha=0.94$); Instruction ($\alpha=0.94$); Management ($\alpha=0.86$); and Student Engagement ($\alpha=0.79$).

There is evidence that the AIMS tool has good psychometric validity. Roehrig et al. conducted an initial validation of the AIMS instrument, asking multiple experts in elementary teaching (including both academic experts and elementary teachers known to be very effective) to validate it [1]. Responses supported the face validity of the four categories, 17 subcategories, and 130 items included in the instrument. Evidence for discriminant validity comes from studies by Bohn et al. [53] and Roehrig et al. [54] who showed in both cases that the tool discriminated between effective and less effective novice teachers. Finally, evidence for construct validity comes from psychometric analyses of AIMS sub-scales relationships reported in detail by Roehrig and Christesen [52-54].

Procedure

The research design involved three distinct phases of data collection: pre-test literacy assessments (October), classroom observations (January-March) and post-test literacy assessments (May). At post-testing, the Reading comprehension tasks were added to reflect students’ progress in reading. All assessments and observations were completed by graduate students in educational psychology who had been trained with the materials.

Classrooms were observed for 4-6 hours during literacy teaching by a pair of observers who took detailed notes on the activities, verbalizations, behaviors and interactions of teachers and students. Teachers were informed that the observations were confidential and that the purpose of the study was to explore the naturally occurring variation in practices in teaching approaches in schools. Teachers were asked to do what they would normally do, and not to change their practices in any way. Teachers were free to refuse consent to take part in the study, but none of the teachers approached chose to do so.

After observing the teacher for typically 3 to 4 periods of language arts, observers then independently completed the long form of the AIMS rating scale [13]. As the observations took place between January and March, this allowed for a range of classroom activities in language arts to be observed, not just a small window of observations at one point in time. All observations were ‘*in vivo*’ and were not videotaped. Observers rated each of the items on a scale from 1-3 (1=seldom representative, 2=somewhat representative, 3=consistently representative) depending on the degree to which they thought that item characterized the teacher and classroom. If there was not enough information to rate a particular item, it was scored as

'0', indicating that this practice was not observed during the course of the evaluation.

Thorough training of all observers was undertaken through exposure to- and repeated review of-, the Classroom AIMS Instrument and the background research for it. After initial training, subsequent observer team meetings were undertaken in order to discuss and deliberate on the meaning of the items to ensure that there was a shared understanding of each of the items. Research assistants also practiced doing classroom observations by viewing videotapes of elementary classrooms and then further discussed and clarified item meanings on individual AIMS ratings. For the purposes of data analysis, an "agreed" rating was obtained for each teacher, using the input from the two observers. The agreed rating was achieved through a meeting where the two observers went through each of the items to compare their ratings and to discuss what they had observed. Where observers did not assign the same rating for a particular item, the pair discussed their reasons for assigning a particular rating, using their field notes as evidence of the frequency of a given behavior, until the two observers were in agreement. When agreement had been reached and a combined rating of the classroom was complete, scores were calculated by averaging across items in each of the categories. Items that had been coded as "0", indicating that not enough information had been gathered to rate this item, were omitted from the calculated averages for each category. Thus, the final product determined from the classroom observations was a set of four scores for each teacher, representing their average "agreed" ratings on the categories of Atmosphere, Instruction, Management and Student Engagement.

Although they were not used in subsequent analyses, observers' independent ratings were compared to estimate the inter-rater reliability of the AIMS scale in this study. Inter-rater reliability data was calculated by obtaining the correlations with Spearman's Rho between Observer 1 and Observer 2 for each category. There were moderate inter-rater correlations for the Atmosphere ($r_s=0.719, p<0.05$), Instruction ($r_s=0.639, p<0.05$) and Student Engagement categories ($r_s=0.521, p<0.05$) and smaller correlations for Management ($r_s=0.353, ns$). These results indicate that there was moderate consistency among observers' ratings of teachers on some scales, prior to coming together to obtain the "agreed" rating, which was the metric used in subsequent analyses.

Results

Two separate research questions were addressed in this study, which focused on two distinct populations, typically developing readers and then a subset of that first sample, students at risk due to attention difficulties. For typically developing students, Hierarchical Linear Modeling (HLM) was used to explore individual- and classroom-level variation in students' development of reading and attention skills. Secondly, a sample of students at risk due to elevated attention difficulties were identified using the procedure outlined below and then split into two groups representing classrooms of contrasting quality (i.e., highly effective teaching compared to less effective teaching based on AIMS ratings) and were compared on key outcomes using analysis of variance.

Simple correlations were first obtained in SPSS using two-tailed Pearson product moment correlation coefficients. Inspection of the pattern of correlations was conducted in order to determine if there were variables that are highly correlated, to determine if data reduction techniques were warranted. Inspection of the pattern of correlations revealed that the different subtests were moderately correlated, with the highest correlations yielding values of .78. Since correlations of 0.90 or greater are considered problematic due to multicollinearity, the present data yield acceptable correlations between subtests [55]. Thus, separate subtest scores were retained for analysis in order to examine different aspects of reading

outcome in relation to teaching quality.

