Dissertations   Theses and Dissertations

2014

Relating Preschool Class Size to Classroom Life and Student Achievement

Jessica Francis
Loyola University Chicago, jfrancis@nieer.org

LOYOLA UNIVERSITY CHICAGO

RELATING PRESCHOOL CLASS SIZE TO CLASSROOM LIFE AND STUDENT

ACHIEVEMENT

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE GRADUATE SCHOOL

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

PROGRAM IN CHILD DEVELOPMENT

BY

JESSICA FRANCIS

CHICAGO, IL

MAY 2014

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

There is an ongoing debate regarding the importance of smaller classes in elementary school, and a lack of solid research to support class size policies in preschool. State spending on preschool has nearly doubled in the last five years and currently more than 80 percent of American 4-year olds go to some kind of preschool. Increasing enrollments in preschool coupled with the high costs of reducing the size of classes creates the need to decide how many children should be placed in a preschool classroom.

The majority of states require that programs implement class sizes of 20 and teacher-child ratios of 1:10 (Barnett et al, 2011), but much of the research on class size suggests that class sizes smaller than 20 might be more beneficial for children. This dissertation examines the effects of preschool class size on classroom life and student achievement by drawing upon data from 21 teachers and 354 children that were collected during the 2008-2009 school year. Regular class sizes contained 20 students and reduced class sizes contained 15 students. Either the AM or PM session was randomly assigned to be 15 students for each teacher, so that each teacher taught both a regular and reduced class size.

Children who attended reduced size classrooms were found to partake in more one-to-one interactions with teachers than children in regular size classrooms, but there were no differences between groups in the quality of classroom interactions as measured by the Classroom Assessment Scoring System (CLASS). Children in smaller classrooms also were found to gain more in literacy skills by the end of preschool. In contrast, there were no significant differences between groups in vocabulary or math gains. These

results indicate that an assigned difference of five children in a preschool classroom can benefit children's cognitive development after just one school year, but these benefits are not explained by changes in the most commonly used measures of classroom quality, which were minimal. Future endeavors to reduce class size in preschool might be enhanced if coupled with professional development strategies that aim to maximize teachers' effectiveness with smaller classes.

CHAPTER 1

INTRODUCTION

This dissertation examines the effects of preschool class size on classroom life and student achievement. The data for this dissertation comes from a large 3-year examination conducted by the National Institute for Early Education Research (NIEER) at Rutgers University and the Chicago Public Schools. The purpose of the larger study was to understand the effects of three preschool inputs (class size, parental engagement, and a professional development initiative) on teachers' classroom practices and children's outcomes over time, ultimately for the sake of adopting policies that benefit children's learning and development.

Results from the larger study showed evidence that the parental engagement component was found to significantly increase children's vocabulary and early literacy skills after one year of preschool. Reduced class size was also found to significantly increase children's early literacy skills. In contrast, there was no evidence that the professional development intervention led to increased learning and development on the measures employed. These results were based on data that was pooled across multiple years of the study, and various child characteristics were controlled for in each statistical model. Each input was also analyzed in terms of overall cost-effectiveness in relation to the gains that were found for each measure that was studied. In general, the parental engagement program was found to be most cost-effective for widespread adoption if child outcome gains are of primary concern.

It is important to note that all three interventions were studied simultaneously after the first year of the evaluation. This allowed for various combinations of reduced class size, parental engagement and professional development inputs that were received by children. For instance, some classrooms had a reduced class size and received the parental engagement input; others received parental engagement coupled with professional development, and so forth. As a result, it is difficult to disentangle the effect of each input after the first year of the study because of the various combinations of inputs that existed across classrooms.

This dissertation focuses specifically on class size, which was the only variable in the first year of the study. It draws upon classroom observation data from 21 teachers and assessment data from 354 children that was collected during the 2008-2009 school year. Three research questions guide the present study:

(1) Does preschool classroom quality vary by class size?

(2) Does the quantity of time children spend in particular types of preschool activities and interactions vary by class size?

(3) Do preschool children's cognitive outcomes vary by class size?

Preschool is an important developmental period for children as they learn and acquire skills that will eventually aid them in school. It is also a time that marks great variability in the experiences children have because some children stay at home with their families, others attend private child care programs, and others attend public preschool programs such as state programs or Head Start. While children's development has been a topic of interest for centuries, children were not viewed as having specific age-defined

needs until the late 1800s, and it was not until the 1960s that researchers in the US began to intensively study the role of early schooling (Condry, 1983; Cahan, 1989).

This interest in early childhood education coincided with increased numbers of women in the labor force in the late 1960s, and therefore the expanded interest in children spending time in formal day care settings outside of their homes. Also in the early 1960s, the well-known High/Scope Perry Preschool study began, which examined the lives of 123 children born in poverty and at high risk of failing in school (Schweinhart et al, 2005).

From 1962–1967, at ages 3 and 4, the High/Scope Perry Preschool study subjects were randomly divided into a program group that received a high-quality preschool program based on High Scope's participatory learning approach and a comparison group who received no preschool program. Gains were found for those children that attended preschool in comparison to their no-preschool peers, specifically in terms of their cognitive development. Impacts over time were also pronounced, affecting things like employment rates, familial relationships, and even crime (Schweinhart et al, 2005). This resulted in major economic benefits for society, with cost-benefit estimates being as large as 1:7 for every dollar spent on high quality preschool. More recent estimates have been as large as 16 dollars in benefits for every dollar spent (Belfield et al., 2006).

Head Start, the federally-funded preschool program for low-income families was also launched around this time, which further transformed the preschool landscape. Head Start was also initially found to have large impacts on children's IQ scores, though it was controversial from the start, and critics called for a large-scale evaluation to prove Head

Start really worked. In the spring of 1969, soon after the inauguration of President Richard Nixon, a report about Head Start was released claiming that IQ gains from Head Start "faded out" after just a few years (Westinghouse Learning Corporation and Ohio University, 1969). However, longer-term effects proved to be much more promising (Oden, Schweinhart, Weikert, Marcus, and Xie, 2000). A number of additional early education studies were also conducted around this time, many of which also provided evidence for the benefits of formal preschool education on children's cognitive development (Pierson and Sperber, 1974; Ruopp et al, 1979; Cawley & Goodstein, 1966; Zigler, 1983; Reynolds, 2000; Campbell et al, 2002).

Over time, this body of literature has grown and demonstrates general consensus that high quality preschool programs for children at 3 and 4-years of age positively impact their school readiness and educational attainment (Gormley, Gayer, Phillips, & Dawson, 2005; Magnuson Meyers, Ruhm, & Waldfogel, 2004; Schweinhart et al., 2005; Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Heckman, Moon, Pinto,Savelyev, & Yavitz, 2010; Reynolds, Temple, White, Ou, & Robertson, 2011; Rolnick & Grunewald, 2003, Pianta, Barnett, Burchinal, & Thornburg, 2009), and especially so for low-income students (Camilli, Vargas, Ryan, & Barnett, 2010). High quality preschool programs have come to be characterized by a number of components, or standards. These include comprehensive early learning standards, teacher education and training, assistant teacher education, class size and ratio, teacher in-service, meals, monitoring, and referral and support services, among others. The NIEER state yearbook that is released annually depicts quality in terms of these benchmark standards, and it

ranks states based on the number of benchmarks that their public programs achieve (Barnett et al, 2011). While various combinations of these standards have been shown to influence quality (Ackerman and Barnett, 2006), there are no rigorous studies that examine the effects of these individual program components on classroom practices and children's development in preschool.

The discourse surrounding preschool quality has shifted over time to one characterized by a process-oriented focus, emphasizing that which occurs in the classroom. This marks a clear divergence from the proxies of quality mentioned above that were so heavily attended to in the past. Even so, policy regulated features such as class size cannot be ignored when developing a preschool program.

This particular variable has large implications for the interactions that occur between teachers and children, and thus the skills that children develop. Like any feature, class size is nested within larger influential spheres such as neighborhoods, schools, and families, all of which contribute to children's experiences and development. Further, class size is a product of political, developmental, and financial constraints and it cannot be understood as an isolated policy. There are developmental necessities that make extremely large class sizes in preschool inappropriate and even dangerous. At the same time, the high costs involved and conflicting political stances influence decisions that are made regarding class size. The classroom approach that is used further impacts class size decisions. In other words, some curriculums are set up to accommodate larger classes, while others work better with smaller classes. Even so, this particular feature varies from

program to program and from state to state, and it is often unclear and seemingly arbitrary why one class size is chosen over another.