The aim of the HLM analyses was to address the primary research question: Are there cross-level interactions between observed classroom-level literacy teaching factors (Atmosphere, Instruction, Management and Engagement) and student-level variation in students' reading and attention skills in grade one, controlling for initial skill levels? This question was answered using a two-level hierarchical model that enabled modeling of the interaction of student-level and classroom-level variance. Both Level 1 (i.e., individual-level predictors, such as gender and pre-test scores) and Level 2 (i.e., classroom-level predictors, such as teaching factors) were added sequentially to the model to compare the results at different levels of analysis. At Level 1, four possible student-level predictors were entered simultaneously into the One-way ANOVA Model with Random Effects: (a) pre-test variable, (b) gender, (c) maternal education, and (d) home language to explore their relationship to the dependent variable.

$$Y_{ij} = \beta_{00} + \beta_{01}(\text{PRE-TEST}) + \beta_{02}(\text{GENDER}) + \beta_{03}(\text{MATEDUC}) + \beta_{04}(\text{HOMELANG}) + r_{ij}$$

For each dependent variable, the pre-test covariate was retained following the results of the ANCOVA analysis. However, it should be noted that in order to retain a parsimonious model and to maximize degrees of freedom, covariates that were not significantly associated with the dependent variable (as these were: gender, maternal education and home language) were dropped from the final models. At Level 2, the Intercepts-and-Slopes-as-Outcomes Model, four classroom-level predictors were simultaneously added to the model: (a) atmosphere, (b) instruction, (c) classroom management, and (d) student engagement.

$$\beta_{00} = \gamma_{00} + \gamma_{01}(\text{ATMOS}) + \gamma_{02}(\text{INSTR}) + \gamma_{03}(\text{MNGT}) + \gamma_{04}(\text{ENGMT}) + u_{0j}$$

This model was fit to determine the effects of the Level 2 predictor variables on several student outcome variables, after controlling for initial skill level using the pre-test covariate: (a) vocabulary, (b) listening comprehension, (c) reading comprehension, (e) teacher attention rating. Since reading comprehension was not assessed at pre-test, the GRADE vocabulary composite score was selected as the pre-test covariate due to its high correlation ($r=0.75$) with reading comprehension. These analyses explored whether there were significant cross-level interactions between student-level characteristics and classroom-level factors. Where possible, predictor variables were left uncentred so that raw scores could be used which have a meaningful zero point value. In the case of the teacher ADHD rating, t-scores were analyzed using grand-mean centering. Likewise, the Level 2 variables (Atmosphere, Instruction, Management and Student Engagement) were analyzed using grand-mean centering since the "0" value within the AIMS rubric means that data was not available to measure that specific item accurately.

Results of the HLM analyses

It was first important to establish that significant between-group variability existed between classrooms with respect to student reading and attention means in order to justify the need for hierarchical linear modeling. This was explored by analysis of the unconditional model which investigates whether there is significant between-group variability on student outcomes. The results indicate that the intercepts were statistically significant for all dependent variables, meaning that there is a statistically significant degree of variability between classrooms on each of the reading and attention outcomes. Additionally, inspection of the Intra-class Correlation Coefficients (ICC) was used to estimate the degree of between-group variability. These results indicated that between- classroom variability accounts for a significant proportion of the variance in reading measures at post-test: Vocabulary (8%), Listening Comprehension (5%), and Reading Comprehension (14%), which suggests that there is sufficient

between-class variability to warrant consideration of classroom level effects using HLM. Between-classroom variability also accounted for a significant proportion of the variance in the Teacher Conners' ADHD Index (12%). The estimates of the variance components provide an additional descriptor of the possible nested nature of the data: All variables demonstrate statistically significant variation between classrooms. Since the unconditional model demonstrates that there is significant variability both within and between classrooms, both student- and classroom-level predictors were added to subsequent models to attempt to explain this variability. This finding illustrates that hierarchical linear modeling is an appropriate technique, since there is significant variation between classrooms, as well as within classrooms. Therefore, modeling at both levels of analysis is appropriate.

To address the first research question, the Intercepts- and Slopes-as-Outcomes Models, depicting the cross-level interaction between student-level and classroom-level characteristics on student outcomes are summarized in Table 1 and 2 (reading outcomes). After controlling for students' pre-test scores on each measure, classroom characteristics (assessed with the AIMS) were not strong or consistent predictors of the student Vocabulary composite measure. However, for reading comprehension, both classroom management and student engagement were significant at the level of the intercept, indicating that these variables are associated with students' outcomes on this measure. However, given that the intention was to explore classroom-level predictors in relation to change in students' reading skills across time, the finding of interest to this research question involves examination of the cross-level interactions with respect to the slopes (i.e., after controlling for individual-level covariates). There were two significant interactions between students' reading skills at the beginning of the year (measured with the vocabulary composite) and the different classroom factors that contributed to higher rates of growth in reading comprehension. Inspection of data plots within HLM 6 where median splits of pupil-level attainment produced 'higher' and 'lower' attainment groups showed that classroom management positively interacted with initial scores on the vocabulary measure, such that for students with higher reading skills at the beginning of the year, classroom management significantly predicted higher outcomes in reading

comprehension. In contrast, for students with lower reading skills at the beginning of the year, there was a negative interaction such that students showed more reading comprehension improvement in classrooms with higher student engagement. These results indicate a cross-level interaction: both individual- and classroom-level characteristics are predictors of students' reading comprehension outcomes.