Class size has been studied extensively, yet a great deal remains to be learned, and especially so in the preschool period. Given the increase in preschool enrollment over the past several decades and the increased attention to children's achievement at an early age, every standard that lends itself to a child's experience must be scrutinized. Class size is an aspect of the classroom experience that appears to be straightforward, yet it actually encompasses a great deal of ambiguity and unanswered questions. For example, is there a particular class size that dramatically alters children's learning and development? Does learning thrive more and more the smaller the class size, or is there a point at which further reductions make minimal differences for children? Are small class sizes equally important across different schools and grade levels? Further, is it more important to provide for a small ratio of teachers to children, or is a reduction in the overall number of children in the classroom more significant for children's learning? Indeed, class size is not the same as ratio, and looking at the number of children along with the number of staff in the classroom provides a different representation than looking at the number of children alone.

Adding to the complexity involved in studying class size, each teacher also has a unique background, characteristics, and teaching style that influence children's experiences. Do the effects of class size change depending upon those particular teacher variables? Are we able to tease out the effects of class size alone? And how does class size actually effect what happens in the classroom?

While the current study will focus solely on class size in preschool, the literature review that follows will assess research that involves children in the early elementary grades, as well. It is important to get a handle on the research that has been done with older children, as it provides the impetus for experiments with younger children. Further, the bulk of research in this area examines the effects of class size in elementary school and there are a limited number of experiments of class size in preschool.

The two major forms of research on class size consist of that related to classroom life and that related to children's outcomes and school achievement. There is a larger body of research on class size and children's school achievement, which is comprised of experiments that examine whether smaller classes result in greater performance among children, in terms of assessments or standardized test scores. Literature that focuses more on class size as it relates to learning time or classroom life, which can be viewed as intermediate factors to children's achievement, is less prominent. These studies address questions such as: How do smaller classes influence interactions between teachers and children, and among children? Do smaller classes result in more on-task teaching and less disengagement among children? Results from the latter type of research can help us understand why differences in class sizes may have an effect on children's achievement. They can help us to define the mechanisms through which class size influences student learning and achievement.

This dissertation will begin with a review of research on class size as it relates to classroom life, followed by studies on class size and children's achievement – both in the early grades. It will then transition into a review of the limited research that exists on

class size in early childhood, beginning with a focus on classroom life and concluding with class size as it relates to children's achievement in preschool. Some of the early childhood research that will be discussed comes from the child care literature, as its history parallels that of public programs.

The most rigorous and well-known study of class size is the Tennessee STAR experiment, which was a large-scale experimental study. The study took place in the 1980s and showed remarkable gains for children that attended smaller classes in comparison to their regular class size peers, providing the impetus for a number of additional studies on class size that transpired into the 1990s. With consistently limited funds available for education and the complexity involved in studying class size, it has been frequently revisited as a topic of research interest. Over time, studies have concluded that smaller class sizes are a key indicator of classroom quality, leading to better experiences and outcomes for children. This is based on the rationale that smaller groups allow for more individual attention and better supervision of children. Intuitively, small class sizes allow teachers to devote more time to each child, thus allowing for longer conversations and more individualized interactions and teaching time. However, there is also a body of literature that largely contests the notion that small classes positively impact children's learning and educational experiences.

There are valid arguments for and against class size reduction, based on research that favors both sides. Because of the intricacy involved in studying class size, it is not surprising that research in this area has taken various forms, asking various questions, and has produced conflicting results. This is largely due to the inconsistency across

studies in defining small versus large class sizes. For instance, some studies define a small class size as 25 or few children, while others define a large class size as ranging from 22 to 26 children. The various research designs and types of analyses employed also contribute to mixed findings.

Class size is also very expensive in relation to other features that have been thought to influence student learning, such as professional development or parental involvement, which is perhaps another reason why it has been studied so extensively. In the 2010-11 school year the average US pupil/teacher ratio in public schools was 15.3/1 and the average teacher salary was approximately $55,000. That translates to an individual cost of $3,600 per pupil in teacher salary alone. With about 49.3 million public school students enrolled, if we were to decrease the present average class size by one student across the board it would cost over $12 billion extra per year in aggregate for the U.S (NCES, 2012).

In the example just given of a one-student reduction in class size across the U.S., only teacher salaries are taken into account, and the costs of reducing class sizes involve much more than teacher salaries. For instance, more classrooms are also needed. Again, for a one-student reduction in class size across the U.S., more than 225,000 additional classrooms would need to be added to the nation's stock. The high cost involved in reducing class sizes warrants widespread research on the feature regardless of how much it is thought to influence student learning, particularly because the majority of states have mandated some type of class size reduction policies in the last decades, mainly targeting students in elementary grades (especially K-3) (Zinth, 2009).

Given the high cost of reducing class sizes, the limited pool of funds available for education, and the mixed findings and unclear implications for specifying class sizes, there is certainly a need for more intensive study of this topic. We often claim that smaller is better, but just how small is small enough? There is a need for more specific information on the effectiveness of different class sizes. This is especially the case for the preschool period as very few rigorous experiments have been conducted for children of this age. Further, more and more children are attending public preschool programs and investments in these programs continue to rise. State spending on preschool has nearly doubled in the last five years and currently more than 80 percent of American 4-year olds go to some kind of preschool. The majority of states require that programs implement class sizes of 20 and teacher-child ratios of 1:10 (Barnett et al, 2011), but as the literature review in Chapter 2 will show, much of the research on class size suggests that class sizes smaller than 20 might be necessary for enhancing classroom practices and increasing children's achievement (Word et al, 1990; NICHD, 1999, 2002). Further, even the smallest class size reductions can have very large cost implications.

The next chapter (Chapter 2) will provide an extensive review of the different ways that class size has been studied in both preschool and the early elementary grades. This review highlights the ambiguity in the way class size has been defined and studied, the mixed findings that have accrued, and the lack of solid evidence that exists for the preschool period. As is made evident in the review of literature, there are a multitude of variables and intersecting influences that surround the examination of class size. In other words, the influence of class size on an individual's development is nested within

interconnected systems. For instance, children influence classroom practices which influence family systems, and vice versa. This can be explained through Bronfrenbrenner's Ecological Systems Theory, which is the overarching framework for the current study, to be presented at the end of Chapter 2. Within the context of this theoretical framework, the evidence presented in the literature review provides the foundation for the research method that will be used (Chapter 3). Results will be presented in Chapter 4, followed by a discussion of the results as well as conclusions, limitations and implications in Chapter 5.

CHAPTER 2

REVIEW OF LITERATURE

The review of research that follows is separated into four sections: class size and classroom life in the early grades; class size and children's achievement in the early grades; class size and classroom life in preschool; and class size and children's achievement in preschool. While the primary interest for the sake of the proposed study is to present an understanding of how group size influences preschool classrooms and children, it is important to get a handle on the research that has been done with older children, as it provides the impetus for experiments with younger children. Further, as will be realized, the bulk of research in this area examines the effects of class size in elementary school and there are a limited number of experiments of class size in preschool.

Class Size and Children's Classroom Experiences in the Early Grades

A number of studies have found relationships between small classes and positive classroom experiences for students. For instance, one of the most rigorous studies of class size to-date, the Tennessee demonstration project (Project STAR; Student-Teacher Achievement Ratio) compared classes of 13 to 17 to classes of 22 to 25. Children in kindergarten from a large statewide sample were randomly assigned to smaller (13 to 17) or larger (22 to 25) classes, and some of these larger classes had a full-time paid aide, resulting in a third comparison group. STAR data has been reviewed and analyzed by

numerous authors in many different ways and details of the study design and implementation will be reviewed later in this paper.

Class Size and Children's Classroom Behavior

Everston and Folger (1989) investigated class size in relation to teaching practices in the STAR sample. Trained observers recorded student initiation, on- and off-task behaviors, and waiting time separately during math and reading lessons in 52 second grade classrooms. Average percentages for small and regular class sizes were computed separately for reading and mathematics. Results showed that children in smaller classes took greater initiative in the classroom and were more frequently on-task. However, this was not the case during both math and reading lessons. In addition, it was found that teachers in small classes spent more time on instruction and less on managerial and organizational tasks.

The latter finding mentioned above was replicated in a study of pre- and post-observations of small (averaging 14 students) and regular size (averaging 23 students) first grade classrooms in two Title 1-eligible schools in North Carolina. Interactions between teachers and students were coded by trained observers every 4-5 seconds, using the following categories: "personal" (not related to academic activities or school), "institutional" (related to daily classroom routines), or "task" (related to academic activities). They were also classified as having an either an "individual", "group", or "mixed" focus. Small classes were found to have more "task events" and fewer "personal" and "institutional" practices. There was also more individual attention given to students and a greater degree of individualized communication among teachers and

students in small classes. Despite the differences in percentages, no tests of statistical significance were performed so it is difficult to rely heavily on these findings (Achilles et al., 1995).

Natural Variation in Class Size

A study conducted by the NICHD Early Child Care Research Network (2004) examined the extent to which natural variation in observed class size was related to classroom processes, and found generally positive associations with small class sizes. The study sampled 651 children in 651 different first grade classrooms and used data from classroom observations to understand the differential impacts of class size, which was collected from teacher reports as well as during observations of classrooms. This sample was part of a large-scale longitudinal study that recruited mothers in hospitals at the time of giving birth, and children were followed over time. The details regarding recruitment and inclusion for this sample will be described in a later section of this paper (see *Class Size and Children's Preschool Experiences*, p. 23) Each classroom was observed by a trained, reliable observer for 3-hours during the morning. The focus was on one specific child's behavior, the teacher's behavior toward that child, the setting in which that child was working, and the overall classroom environment. Dependent variables included the amount of time spent in various classroom processes (i.e. activities, engagement, interactions) and ratings of classroom environment quality.

A number of significant findings emerged, some of which positively favored smaller classes and some that attributed greater benefits to larger classes. Findings differed between teacher report and observations of class sizes. For instance, in teacher-

reported class sizes of fewer than 23 children, there was significantly more negative and disruptive behavior in classrooms. This finding was in the same direction, but not statistically significant for observed teacher-child ratios. Overall, findings showed that when more children (above 25) were in the classroom there were more interactions with the teacher, though the quality of instructional support decreased.

The authors also analyzed the data in a different way by breaking their sample into the Tennessee class size groupings (22-27 children versus 13-17 children). The only significant finding between these group sizes was for teacher-reported class size and ratings of externalizing behavior. Specifically, teachers of children in smaller classes rated children as showing more externalizing behavior, which was not what was expected. In general, findings indicated that small classes can lead to positive experiences for children in some ways, but not in other ways.

Class Size and Student Individualization

The Wisconsin SAGE Project also showed evidence of positive classroom experiences attributable to small class sizes. The SAGE Project adopted a program in the 1996-97 school year that reduced student-teacher ratios to 15 students per adult in kindergarten and first grade (second and third grade reduced classes were added in the subsequent two years). This allowed for various configurations such as 15:1, 30:2, and 45:3. The relationship of classroom life to reduced class size was examined through various data sources, such as logs, classroom observations, teacher questionnaires, and teacher interviews. The latter two focused on teachers' perceptions of classroom changes related to reduced class size.

Findings from teacher questionnaires (n=150) and in-depth teacher interviews (n=28) revealed minimal differences between the various configurations of teacher-child ratios. In general, all teachers indicated that their teaching had changed as a result of having a small-sized class, most prominently in the area of individualization. Reduced class sizes allowed for some movement toward more student-centered teaching, yet many techniques and methods that were used continued to be the same as those used in regular-sized classrooms. These techniques, however, were more frequently targeted at individual children in smaller classes. This greater focus on students as individuals arose from teachers' reports of having greater knowledge of their students, more instructional time, and more pleasure and satisfaction regarding teaching (Molnar et al., 1999).

In contrast, the degree of individualization practiced by teachers was not evident for smaller classes in a study conducted by Bourke (1986), though other positive associations emerged. The effect of class size on achievement by way of teaching practices was examined with a sample of 63 teachers in 33 elementary schools in the Melbourne (Australia) metropolitan area. In this study, class size was defined as the number of children per adult for each 5 minutes of lesson time observed. Classroom size ranged from 12 to 33, with an average of 25, thus small classes were identified as those with 25 or fewer children.

Data were collected on a wide range of teaching practices and student achievement over a 12-week period. The frequencies of many different teaching practices in terms of the classroom context, the participants in interactions between teachers and students, and the nature of each interaction were recorded by trained observers during

mathematics lessons. Those practices that were found to be significantly related to class size were considered for inclusion in a composite variable formed from the practices. This block variable was then included in a regression analysis, along with class size, student ability, and background factors, as pathways to achievement.

Nine teaching practices that had significant correlations with class size were identified. Specifically in smaller classes, there was a greater use of whole class teaching, more teacher follow-up of questions, greater use of homework, assignments and oral tests for assessment purposes, and more direct interaction. Larger classes had a greater use of class grouping, more student questions (usually seeking help or clarification), more lecturing, more interactions overall between teachers and students, and higher noise levels. Overall, results showed that class size differences were related to teaching practices and, through teaching practices, related to student achievement. Class size alone did not affect student achievement directly and significantly. Rather, teachers with smaller classes tolerated less noise, provided whole class instruction and probed with follow-up questions more frequently, and their classes had higher achievement. In addition, students in smaller classes were given more homework and had higher achievement. As mentioned, the degree to which teachers practiced individualization and engaged students were not found to be of consequence in this study.

Class Size and Teacher Attitudes and Instruction

Also focusing on class size and intermediate factors to achievement, Glass and Smith (1980) conducted a meta-analysis in an attempt to understand the relationship between class size and teacher attitudes and instructional processes. They separated class

size effects from 59 studies into three areas: (1) affective effects on students (i.e. self-concept, interest in school, participation), (2) effects on teachers (workload, morale, attitudes toward students), and (3) effects on classroom environments and processes (attempts to individualize instruction, classroom climate).

Results showed that practices in each of these categories improved as class size decreased, and effects were greatest on teacher attitudes. Further, improvements increased dramatically in each category when class size decreased to a size of 10 or fewer students, and effects were largest for students 12 years and under (compared to separate groups aged 13 to 17, and 18 and older). These findings point to the positive effects of small classes on teachers' attitudes and instruction. However, it is important to note that when data was analyzed separately for randomized versus uncontrolled experiments, findings were more robust for uncontrolled experiments that were included in the meta-analysis. This could mean that the overall class-size effect is inflated and should be interpreted with caution.

Minimal Class Size Effects

In contrast to the studies that were just reviewed, other research has shown minimal differences in classroom processes attributable to class size. For instance, Shapson et al. (1980) examined classroom process variables with the Toronto Classroom Observation Schedule (designed specifically for this study), an observation checklist, a classroom atmosphere rating scale, and a classroom quality rating scale. Combined, the first three measures were used to assess teachers' verbal behavior, pupil participation, pupil aggressive behavior, method of instruction, subject emphasis, use of educational

aids, physical condition of the classroom, and classroom atmosphere (e.g students' regard for the teacher). The quality rating scale assesses classroom activity in terms of individualization, interpersonal regard, creative expression, and group activity. Data was collected from 62 fourth and fifth grade classes (16 each of class sizes 16, 23, 30, 37) by trained external observers over eight half-day visits to each classroom during each year of the study. .

Significant differences between class sizes were detected in the following areas: proportion of students addressed as individuals, lecture by teacher, and supervision by the teacher while students worked, and proportion of written aids used. Within these areas, only the proportion of students addressed as individuals was significantly favored by the smallest class sizes. In contrast, supervision by teachers and proportion of written aids used were used more frequently in class sizes of 30. Thus, while it was hypothesized that the two small class sizes (16 and 23) would result in great changes in classroom processes, there were minimal differences outside of the proportion of students that were addressed as individuals during the classroom day.