Analysis of an at-risk subgroup

For the second analysis, a subgroup of students from the main sample identified as at risk for attention and behavior problems were examined to look at their attainment in different classroom environments. The at-risk subgroup was defined as any students across the 18 classrooms whose pre-test scores fell into the at-risk range ($t > 65$) on the Conners' Global Index teacher rating. This resulted in a modest sample size ($n = 31$) of at-risk students, with 18 boys and 13 girls (Table 3 and 4). An independent variable assessing classroom quality was constructed by performing a median split to define two specific at-risk groups: (a) teachers whose classrooms were globally rated as effective on the AIMS scale; (b) teachers whose classrooms were globally rated as less effective on the AIMS scale. The median split was based upon the sum of the 4 AIMS subscale scores. The two groups were found to be significantly different on the AIMS measure itself, $F(1,30)=68.62, p<0.001$, indicating that the median split had effectively divided the sample of teachers into two meaningfully different groups that varied across the characteristics of the AIMS.

The two groups were first compared on pre-test scores to determine whether it was necessary to include pre-test scores as a covariate in this analysis of at-risk children. Raw scores were used to maintain consistency with the HLM analyses, with the exception of the *Conners'* measures where t -scores were used. The two AIMS groups were found to differ with respect to the level of maternal education, $F(1,30)=6.66, p<0.05$, with higher levels of maternal education in the high AIMS classrooms ($M=4.17, SD=1.19$) compared to the low AIMS classrooms ($M=2.84, SD=1.50$); therefore, this variable was used as a covariate in subsequent analyses.

A series of univariate ANOVAs were conducted with maternal education as the covariate in each analysis. Students in the two at-risk groups were found to differ in their scores in listening comprehension at post-test, $F(1,30)=4.55, p<0.05$, after controlling for maternal education levels at pre-test. Results indicated that students in classrooms rated as being highly effective on the AIMS performed better on the listening comprehension subtest ($M=14.83, SD=3.88$) relative to students in classrooms rated lower overall on the AIMS ($M=13.00, SD=3.51$). Partial

Pre-test Variables	M	SD
GRADE Vocabulary subtest	26.31	
GRADE Listening Comprehension subtest	13.96	10.01
Post-test Variables		2.94
GRADE Vocabulary subtest	41.36	
GRADE Listening comprehension	15.44	6.97
GRADE Reading comprehension	26.21	1.93
Teacher <i>Conners'</i> ADHD index T-score	52.98	10.84

Table 1: Imputed Means for Pre- and Post-test Variables with standard deviations (Raw scores unless otherwise stated).

Parameter	VOC	LC	RC
Fixed Effects			
For Intercept, γ_{00}	29.73 (1.66)**	2.33 (0.15)**	5.07 (1.42)**
ATMOS, γ_{01}	-18.85 (13.84)	-0.14 (1.09)	-23.71 (12.32)
INSTR, γ_{02}	14.67 (15.05)	2.70 (1.22)	15.75 (13.19)
MNGT, γ_{03}	-10.57 (7.48)	0.05 (0.60)	-14.74 (6.62)*
ENGMT, γ_{04}	12.39 (8.98)	-0.55 (0.82)	17.17 (7.76)*
For fall pre-test slope, γ_{10}	0.44 (0.05)**	-0.59 (0.01)**	0.80 (0.04)**
ATMOS, γ_{11}	0.50 (0.39)	0.06 (0.08)	0.68 (0.37)
INSTR, γ_{12}	-0.35 (0.42)	-0.15 (0.09)	-0.46 (0.40)
MNGT, γ_{13}	0.26 (0.22)	-0.02 (0.04)	0.42 (0.21)*
ENGMT, γ_{14}	-0.36 (0.25)	0.05 (0.06)	-0.49(0.23)*
Variance Components			
Intercept (u_{00})	35.28	0.00	10.16
Level-1 (R)	23.74	0.20	43.39

ATMOS: Classroom atmosphere; INSTR: Literacy instruction; MNGT: Classroom management; ENGMT: Student engagement

Table 2: Results of the Intercepts- and Slopes-as-Outcomes Model for Reading Outcomes.

	Low AIMS Classrooms ^a		High AIMS Classrooms ^b	
	Pre	Post	Pre	Post
MED	2.84 (1.50)		4.17 (1.19)	
VOC	18.21 (7.74)	37.79 (8.80)	19.50 (7.86)	34.17 (9.68)
LC	11.70 (3.59)	13.00 (3.51)	13.42 (2.68)	14.83 (3.88)
RC		19.53 (9.59)		17.51 (8.26)
CGI-T	71.74 (6.03)	71.42 (9.42)	76.50 (9.03)	71.58 (13.34)

MED: Maternal Education level; VOC=Vocabulary; LC: Listening Comprehension; RC: Reading Comprehension; CGI-T= Teacher *Conners'* ADHD index, ^a $n=19$, ^b $n=12$

Table 3: Student-level Means and Standard Deviations for the At-risk Groups.

	Low AIMS Classrooms ^a	High AIMS Classrooms ^b
ATMOS	1.71 (0.28)	2.19 (0.21)
INSTR	1.85 (0.28)	2.18 (0.19)
MNGT	1.81 (0.38)	2.44 (0.16)
ENGMT	1.75 (0.36)	2.54 (0.14)
GLOB	7.12 (1.17)	9.34 (0.60)

ATMOS: Classroom atmosphere; INSTR: Literacy instruction; MNGT: Classroom management; ENGMT: Student engagement; GLOB: AIMS global rating, ^a $n=19$, ^b $n=12$

Table 4: Classroom-level AIMS Means and Standard Deviations for the At-risk Groups.

eta squared (η^2) was also used to provide an estimate of effect size. Analysis of listening comprehension effect sizes with partial eta squared indicated that classroom effectiveness explained 14% of the variance in performance on the listening comprehension subtest ($\eta^2=0.14$).

Discussion

The primary question in the present study was to explore whether there were cross-level interactions between observed classroom-level literacy teaching factors (Atmosphere, Instruction, Management and Student Engagement) and individual-level variation in students' reading skills at the end of grade 1, controlling for initial skill levels at the beginning of the year. Results of hierarchical analyses showed that both effective classroom management and high levels of student engagement are differentially associated with growth in reading comprehension skills for first grade students who present with differing reading abilities.

Specifically, there was a positive interaction between students' initial reading vocabulary scores (i.e., word reading and word meaning) and classroom management: Students with stronger reading skills at the beginning of grade one made greater reading comprehension growth in well-managed classroom environments. In this context, Management on the AIMS refers to "the order, rules, routines and procedures, and what keeps the instruction moving in an orderly fashion". Notably, teachers' use of monitoring for on-task behavior and promotion of behavioral self-regulation are two of the key subscales in this construct. Therefore, classrooms where teachers have more effectively managed students and promoted self-regulation and task-oriented behavior are potentially more likely to have developed students' capacity to focus during sustained and challenging academic tasks, such as reading comprehension.

The second key finding that emerged revealed a negative interaction between students' initial reading ability and student engagement. In contrast to the above results, students with lower reading skills at the beginning of grade one demonstrated greater reading comprehension gains in classrooms that had higher student engagement ratings on the AIMS. Student Engagement is defined on the AIMS as "observable indicators of student engagement, including participation, excitement, and staying on task." Importantly, several of the AIMS items speak to student engagement as going beyond just productivity, to emphasize student excitement and enthusiasm to participate, suggesting that engagement may reflect the degree to which the teacher has been able to interest children in literacy in a broader sense. Classrooms that are highly engaging are hypothesized to have teaching that is also positive, motivating and which stimulates literacy behavior on the part of students [56]. Teachers perceived as highly effective tend to encourage, require, and facilitate children's active participation in learning [55]. These processes may be more important for students at risk [57]. The present results suggest that students whose reading skills are weaker at the beginning of grade one benefit from environments that can seek to engage them actively and enthusiastically in literacy instruction.

Our results are thus consistent with current theoretical work that emphasizes an individual differences perspective and recognizes that what is deemed "teaching quality" does not affect all children equally [11,26,58]. Consistent with the results of Conner et al.'s findings, strong readers in the present analysis made greatest growth in reading comprehension in well-managed classrooms where self-regulation was emphasized. Further studies evaluating the interactions between reading ability, self-regulation, management style and types of instruction (i.e., child-managed vs. teacher-directed) could potentially explore these patterns.

Our results did not report a relationship between quality of teaching as identified by the AIMS tool and growth in word-level reading and understanding, measured using the Vocabulary test or in listening

comprehension. Reasons for this are probably several but must remain necessarily speculative at this point. Candidate explanations include the possibility that AIMS is most sensitive at this point to only the most global measures such as reading comprehension, rather than to measures such as word reading or individual word meanings that are best seen as components of reading comprehension [50]. It may also be that reading comprehension is the best cumulative index of teachers' efforts to improve literacy most generally and emerges for this reason here. Promoting reading comprehension is the ultimate goal of reading instruction, and in relation to more specific word-level and language skills, reading comprehension is often the hardest ability for teachers to improve [59]. These results here are thus nevertheless very positive. The study adds to the field by being the first quantitative study in the literature reporting effects of variation in teaching quality in language arts on reading comprehension growth using the AIMS observation tool. This finding is important because the AIMS tool was designed to reflect best practices in reading teaching, and therefore provides a rich picture of the teaching practices, and of the differentiation needed for less- and more- literate students in grade 1.