Similarly, findings from an evaluation of the California class size reduction (CSR) program in kindergarten through third grade found that classroom instruction in small classes (20 or fewer students) was generally no different from that in larger classes (more than 20 students). Results from surveys that were completed by teachers in both reduced and non-reduced classes showed that teachers in both size classes covered about the same number of mathematics and language arts topics. They also spent about the same amount of time on each major curriculum element and used generally the same teaching

strategies. There were differences favoring reduced size classes in terms of the amount of

time teachers spent with small groups or individual students, but these differences were

small (Stecher et al., 2001).

Pianta et al. (2002) also found mixed relationships between class size and

classroom environments. They studied a subset of 223 children from the NICHD study in

223 kindergarten classrooms, in an attempt to understand relationships between

kindergarten classroom environments and teacher, family, and school characteristics, and

child outcomes. Class size and teacher to child ratios were two of several classroom

characteristics that were examined in relation to global classroom ratings as well as child

and teacher behaviors, activities, and setting conditions. Class size ranged from 5 to 31,

with an average of 19.86. The range of ratios was not specified.

Each classroom was observed for approximately 3 hours by a trained, external

observer. Global ratings on a scale of 1 to 7 were assigned to classrooms in terms of

child-centered climate, instructional climate, and teacher positivity. A classroom

characterized by high-level child-centered climate was one that had low classroom over-

control, low negative emotional climate, and high ratings on classroom management and

support of children's responsibility. A classroom characterized by high-level instructional

climate was one that included high ratings on literacy instruction, evaluative feedback,

and instructional conversation. Teachers' positivity was rated based on their

sensitivity/responsivity, over-control/intrusiveness, and disengagement/detachment.

Time-sampled codes were also assigned to individual "target" children based on

their setting and activities (i.e. structured teacher-directed activity vs. unstructured

classroom activity), teacher behaviors (i.e. reads aloud, interacts with small group), and child on-task behavior. Findings showed that class size was not related to the frequency of any time-sampled codes and there was a considerable range of child experiences within any given classroom. In contrast, a lower teacher to child ratio (fewer adults per child) was related to lower teacher positivity and lower instructional climate.

The research that was just reviewed shows that small classes have been linked to positive classroom experiences for children in some cases and not in other cases. For instance, some research found more direct, or academic-focused interactions in small classes (Glass & Smith, 1980; Achilles et al, 1995; Bourke, 1986), while other research found more interactions occurring in larger class sizes (NICHD, 2004). Similarly, some studies found more individualized interactions in small classes (Molnar et al, 1999; Stecher et al, 2001), while others found no differences in individualization by class size (Bourke, 1986). It can be said that reduced class sizes have the potential to influence classroom life, though this influence is not inevitable and does not occur in all classroom circumstances.

Class Size and Student Achievement

In light of such mixed findings regarding class size and children's classroom experiences, it is not surprising that small classes also show variable impacts on children's achievement, based on the rationale that small classes impact student achievement through the experiences that different class sizes allow for. Indeed, some research shows that small classes positively impact children's achievement levels, while other research claims just the opposite: that small classes have small to no or even

negative impacts on child achievement levels. This body of research will be reviewed next.

Tennessee STAR Findings

While Project STAR touches upon intermediate factors, as depicted in the previous section, its primary aim was to understand class size as it relates to student achievement. This study marks the largest, and only randomized experiment of class size to-date. The Tennessee STAR experiment alone has been the impetus for major class size reduction policies in various states, therefore it is depicted in great length in the following section. Mosteller (1995) summarizes this 3-phase study in depth, as follows. Phase 1 marks the largest, most rigorous, and most studied phase of the study, and this phase is the only one referred to as Project STAR. Phases 2 and 3 were add-on components with more specific research questions.

In 1985, children in Project STAR (phase 1) were randomly assigned to classes of 13 to 17 (small) or 22 to 25 (regular). Regular classes took two different forms, in that some of them had a full-time, paid aide, allowing for smaller teacher to child ratios. Overall, more than 6000 students in 329 classrooms, across 79 schools participated in the study in kindergarten. Children remained in their class types (small, regular, or regular with an aide) through third grade, with some exceptions which will be discussed in detail later in this review. Any student who entered a participating school in any relevant grade was added to the study and randomly assigned to one of the three class types. Participating students who were retained in a grade, skipped a grade, or left their school were removed from the sample. Overall, 2200 new students entered the project in first

grade, 1600 in second grade, and 1200 in third grade, resulting in a total of 11,600

children that were involved in the study over all four years. A new teacher was randomly

assigned to each small class each year. Variations from the ideal research design will be

discussed in greater detail later.

Students in Project STAR were located in inner-city, suburban, urban, and rural

schools. For the sake of study comparisons, inner-city (defined as those in which more

than half of the students received free or reduced-price lunches) and suburban (outlying

areas of inner-cities schools were combined in the category of metropolitan areas. To be

eligible to participate, a school had to be able to accommodate 3 class types (small,

regular, regular with aide) for the duration of 4-years. Of the 180 schools that offered to

participate, 100 met the qualifying criteria and 79 actually participated.

All students returned to regular size classes in fourth grade, at which point follow-

up studies were conducted to examine long-term outcomes. The Lasting Benefits Study

(phase 2) was begun in 1989 to determine whether perceived benefits lasted when

children returned to regular size classes. Under phase 3, Project Challenge, the 17

economically poorest school districts were given small classes in kindergarten, first,

second, and third grades. Results from phase 2 will not be reviewed since older children's

development is not within the scope of this literature review.

Authors have examined this data in numerous ways. The majority of research that

draws upon STAR data favors small classes (e.g. Mosteller, 1995, Word et al, 1990; Finn

& Achilles, 1990), though some authors claim small classes have little importance overall

for children (Hanushek, 1999). Implications for small class sizes vary depending upon the specific question that is asked and type of analyses that are employed.

The most straightforward analyses of STAR data examined the basic impacts of small classes on children's math and reading achievement scores. Finn and Achilles (1990) reported results at the end of kindergarten and at the end of Grade 1, after the second year of the study. Student achievement scores were analyzed cross-sectionally for the entire kindergarten and first grade sample, and longitudinally for the subset of pupils who were in the study for both kindergarten and first grade. Both analyses employed multivariate analyses of variance. Results showed that kindergarten and first grade students in smaller classes scored significantly better on standardized reading and math tests. Those who were in small classes for 2 years showed significantly greater growth during first grade on standardized reading measures when compared with their regular class peers. There were no significant differences between teacher aide and regular classes in the first grade cross-sectional analyses. However, in the 2-year longitudinal sample, teacher-aide classes gained as much as small classes from the end of kindergarten to the end of first grade on standardized reading tests. This finding is difficult to reconcile. Perhaps the consistency of class type over time is beneficial for children's gains, and simply attending a class with low ratios due to an aide's presence impacts children more over time than initial results might show. However, this notion is speculative and would be a valuable topic for future exploration.

The analyses reported during the original project (Word et al., 1990) provided results from cross-sectional analyses of achievement, at the end of each year of

experimentation. Nested ANOVA and MANOVA models showed that students in small classes in grades K through 3 had superior academic achievement in every school subject, compared to students in regular-size classes. Mean differences between children who attended regular classes and those who attended small classes were about the same in kindergarten through third grade. On average, those in small classes scored 0.17 standard deviations above those who attended regular size classes. One major finding was that significant effects were not found for children who attended regular-sized classrooms with a full-time aide compared to those without an aide, which points to the importance of actual class size rather than small teacher-child ratios. However, these effects may be attenuated because of the frequent availability of part-time aides in regular size classes, which were not accounted for (Krueger, 1999).

Another way of summarizing the results gives percentile ranks for the average score based on national norms for the test (Word et al., 1990). Results showed that, on average over the four grades, children in small classes gained about eight percentiles over those in regular size classes. This was slightly more for reading and slightly less for math. These findings are in agreement with those found by other authors (e.g. Nye, Hedges, and Konstantopoulos, 2002, 2000, 2000a), which showed that the average effect of small classes on math and reading scores was significant and positive in almost all grade levels.