Teaching quality and students at risk of attention difficulties

A second analysis was undertaken that focused on students who may be at risk for academic difficulty due to poor attention and contrasted their outcomes in classroom environments that were rated as differing in global quality based on observation. Results indicated that classroom quality (as measured by high or low total AIMS rating) had a significant impact on the outcomes of students at risk of attention difficulties. Specifically, the results showed that children who were equally at risk for attention difficulties at the beginning of the year had stronger listening comprehension skills at the end of the year if they experienced classrooms that were rated as more effective compared to classroom environments that were rated as less effective. These results suggest that the overall quality of teaching students experience with respect to the classroom atmosphere, instruction and management had a significant impact on students' ability to develop their listening skills in first grade. Classrooms appear to differ in the extent to which they facilitate children's ability to sustain attention to standardized listening comprehension tests. The fact that teaching effectiveness accounted for 14% of the variance in listening comprehension between the two groups indicates a large effect size that is potentially practically important for students at risk.

Differences between groups did not extend to reading achievement in this analysis, nevertheless the results from the full sample (research question 1) and for the at-risk sample (question 2) provide at least some consistency as in both cases, the domain of cognition affected most by teaching quality as assessed in AIMS was at the level of comprehension rather than, for example at the level of word reading. There are lots of independent reasons to see Reading Comprehension and Listening Comprehension as closely-related constructs in models of reading [60]. As noted earlier, it has often been found in reading intervention research to be much harder to improve comprehension than word reading skills. It is thus encouraging that the current study identifies that higher-order language processes may be influenced by AIMS-assessed teacher quality.

It is also noteworthy that listening comprehension was differentially affected by the classroom context, since this has been identified as a skill deficit for children with attention problems [61]. Although the design did not specify the specific teaching strategies that were optimally effective, it is consistent with research that emphasizes the beneficial impact of high quality classrooms on children at risk [20,46]. As there is currently modest empirical evidence as to the types of specific classroom contexts and teaching strategies that are effective with students who are at risk of attention difficulties, identifying the observable characteristics of teachers

who exhibit a “goodness of fit” for children can thus help to elucidate best practices for these students. This is the major additional contribution of these findings from the present paper.

Limitations and implications for further study

Some limitations associated with the present study should be noted when interpreting the results. First, the present study is limited by a modest sample size at the classroom level ($n=18$), which, although consistent with other published studies in the literature [25,47], may have resulted in limited power to detect differences using HLM. As is the norm in this age group in this region all the teachers were female. The results are however broadly consistent with two larger studies that assessed $n=31$ and $n=36$ classrooms respectively with the AIMS tool (current authors, papers in preparation). Effect sizes for some main effects reported here, most notably for the attention at-risk children, were large in size as measured on standard metrics by eta squared, suggesting they are likely to be replicable. In the most general sense, our results suggesting the centrality of student engagement and strong classroom management in effective teaching for typical and atypical learners themselves replicate similar findings reported over several decades of research [62-64]. The present study nevertheless places such findings on a more secure empirical basis by using hierarchical modeling of the shared classroom-level variance across diverse regular classrooms and also extends this knowledge base to show exactly for which students these elements of effective teaching fit best, albeit in this particular sample of children and teachers.

A second limitation with the present study is the measures included to assess children and the learning they experienced. The research was carried out with the original version of the Classroom AIMS Instrument, which has since been modified [52]. As the AIMS continues to undergo refinement and validation, further research is necessary to establish the predictive validity of the AIMS in relation to student achievement. Following from the current results, the AIMS appears to be a better predictor of growth in reading comprehension, as opposed to word-level reading skills in grade 1. As such, the AIMS may benefit from greater precision regarding teaching strategies that are important for specific grade levels, since the current version is intended as an evaluation tool for teachers from K-12. Given that literacy teachers may require differentiated teaching strategies and specific skill sets to help children be successful at each grade level [27,65], the broad nature of the AIMS may not adequately emphasize key skills that are associated with literacy teaching in grade 1 specifically. In particular, the relative lack of emphasis on phonics instruction in the AIMS observation tool could explain the fact that the AIMS did not significantly predict students' scores on word-reading. These skills may be more associated with specific implementation strategies and with absolute amount of instruction [26]. Other studies have also observed weak classroom effects at the word level in grade 1 from measures of classroom quality [47,66,67].

Beyond the study use of the Connors questionnaire to assess at-risk status for ADHD there was not further screening for the presence of ADHD. Similarly, wider assessment of text processing and visual-attention and working memory processing was not undertaken in this study. Future studies should seek to undertake more detailed analyses of both the clinical and wider cognitive processing characteristics of samples of children with attention difficulties in relation to the classroom teaching experiences they receive.

A third potential limitation of this study was that the timing of observations, which were all conducted in the winter term, could also have affected the types of teaching practices that were observed. For example, Juel and Minden-Cupp (2000) observed that first grade teachers tend to change the focus of their instruction throughout the year, with an initial focus on phonics in the fall shifting to an emphasis on vocabulary instruction and

text discussions as children's reading skills improve. Observing teaching practices at the beginning of the year, when classroom rules and routines are being established [53,68] could have been particularly salient in elucidating the contribution of distinct classroom management approaches to students' development of self-regulation. Additionally, teachers' ability to change their instruction over time in response to students' developing reading skills, has also been associated with stronger reading outcomes, and could be assessed in future studies [8].

The main implication for further study is that student by classroom interactions for reading comprehension and for attention difficulties demonstrated here suggest that observable aspects of classroom teaching affect reading and attention development for different children in distinct ways, consistent with current goodness-of-fit models. For children with relatively strong reading skills at the beginning of the year, strong classroom management was most effective in supporting literacy development. For those students with weaker literacy skills in the fall of year 1, observed student engagement best predicted growth. These complex interactions thus highlight the need for a richer picture of the effects of pedagogical practices on attention, language and literacy attainment, taking into account the dynamic and multifaceted nature of classroom teaching in future replications and extensions of these findings.

Practical Implications

The practical implications of the present study are numerous. Firstly, the present study confirms that there are important classroom-level differences in growth in literacy and related language skills in Grade 1 among unselected samples of 'typical' teachers. These effects are not restricted to the 'exceptional' teachers used in some previous studies. Secondly, findings further suggest that at least some of these differences can be assessed using research-based observation tools such as AIMS. Potentially such tools can be used by teachers and other professionals to aid in the development of superior teaching through mentoring or other supports [1,54], and may thus aid school improvement or better school board response-to-intervention initiatives. Thirdly and as already noted above, more specifically, our results provide a clue to the differentiated support that probably characterizes the 'goodness-of-fit' of teaching to children with relatively lower and higher initial literacy levels. Our results suggest that observed engagement and enthusiasm for learning mark effective teaching for the former children and teacher development of self-regulated behavior mark effective teaching for the latter more literate grade 1 children respectively. One size does not thus fit all in teaching effectiveness. Fourthly, all AIMS constructs are implicated in the large effects that classrooms have for the language development of children identified by teachers as at-risk of attention difficulties, and potentially in building the resilience of these children. Together these results testify the developmental implications of effective teaching for diverse learners in regular grade 1 classrooms.

References

1. Roehrig AD, Pressley M, Dolezal SE, Mohan L, Bohn CM (2001) Classroom AIMS Instrument: A checklist of effective classroom practices (atmosphere, instruction/content, management) and student engagement for the early primary grades (K-4).
2. Hinshaw SP (1992) Externalizing behavior problems and academic underachievement in childhood and adolescence: causal relationships and underlying mechanisms. *Psychol Bull* 111: 127-155.
3. Lonigan CJ, Bloomfield BG, Anthony JL, Bacon KD, Phillips BM, et al. (1999) Relations among emergent literacy skills, behavior problems, and social competence in preschool children from low- and middle-income backgrounds. *Topics in Early Childhood Special Education* 19: 40-53.
4. McGee R, Prior M, Williams S, Smart D, Sanson A (2002) The long-term significance of teacher-rated hyperactivity and reading ability in childhood: findings from two