Class size effects on different groups of students. Research also examines STAR data in an attempt to understand the impact of class size on different groups of students or in varying circumstances. For instance, Finn and Achilles (1990) found that that not only did achievement reach statistical significant for every test, there were also

significant interactions with race and urbanicity, indicating the greatest benefits were provided for minority students attending inner-city schools. More specifically, cross-sectional analyses revealed greater differential impacts of small classes on minority students on curriculum-based reading and mathematics tests. Further, minority students in the longitudinal sample experienced significantly greater relative growth on standardized tests in the second year of small class participation. All of these results were based on multivariate analysis of variance for mean scores on each outcome measure. Krueger (1999) also found that smaller classes had larger initial and cumulative effects on black, low income, and inner-city students, as compared to students that were white, not receiving free lunch, or living in other areas.

Larger impacts on low income children also emerged from phase 3 (Project Challenge) of the original Tennessee STAR experiment, which implemented reduced class sizes in grades K through 3 in the 17 poorest school districts (i.e. those with the lowest per capita income and greatest percentage of students receiving free or reduced-price lunch). Compared with previous performance by children in these districts, effect sizes were 0.4 for reading and 0.6 for math. Further, prior to the small classes, these districts performed well below the state average in math, and they moved above average after the intervention (Achilles, Nye, and Zaharias, 1995).

These districts were also compared to the other districts in the state by ranking them. Of the 139 districts, 1 is the best ranked and 139 is the worst ranked. At the start of the study (1989-90), these districts ranked 99 in reading and 85 in math, for second grade students. By the fourth year of the study, they found that the average rank in second grade

went down to 78 in reading and 56 in math, marking pronounced improvements in both

subject areas overall. It should be noted, however, that this was not a carefully controlled

experiment like Project STAR. As mentioned, data was compared with previous

performance among children and districts, and results must be interpreted accordingly.

In contrast to that which was just reviewed, other research provides minimal

evidence of differential effects of class size for different groups of students. For instance,

Nye, Hedges, and Konstantopoulos (2000a) examined differential effects of small classes

on minorities and students of low socioeconomic status in each grade of the study, as well

as pooled data across all grades. They tested several hierarchical linear regression models

to account for the clustering of children within schools. Models included both student-

level and school-level characteristics. Results showed that there were no statistically

significant differential effects of small classes on minority students in either math or

reading in any grade level. In other words, the small class advantage was not significantly

greater for minority students than for white students. However, while they were not

statistically significant, the small class effect for minorities was always larger than for

whites in both reading and math. Statistically significant differential effects on minority

students emerged for reading when data was pooled across grade levels.

Using the same models but substituting socioeconomic status for minority status,

the same authors found no statistically significant effects that favored small classes for

economically disadvantaged students over their more affluent peers. Although the

socioeconomic status and small class size interaction was statistically significant for

reading in grade 1, the small class advantage was actually greater for high SES students.

The analyses employed by these authors were more rigorous than those previously conducted, thus findings are noteworthy. Despite a lack of statistical significance, the small class effects for minorities were always larger. This was not the case for low SES students.

Konstantopoulos (2008) also examined STAR data in an attempt to understand the differential impacts of small classes on different types of students, specifically high and low achievers. The author examined not only mean differences between small and regular size classes in the sample, but also the degree of variability between class sizes. In other words, do student achievement scores converge or diverge as a result of attending small versus regular classes? Results showed that small classes produced significantly higher variability in achievement than regular classes in kindergarten in mathematics and in first grade in reading. Differences in variability favored small classes in second and third grade as well, though they were much less pronounced and statistically insignificant.

Small classes increased overall achievement as well as variability in students' achievement in kindergarten through third grade. Increased overall achievement (i.e. increased mean scores and a shifted achievement distribution) suggests that all students benefit from small classes, and increased variability within small classes suggests that high-achieving students may have benefited more from the small classes. This notion was further supported by results from quantile regression analyses, which were performed on five groupings of student scores across the achievement distribution. Results showed that there was a larger small class advantage at the upper quantile compared to that at the middle or lower quantile, indicating that in general high achievers benefited more from

being in small classes. However, it should be noted that differences between quantiles were statistically significant in only some grade levels, and at times in math and other times in reading, making it difficult to explain the findings.

Similarly, Nye, Hedges, and Konstantopoulos (2002) also looked at the differential effects of small classes on low achieving students, specifically those students in the bottom half or bottom quarter of their class's achievement distribution. Data was analyzed for each grade separately (kindergarten through third grade) and for pooled data across all four years. In terms of differential impacts on low achievers, the small class effect in reading for lower achieving students (bottom half) was larger than for high achievieg students, at every grade level. In contrast, the small class effect in math was larger for high achieving students than for low achieving students, in every grade level. None of these differences were statistically significant.

The same pattern emerged for children in the bottom quarter of their class's achievement distribution, indicating that class size did not differentially impact the lowest achievers. For both the bottom half and bottom quarter of student achievement levels, an analysis of data pooled across grades again did not show any statistically significant differences attributable to class size. These findings are similar to those reported by Konstantopolous (2008), described previously, in that small classes did not prove to have differentially larger effects for lower achieving students, and effects on lower achievers in mathematics were actually smaller than those for high achieving students.

Class size effects by school. Konstantopoulos (2011) drew upon STAR data to examine class size effects differentially by school. Results showed that there was a great

deal of variability between schools. On average, the small class effect was significant and

positive in early grades. However, within individual schools small class effects were at

times zero, at times significant and positive, and at other times significant and negative.

This was also true for regular class with full-time aide effects, which on average was

small and nonsignificant. In sum, students benefit from small classes and regular classes

with a full-time aide in many schools, though in some schools it is not beneficial and can

even be a disadvantage. There was no general pattern found in terms of the types of

schools that benefitted more or less so from small class sizes.

Number of years of attendance in small classes. Additional research that

utilizes STAR data examines effects in relation to the number of years students attended

small classes. Analyses such as these address some of the complex issues of the STAR

sample, for instance, many students moved and were replaced by different students

within the same class type as time went on, due to student mobility and grade retention.

Students that were retained in any grade were also removed from the study and replaced

by new students in their same class type. In addition, kindergarten was not mandatory in

Tennessee at the time of the experiment, so students who entered the study in grades 1, 2,

or 3 had a range of different possible experiences, thus creating new variables that needed

to be accounted for.

 Finn et. al (2001) attempted to understand the magnitude of small class size

effects relative to the number of years students attend small classes. By examining the

effects of the number of years students attended small classes, they were able to control

for student mobility and address the issue of in-grade retention. Using hierarchal linear

modeling, the authors report grade-equivalent SAT test results that estimate the benefits of attending a small class in "months of schooling."

Results showed that students who attended small classes outperformed their regular class size peers in kindergarten thru third grade on all measures of achievement. No overall differences were found between children who attended regular classes and those with a full-time aide. When students entered small class sizes earlier and remained in them over a few years, they achieved significantly better than those who did not, and the gap in achievement levels widened over time. Specifically, by third grade the advantages of attending a small class for two years (since grade 2), for three years (since grade 1), and for four years (since kindergarten) were approximately 3.3 months, 5.2 months, and 7.1 months, respectively. Similar results were found by Nye, Hedges, and Konstantopoulos (2001) in that, at every grade level, the effects of small classes were greater for more years spent in small classes.

In contrast, Krueger (1999) suggested that the largest benefits from small classes appear to occur in the first year of treatment. The author provides an econometric analysis of achievement for the 4-year Tennessee STAR sample of 11,600 students. This allows for the examination of class size in light of school resources. Accounting for several flaws in the study's experimental design (to be discussed in greater detail later in this paper), findings show that the flaws did not jeopardize its main results: that students in small classes scored higher on standardized tests than students in regular size classes. However, the first year of attending a small class had the largest effect in this analysis, and time spent in small classes in subsequent years had positive, but smaller effects on

children's achievement scores. Specifically, performance on standardized math and reading tests increased by four percentile points the first year students attended small classes, and about one percentile point each year thereafter. In short, results show that the only notable benefits from small classes occur in the first year of students' attendance. These findings allow for various implications, one of which is that attending small classes in very early grades may have a one-time effect that permanently raises the level of student achievement without significantly impacting the achievement trajectory.

On average, the research that has just been reviewed provides evidence for positive effects of small classes on children's achievement in the early grades. There is evidence that small classes matter, and that the actual size of the class should be considered rather than just attending to ratios. For instance, a ratio of 20 students to 1 adult appears to be better than a ratio of 40:2 for children's overall achievement. However, findings vary regarding small class effects on different groups of children, for example minorities, disadvantaged children, or children from different areas and in different schools.

Despite the overall generally positive findings, there are conflicting conclusions in the literature that was just reviewed. These conflicting conclusions coupled with policies to reduce class size based largely on the findings from this one large-scale study have led to more recent analyses concentrated on evaluating threats to the validity of STAR findings. Examples include implementation problems, attrition, and switching between groups.

Validity of Findings from Project STAR

The credibility of findings from Project STAR is based largely on its status as a randomized trial study with a rigorous research design; therefore, the efficacy of its design is tested in several different experiments. In other words, a number of research studies have examined whether students in small classes were systematically different from those in regular-sized classes in any significant ways. If systematic differences exist, randomization may not have been effective and estimates of class size effects are questionable due to potential selection bias. This can be tested by examining differences between sample characteristics or attrition, for example. There was a great deal of attrition in the STAR project, and as students dropped out new students were randomly assigned to replace them. Even though this process was random, little is known of students' prior school experiences, which adds uncertainty to the experiment. Analyses of pretest data would give a valid comparison of groups prior to treatment, however the fact that there was no pretest data was collected for any children makes it difficult to verify the randomization of the study.

It is also important to look at the fidelity of class size implementation. In other words, are the actual sizes of small and large classes what they were assigned to be? And did children who were assigned to small classes actually remain in small classes throughout the course of the study? Additional criticisms of the research design have included the fact that schools in the study were not random; they had to agree to participate and accommodate at least three classes in each grade, making the generalization of results questionable. Teachers were also not random and there is limited

information available on teachers' classroom practices and behaviors. Adding to this uncertainty is the fact that there is very little detail on how random assignment of teachers was actually done.

Findings from analyses that have addressed fidelity issues such as these have been mixed. Nye, Hedges, and Konstantopoulos (2000) found that the actual sizes of the classes assigned to be large versus small were within the intended range the majority of the time. They also found that the percentages of children that were assigned to small and regular size classes remained generally consistent across time. Additionally, an examination of class size effects based on actual treatment versus 'intent to treat', or the class size to which students were originally assigned showed that the effect of small classes was significant and positive in both mathematics and reading at every grade level. The same pattern of results emerged regardless of whether initial assignment or actual classroom type was used for the analyses. Thus, they concluded that treatment switching did not greatly influence estimates of the effect of small classes.

Konstantopoulos (2011) also reported on transitions from small to regular class sizes (and vice versa), albeit differently. Findings revealed that nearly 15% of students who were in regular size classes in kindergarten moved to small classes in first grade, and this pattern persisted in second and third grades. At the same time, the percentage of students who moved from small classes to regular classes was around 15% in first grade, but only 2-4% in second and third grades, suggesting a potential existing bias in the estimated effect of small classes.

Krueger (1999) examined the effectiveness of the randomization among the three treatment groups (small, regular, and regular classes with a full-time aide) by looking at socio-economic status (i.e. free lunch), minority group status, and age as they were related to school effects, attrition, re-randomization after kindergarten, nonrandom transitions, and variability in actual class size. Results showed that there were some significant differences between groups regarding the three background variables. However, differences were not statistically significant when school effects were accounted for. Additionally, there were no significant differences regarding teacher characteristics (experience, race, or education) by class type. Further, values for missing test score data (in light of attrition) were imputed to see if the results of the experiment would change when the imputed values were added. Findings showed that the treatment effects were largely unaffected by attrition, but that they seemed to be greatest among students who began the experiment in kindergarten. In contrast, Konstantopoulos (2011) examined similar variables at the school level and found systematic differences for SES and age. This suggests that randomization at the school level was perhaps unsuccessful.

Attrition was also examined as a way to test the effectiveness of randomization, again producing mixed findings. Nye, Hedges, and Konstantopoulos (2000) found that for every year and every subject matter the STAR treatment effect for children who left the study was nearly identical to those who remained in the study, based on their previous year's test scores. In contrast, Blatchford, Goldstein, and Mortimore (1998) and Kreuger (1999) found that attrition was not random. Specifically, those children dropping out of the experiment in first grade had kindergarten achievement scores noticeably below

average, and the differential below average was larger for those who started in regular

size classrooms than for those who were in small kindergarten classrooms.

Konstantopoulous (2011) examined the percentage of students from each school that

dropped out each year in relation to class size effects, achievement, and school

composition. Overall, percentage attrition in a school was related to school achievement

and school composition in some grades; however it was not associated with small class or

full-time aide effects. That class size was not associated with attrition rates lends

credibility to the randomization; however, since attrition was related to achievement and

school composition selection bias from grade to grade is indeed possible.

Another way to address all of the randomization effectiveness issues is to

compare the full STAR sample with those students who remained in the sample for the

entire 4 years. In doing so, Hanushek (1999) finds that the kindergarten differential for

the 4-year sample is slightly larger than that for the annual sample, in both reading and

math. However, the differential effect decreases over time and is much smaller by the

time children are in third grade. In contrast, findings from the Lasting Benefits Study

(Nye, Zaharias, Fulton, & Achilles, 1993) showed that students who attended small

classes in the early grades outperformed those who attended regular classes in 6th grade.

Differences were found to be very close to those reported in original third grade analyses

of STAR data (.22 and .18 standard deviations in reading and math, respectively). That

differentials do not get larger over time, as might be expected as a result of cumulative

effects of small class sizes, is a concern that has been expressed regarding these results

(Hanushek, 1999).

Project STAR was a well-designed study, though as with any well-designed study there were real life circumstances that caused deviation from the intended research plan. These deviations have been studied in depth and while they do represent valid concerns, the majority of research shows that they do not give cause to dismiss the overall positive findings.   In general, results from Project STAR can be used to support policies for class size reductions. However, despite how large and well-researched Project STAR was, one sample alone does not prove that small classes are better for students. Research in addition to Project STAR that has focused on the effects of different class sizes will be reviewed next.

Research in Addition to Project STAR

In addition to the Tennessee STAR experiment, a number of studies point to the positive impact of small classes. Wisconsin's Project SAGE is one example of a reduced class size initiative that was shown to have positive impacts on children's achievement. As mentioned in a previous section, Wisconsin's Project SAGE (Student Achievement Guarantee in Education) adopted a program that reduced student-teacher ratios to 15:1 in kindergarten and first grade classrooms. These ratios allowed for various configurations, such as 15:1, 30:2, and 45:3. On average, ratios were between 12:1 and 15:1.

The SAGE program was targeted toward school districts that had at least one school with 50% of children or more living below the poverty level. In all, 30 schools in 21 districts became part of the program, and they were compared to 14 to 17 pre-selected comparison schools with similar student and school characteristics. Student-teacher ratios in comparison classrooms averaged between 21:1 and 25:1.

A quasi-experimental, comparative change design was used to examine the effects of small classes on children's achievement. First grade results from regression analyses in both 1996-97 and 1997-98 indicated positive impacts of small classes on student achievement in math, reading, and language arts, especially for minority students. Interestingly, in contrast to results from Project STAR, there were no statistically significant differences attributable to ratios of 15:1 versus 30:2 (Molnar, et al., 1999).

California implemented a class size reduction (CSR) program in June of 1996 for all kindergarten through third grade classrooms. At the time, the average class size was 28 students per classroom, and the classrooms were to be capped at 20 students. The program was voluntary and provided a financial incentive for schools to participate. By the 1998-1999 school year, 92% of students in kindergarten through third grade throughout the state were participating. Results showed that third grade students enrolled in reduced classes performed better on the Stanford Achievement Test (SAT) than did students in regular size classes (Stecher et al., 2001). This was the case in both the second and third year of CSR, and whether students attended small classes for one versus two years did not appear to make much difference. Gains persisted into fourth grade, and gains were similar among all students, regardless of income, minority status, or fluency in English.