- longitudinal studies. *J Child Psychol Psychiatry* 43: 1004-1017.
5. Rabiner D, Coie JD (2000) Early attention problems and children's reading achievement: A longitudinal investigation. The Conduct Problems Prevention Research Group. *J Am Acad Child Adolesc Psychiatry* 39: 859-867.
 6. Wharton-McDonald R, Pressley M, Hampston JM (1998) Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal* 99: 101-128.
 7. Rutter M (1983) School effects on pupil progress: Research findings and policy implications. *Child Development* 54: 1-29.
 8. Connor CM, Piasta SB, Fishman B, Glasney S, Schatschneider C, et al. (2009) Individualizing student instruction precisely: effects of Child x Instruction interactions on first graders' literacy development. *Child Dev* 80: 77-100.
 9. Downer JT, Pianta RC (2006) Academic and cognitive functioning in first grade: Associations with earlier home and childcare predictors and with concurrent home and classroom experiences. *School Psychology Review* 35: 11-30.
 10. Byrne B, Coventry WL, Olson RK, Wadsworth SJ, Samuelsson S, et al. (2010) "Teacher Effects" in Early Literacy Development: Evidence from a Study of Twins. *J Educ Psychol* 102: 32-42.
 11. Rutter M, Maughan B (2002) School effectiveness findings 1979-2002. *Journal of School Psychology* 40: 451-475.
 12. Hall K, Harding A (2003) A systematic review of effective literacy teaching in the 4 to 14 age range of mainstream schooling. *EPPI* 1-105.
 13. Pressley M, Wharton-McDonald R, Allington R, Block CC, Morrow L, et al. (2001) A study of effective first-grade literacy instruction. *Scientific Studies of Reading* 5: 35-58.
 14. Taylor BM, Pearson PD, Clark K, Walpole S (2000) Effective schools and accomplished teachers: Lessons about primary-grade reading instruction in low-income schools. *The Elementary School Journal* 101: 121-165.
 15. National Institute of Child Health and Human Development 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.
 16. Pianta RC, La Paro KM, Payne C, Cox MJ, 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.
 17. Connor CM, Son S, Hindman AH, Morrison FJ (2005) Teacher qualifications, classroom practices, family characteristics, and preschool experience: Complex effects on first graders' vocabulary and reading outcomes. *Journal of School Psychology* 43: 343-375.
 18. Stuhlman MW, Pianta RC (2009) Profiles of educational quality in first grade. *CASTL Research Brief* 109: 323-342.
 19. Burchinal M, Howes C, Pianta R, Bryant D, Early D, 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.
 20. Wilson HK, Pianta RC, Stuhlman M (2007) Typical classroom experiences in first grade: The role of classroom climate and functional risk in the development of social competencies. *The Elementary School Journal* 108: 81-96.
 21. Mashburn AJ, Pianta RC, Hamre BK, Downer JT, Barbarin OA, et al. (2008) Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Dev* 79: 732-749.
 22. Allhusen V, Belsky J, Booth-LaForce CL, Bradley R, Brownwell CA, et al. (2004) Does class size in first grade relate to children's academic and social performance or observed classroom processes? *Dev Psychol* 40: 651-664.
 23. Atkinson IM, Robinson JA, Shute RH (1997) Between a rock and a hard place: An Australian perspective on education of children with ADHD. *Educational and Child Psychology* 14: 21-30.
 24. Juel C, Minden-Cupp C (2000) Learning to read words: Linguistic units and instructional strategies. *Reading Research Quarterly* 35: 458-488.
 25. Al Otaiba S, Connor C, Lane H, Kosanovich ML, Schatschneider C, et al. (2008) Reading First kindergarten classroom instruction and students' growth in phonological awareness and letter naming-decoding fluency. *J Sch Psychol* 46: 281-314.
 26. Connor CM, Morrison, FJ, Katch LE (2004) Beyond the reading wars: Exploring the effect of child-instruction interactions on growth in early reading. *Scientific Studies of Reading* 8: 305-336.
 27. Connor CM, Morrison FJ, Underwood PS (2007) A second chance in second grade: The independent and cumulative impact of first- and second-grade reading instruction and students' letter-word reading skill growth. *Scientific Studies of Reading* 11: 199-233.
 28. American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders. (4th edn.), American Psychiatric Publishing, Inc, Arlington, USA.
 29. McGee R, Share DL (1988) Attention deficit disorder-hyperactivity and academic failure: which comes first and what should be treated? *J Am Acad Child Adolesc Psychiatry* 27: 318-325.
 30. Hudziak JJ, Derks EM, Althoff RR, Rettew DC, Boomsma DI (2005) The genetic and environmental contributions to attention deficit hyperactivity disorder as measured by the Conners' Rating Scales-Revised. *Am J Psychiatry* 162: 1614-1620.
 31. Sonuga-Barke EJ (2005) Causal models of attention-deficit/hyperactivity disorder: from common simple deficits to multiple developmental pathways. *Biol Psychiatry* 57: 1231-1238.
 32. Thapar A, Langley K, Asherson P, Gill M (2007) Gene-environment interplay in attention-deficit hyperactivity disorder and the importance of a developmental perspective. *Br J Psychiatry* 190: 1-3.
 33. Deault LC (2010) A systematic review of parenting in relation to the development of comorbidities and functional impairments in children with attention-deficit/hyperactivity disorder (ADHD). *Child Psychiatry Hum Dev* 41: 168-192.
 34. Johnston C, Mash EJ (2001) Families of children with attention-deficit/hyperactivity disorder: review and recommendations for future research. *Clin Child Fam Psychol Rev* 4: 183-207.
 35. Savage R, Deault L (2010) Understanding and supporting children experiencing dyslexia and ADHD: The challenge of constructing a unifying model of constitutional and classroom influences. In: Littleton K, Wood C, Staarman JK (Eds.), *International Handbook of Psychology in Education*, Emerald Group Publishing Limited London, UK, pp 569-608.
 36. Zentall SS (2005) Theory- and evidence-based strategies for children with attentional problems. *Psychology in the Schools* 42: 821-836.
 37. Muyskens P, Ysseldyke J (1998) Student academic responding time as a function of classroom ecology and time of day. *The Journal of Special Education* 31: 411-424.
 38. Rimm-Kaufman SE, La Paro KM, Downer JT, Pianta RC (2005) The contribution of classroom setting and quality of instruction to children's behavior in kindergarten classrooms. *The Elementary School Journal* 105: 377-394.
 39. Bugental DB, Lyon JE, Lin EK, McGrath EP, Bimbela A (1999) Children "tune out" in response to the ambiguous communication style of powerless adults. *Child Dev* 70: 214-230.
 40. Barkley RA (2002) Psychosocial treatments for attention-deficit/hyperactivity disorder in children. *J Clin Psychiatry* 63 Suppl 12: 36-43.
 41. Martinussen RL, Tannock R, Chaban P, McInnes A, Ferguson B (2005) Increasing awareness and understanding of attention deficit hyperactivity disorder (ADHD) in education to promote better academic outcomes for students with ADHD. *Exceptionality Education Canada* 16: 107-128.
 42. Raggi VL, Chronis AM (2006) Interventions to address the academic impairment of children and adolescents with ADHD. *Clin Child Fam Psychol Rev* 9: 85-111.
 43. Trout AL, Lienemann TO, Reid R, Epstein MH (2007) A review of non-medication interventions to improve the academic performance of children and youth with ADHD. *Remedial and Special Education* 28: 207-226.
 44. Tannock R, Martinussen R (2007) Promising practices in education for students with ADHD. *Orbit* 37: 32-35.
 45. Bub KL (2009) Testing the effects of classroom supports on children's social and behavioral skills at key transition points using latent growth modeling. *Applied Developmental Science* 13: 130-148.
 46. Hamre BK, Pianta RC (2001) Early teacher-child relationships and the trajectory of children's school outcomes through eighth grade. *Child Dev* 72: 625-638.
 47. Perry KE, Donohue KM, Weinstein RS (2007) Teaching practices and the promotion of achievement and adjustment in first grade. *Journal of School Psychology* 45: 269-292.