While results from CSR positively favored small classes, it is difficult to put much stake in these findings since the experiment was not controlled. Various differences between children who did and did not attend small classes could have biased the findings in favor of small classes. For instance, districts serving the greatest number of low-

income and minority students were more likely than others to report that the cost to implement CSR was greater than revenues generated by the initiative. Thus, comparing students in regular classes in these districts with students in small classes in districts that could afford the initiative clearly introduces extraneous variables into the equation. However, since the qualifications of teachers declined largely with the implementation of the initiative (due to the need for many more teachers in such a short period of time) and small classes still had positive impacts for children, small classes may have actually contributed more to children's achievement than we initially perceive (Stecher et al., 2001).

A study conducted by the NICHD Early Child Care Research Network (2004), mentioned previously in relation to classroom life, examined the extent to which natural variation in observed class size was related to children's outcomes, and found that children who attended class sizes smaller than 21 had significantly higher literacy scores in first grade. However, no significant child outcome findings were found across the full distribution of class size but were detected only within certain limits. In other words, there were specific class sizes that proved to have significant effects, but not necessarily the smallest of them. For instance, teachers reported better social adjustment for children as class size increases up to a point near 20 or so, after which the effects of class size become negligible. This cut-off occurred in class sizes of 21 or 22 children for cognitive outcomes. In other words, there was a significant drop in children's WJ achievement scores when class sizes were larger than 21, but any class sizes that were smaller than 21 did not appear to impact children much differently.

Minimal class size effects. In contrast to the research that was just reviewed, several studies claim that small classes have little to no effects on student achievement. Hanushek (1999) argues against the reduction of class sizes for a number of reasons. He reviews findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. The author argues that both pupil-teacher ratios and class sizes have fallen over time and student performance has not improved, therefore class size does not lead to improved academic performance. However, this information alone does not provide evidence that reducing class sizes has no effect for a number of reasons. For example, one could point to the expansion of special education instruction, which might mask the decrease in pupil-teacher ratios that may have not actually decreased in mainstream classrooms. There is also the possibility that the student population could have changed over time, in terms of motivation and preparation, resulting in decreased student achievement levels. For instance, the percentage of children living in poverty rose from 14.9% to 19.9% between 1970 and 1990.

Both of the aforementioned possibilities have been addressed to some degree. For instance, a study by Hanushek and Rivkin (1997) examined how much the changes in special education could have affected the observed pupil-teacher ratio during the 1980s and concluded that changes in ratios cannot be attributed simply to overall changes in special education. Additionally, according to Grissmer, Kirby, Berends, and Williamson (1994), student backgrounds, in terms of poverty, appear to have improved for the nation as a whole over time. Therefore, it does not seem to be a viable contention that adverse

student backgrounds offset the potentially beneficial effects of smaller pupil-teacher ratios.

Hanushek (1999) also summarizes the available results for estimates of the effects of teacher-pupil ratios on student outcomes, as a basis for his argument against the reduction of class sizes. Drawing upon 277 separate estimates of the effect of class size or teacher-pupil ratio through 1994, he shows that only 15% of all studies found a positive and statistically significant relationship between teacher-pupil ratio and student performance. When less rigorous studies were dropped from analyses, this percentage dropped to 12%. Thus, he concludes that econometric evidence as a whole gives little support to the idea that smaller classes will lead to better academic achievement.

Shapson et al. (1980) also found minimal positive effects attributable to smaller classes. The authors examined the effects of four class sizes (16, 23, 30, and 37) on students' achievement in reading, math, composition, and art. Students from 11 schools in 62 classrooms (16 classes of each class size) in Metropolitan Toronto were randomly assigned to one of the four class sizes in fourth grade, and remained in the same class size through fifth grade. It was required that no student could be in a class size of 16 or 37 for both years of the study and no teacher could teach two larger or two smaller size classes for both years of the study. Variability due to year of the study was removed prior to analyses of children's outcomes, using a multiple linear regression technique. Subsequently, findings from an analysis of variance revealed no significant differences attributable to class size for art (samples were rated on a developmental scale), composition (samples on specific topics were rated on a five-point scale), vocabulary,

reading, and mathematics problem solving. Students in class sizes of 16 had significantly higher scores for mathematical concepts, compared to those in class sizes of 30 and 37, though not compared to those in class sizes of 23.

In addition, an analysis of data from the Early Childhood Longitudinal Study-Kindergarten Cohort found that kindergarten class size had only small effects on reading and none on math (Walston & West, 2004). These effects were in relation to half-day and full-day kindergarten programs. In other words, class size was not found to mitigate the difference in gains found between children in half-day and full-day programs. The majority of classes ranged from 18 to 24 students, though there were small percentages of classrooms with less than 18 or more than 25 students. About fifty percent of classrooms were reported to have an aide that works directly with the students for at least an hour per day.

Results from meta-analyses. Class size was a topic of interest prior to the Tennessee STAR study, but there was a lack of rigorous methodology used, and there were mixed reviews on the effects of small classes. In light of this, Glass & Smith (1979) conducted a meta-analysis on the effects of class size. A meta-analysis can be defined as that which combines the results of numerous research studies that address a set of related research hypotheses. Their approach began with a thorough gathering of literature related to class size, resulting in over 300 documents that were obtained and read. Of these documents, about 150 of them did not report on actual data and were more review type documents. Of the remaining 150, about 70 of them examined class size in relation to classroom process variables, which was the focus of the previous section of this paper. As

a result, 77 studies of class size and student achievement were included in their meta-analysis.

They began by coding studies on a wide range of different properties that might interact with the relationship between class size and achievement, such as subject of instruction, year of study, or the number of pupils upon which small and large class achievement means were based. In effect, a single study ended up with multiple comparisons, depending, for instance, upon how many different class sizes were compared or how many different achievement tests were reported.

Based on an accumulation of 725 comparisons across 77 studies that drew data from nearly 900,000 pupils in over 12 countries, a number of interesting findings emerged. There was an enormous range in class sizes, and those found to produce the greatest differences in children's achievement were found when class sizes of 40 were compared to class sizes of 1. However, notable differences were also found for class sizes of 15 or fewer. Specifically, the difference in achievement resulting from instruction in groups of 20 pupils to that of 10 pupils was, on average, 10 or more percentile ranks. In other words, those who attended small classes achieved at the 60th percentile, in comparison to those who attended regular size classes and scored at the 50th percentile.

It is also interesting to note that findings were more pronounced for secondary level children (ages 12 and older). Further, the full scale analyses included both well-controlled (i.e. random assignment) and uncontrolled studies, which could lead to uncertainty in the interpretation of findings. In effect, the authors examined only those that employed rigorous methodology separately (n=14) and found even greater effects

favoring small classes. Again, differences were pronounced for class sizes of 15 or fewer, and essentially null for those with 20 or more.

By contrast, analyses of Glass & Smith data by other authors did not produce the same findings. For instance, Slavin (1990) analyzed selected studies with stringent inclusion criteria and found effects of very small groups and 1-1 teaching but minimal differences for class sizes of 15 up to 30. Meta-analyses have become a more popular technique in recent decades, however there have been none specifically used for the examination of class size since Slavin's 1990 analysis.

Class Size and Children's Preschool Experiences

Given the large body of research on class size in later grades, and its frequent positive associations, it is not surprising that its importance has trickled down into the early education arena. Theory and evidence indicate that preschool children should benefit from small class size even more than kindergarten children do. For instance, the kinds of teacher and child behaviors that were affected by the STAR class size reduction (i.e. individualized instruction, greater engagement among students) present great potential for producing educationally effective preschool programs (e.g. Boat, Dinnebel, & Bay, 2010; Howes et al., 2008; Lo-Casale Crouch et al., 2007). In addition, developmentally preschool children are still learning to be part of a group and work independently so this period is crucial for the advancement of such skills. Further, parents' demand for early care and education services has surged over the last 30 years, as higher percentages of parents have entered the workforce and families have sought experiences that promote early learning and development. This creates a greater demand

to understand how different classroom features promote positive early learning experiences, for the sake of investing money in what benefits children most.

As will be revealed in the following paragraphs, several early childhood studies draw upon NICHD data to examine class size. In general, results from these studies point to positive associations between small classes and children's classroom experiences. This varies from conclusions that have been drawn from NICHD data for older children, which in some cases have found associations between larger classes and more positive classroom experiences for children.