48. National Institute of Child Health and Human Development Early Child Care Research Network (2005) Predicting individual differences in attention, memory, and planning in first graders from experiences at home, childcare, and school. *Dev Psychol* 41: 99-114.
49. Rudasill KM, Gallagher KC, White JM (2010) Temperamental attention and activity, classroom emotional support, and academic achievement in third grade. *J Sch Psychol* 48: 113-134.
50. Williams KT (2001) Group Reading Assessment and Diagnostic Evaluation: Teacher's scoring and interpretive manual. American Guidance Service, USA.
51. Conners CK (1997) Conners' Rating Scales-Revised: Technical Manual. Multi-Health Systems Inc., Toronto, USA.
52. Roehrig AD, Christesen E (2010) Development and use of a tool for evaluating teacher effectiveness in grades K-12. In: Shute VJ, Becker BJ (Eds.), *Innovative Assessment for the 21st Century*. Springer, New York, NY, USA, pp 207-228.
53. Bohn CM, Roehrig AD, 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: 269-287.
54. Roehrig AD, Bohn CA, Turner JE, Pressley M (2008) Mentoring beginning primary teachers for exemplary teaching practices. *Teaching and Teacher Education* 24: 684-702.
55. Tabachnick BG, Fidell LS (2006) *Using multivariate statistics* (5th edn.), Allyn & Bacon, Boston, USA.
56. Bogner K, Raphael LM, Pressley M (2002) How Grade 1 teachers motivate literate activity by their students. *Scientific Studies of Reading* 6: 135-165.
57. Kewley G (2005) *Attention deficit hyperactivity disorder: What can teachers do?* (2nd edn.), David Futon Publishers, London, UK.
58. Morrison FJ, Connor CM (2002) Understanding schooling effects on early literacy: A working research strategy. *Journal of School Psychology* 40: 493-500.
59. Savage RS, Abrami P, Piquette-Tomei N, Wood E, Sanghera-Sidhu S, et al. (2013) A (Pan-Canadian) cluster randomized control effectiveness trial of the ABRACADABRA web-based literacy program. *Journal of Educational Psychology* In Press.
60. Kirby J, Savage RS (2008) Can the simple view deal with the complexities of reading? *Literacy* 42: 75-82.
61. McInnes A, Humphries T, Hogg-Johnson S, Tannock R (2003) Listening comprehension and working memory are impaired in attention-deficit hyperactivity disorder irrespective of language impairment. *J Abnorm Child Psychol* 31: 427-443.
62. Doyle W (1977) Learning the classroom environment: An ecological analysis. *Journal of Teacher Education* 28: 51-56.
63. Moos RH (1973) Conceptualizations of human environments. *American Psychologist* 28: 652-665.
64. McKinney JD, Osborne SS, Schulte AC (1993) Academic consequences of learning disability: Longitudinal prediction of outcomes at 11 years of age. *Learning disabilities Research & Practice* 8: 19-27.
65. La Paro KM, Rimm-Kaufman SE, Pianta RC (2006) Kindergarten to 1st grade: Classroom characteristics and the stability and change of children's classroom experiences. *Journal of Research in Childhood Education* 21: 189-202.
66. Foorman BR, Schatschneider C, Eakin MN, Fletcher JM, Moats LC, et al. (2006) The impact of instructional practices in grades 1 and 2 on reading and spelling achievement in high poverty schools. *Contemporary Educational Psychology* 31: 1-29.
67. Pianta RC, Stuhlman MW (2004) Teacher-child relationships and children's success in the first years of school. *School Psychology Review* 33: 444-458.
68. Cameron CE, Connor CM, Morrison FJ, Jewkes AM (2008) Effects of classroom organization on letter-word reading in first grade. *J Sch Psychol* 46: 173-192.