Class Size and Children's Classroom Behavior

A small body of early education research reveals positive relationships between reduced class sizes and children's classroom experiences. For instance, Blatchford, Edmonds, and Martin (2003) examined the relationship between class size for children ages 4-7 and their behavioral adjustment to school. Children began in the study at age 4 and were studied longitudinally for three years. Children's off-task behavior and peer relations were analyzed using data from observations and teacher reports.

Observations were quantitative in nature and involved the coding of behaviors and interactions of a random group of 6 individual children in each classroom. Each classroom was observed 5 times per day for 3 days during various classroom-based work activities such as language, math, and free play. Naturally occurring class sizes of less than 20 (small) were compared to those that included more than 30 (large). In sum, children in small classes were significantly less likely to be off-task, and significantly more likely to be attending to the teacher or to their work when on their own. In contrast,

children in large classes were more likely to engage in social interactions with their peers. There were no significant differences in terms of peer relations; however, it was more likely for children in larger classes to be rated by teachers as less aggressive, asocial, and less excluded.

Class Size and Teaching Time

Drawing from a subset of the same sample, Blatchford, Moriarty, Edmonds, and Martin (2002) examined relationships between class sizes and classroom processes, using data from teacher questionnaires, case studies of individual classrooms, teacher estimates of time allocation, and classroom observations. Group sizes (again, less than 20 versus more than 30) were examined in relation to differences in time spent on teaching or instructional activities; time in individual, group, and class contexts; amount of teacher-child contact and individual attention from teachers; and qualitative dimensions of interactions between teacher and children.

Results from observations (described for the previous study) and teacher estimates of time allocation (also quantitative in nature) showed that children in small classes were more likely to interact with their teachers and these interactions were more task-related and involved greater initiation among students, more on-on-one teaching took place, and children were more often the focus of a teachers' attention. All of these differences were statistically significant. Interactions between teachers and children in small classes were also more social in nature, indicating that they were more personalized. In addition, teaching in both individual and group contexts increased as class size decreased.

Findings from questionnaires and case studies, which were more qualitative in nature, suggested that children in small classes were given more individual attention, teachers were more responsive to children in small classes, and interactions between teachers and children in smaller classes were more purposeful and sustained. However, it is important to note that these richer interactions were not simply the result of more one-on-one teaching since, as mentioned previously, there was also more time spent in group contexts in small classes. Further, case studies showed that teachers varied in how they responded to small classes, and some seemed to adapt better than others. For instance, small classes in some cases led to more frequent interruptions given their tendency to allow immediate feedback. This could be due to the fact that children in small classes might expect to have their demands met instantly, which is not feasible in larger classes. As a result, small classes did not lead to inevitable benefits for children; rather, teachers had to work strategically to manage children's learning effectively.

Class Size and Teacher-Child Interactions

Results from classrooms across 57 sites that participated in the National Day Care Study showed that group size and ratios impacted classroom life (Ruopp et al., 1979). The 4-year study utilized three phases to understand the relationship between structural center characteristics and children's classroom experiences and educational outcomes. The first two phases were used to pilot and fine-tune the design, and phase three was used to test the research questions. Phase three included 57 strategically-sampled centers and children in Atlanta, Detroit, and Seattle, focusing primarily on low-income, urban families. Three groups were compared in terms of adult to child ratios: one group had a

naturally-occurring mean ratio of 1:9.1; another group had a naturally-occurring mean ratio of 1:5.9; and the third group had an experimentally assigned mean ratio of 1:5.9, allowing for a randomized trial. Group size was distributed evenly across these three groups, so that the effect of ratio could be singled out. Most centers had group sizes between 12 and 24 children.

Each classroom was observed by a trained data collector during both morning and afternoon activities. Classrooms were rated in terms of the overall quality of the environment as well as teacher-child interactions. Results showed that when children were in larger classes they received less interaction from adults; exhibited more crying, apathy, and wandering; engaged in less conversation, less interaction, and less focused play. Teachers were also more often engaged in dealing with behavior problems and there was less cognitive and language stimulation in larger classes. In smaller classes, children were more cooperative, more verbal, and more responsive to adults and peers. Group size did not appear to have an effect on the amount of individual attention teachers gave to students; however, teachers in larger classes paid more attention to large groups and less to small groups.

Effects were found to be greater in relation to group size, versus ratio. In other words, smaller group sizes had pronounced effects on children's classroom experiences, while the addition of more adults to large groups of children did not have the same impact. However, since most classrooms in the study maintained ratios of 5 to 9 children per adult, this narrow range may have caused the effects of small teacher-child ratios to be underestimated (Ruopp et al., 1979).

Similarly, *Eager to Learn*, a report by the National Research Council on preschool education, examined over 20 research studies and concluded that child-initiated activities were more common in smaller classes. They also found that with more time for each child, teachers were better able to work on extending children's language experiences. In addition, when there were fewer children in the room, teachers were able to more closely mediate children's social interaction. Small group size was also correlated with children being more active in their learning and there was less restrictive and controlling behavior among teachers (National Research Council, 2001).

Phillipsen et al. (1997) examined the relationship between adult-child ratios (as well as a number of other classroom characteristics) and three measures of preschool classroom quality. Classroom quality was measured with the Early Childhood Environment Rating Scale (ECERS), the Caregiver Interaction Scale (CIS) and the Teacher Involvement Scale (TIS). The ECERS is a comprehensive measure of quality that gauges teachers' practices related to classroom activities, health and safety, language and reasoning, interactions with children, as well as the overall structural provisions in the classroom. The CIS measures teacher interactions with children in terms of sensitivity, harshness, and detachment. The TIS captures the amount and quality of individualized teacher-child interactions.

Data was collected from a stratified (profit/non-profit) random sample of 100 programs across four different states, in a total of 360 classrooms. Each classroom was observed by a trained data collector for one day in the morning. Adult-child ratios and group sizes were collected during the time of the quality observations. Results from

hierarchical regressions showed that higher adult-child ratios (fewer children per adult) were predictive of higher ECERS and CIS scores. High ratios were also significantly related to the proportion of time in which preschool teachers were observed as being responsive to individual children, as shown through the TIS. These effects were largest when the teacher-child ratio decreased from 1:10 to 1:5, and much less pronounced when decreasing from 1:5 to 1:3.

Similarly, results from a large-scale study conducted by the National Institute of Child Health and Human Development (NICHD) revealed associations between small classes and positive interactions among teachers and children at 36-months of age (NICHD ECCRN, 2000). The NICHD study was an extensive experiment of early child care, following more than 1,000 children through their first four years of life. Mothers were recruited from hospitals in multiple US locations during selected 24-hour sampling periods in 1991. Of 8,986 mothers that were giving birth, 5,265 were contacted for the study after returning home from the hospital. Mothers were excluded if they were under age 18, had multiple births, anticipated moving from the area soon, had medical or substance abuse problems, did not speak English, lived far away, or in unsafe neighborhoods. When infants were 1-month old, 1,364 families were enrolled in the study. The NICHD sample was referred to earlier in this paper in the context of a study on class size and children's classroom experiences in the early grades.

One component of the NICHD study included child care observations at multiple time points. When children were 36-months old, classroom interactions were rated by external observers at both the individual child level and classroom level in terms of

positive caregiving frequency (i.e. amount of staff attentiveness and responsiveness to children; expressed positive affect and affections; degree of restriction and intrusiveness, and teaching of academic skills), and qualitative ratings of positive caregiving quality (i.e. sensitivity, positive regard, affect). Group size (mean of 7.3) and ratio (mean of 1:4.6) were significantly related to frequencies of positive caregiving when children were 36-months. In other words, classrooms with smaller groups and fewer children per teacher were characterized by greater amounts of positive caregiving. However, there were no significant relationships with qualitative ratings of positive caregiving (NICHD ECCRN, 2000).

Children participating in the NICHD study were followed through 54-months of age, at which point child care settings were observed once again (NICHD, 2002). At this time, child care settings were assessed in terms of caregivers' relationship with children (sensitivity to nondistress, detachment, stimulation of cognitive development, and intrusiveness), and classroom setting (chaos, overcontrol, positive emotional climate, and negative emotional climate). Child-staff ratios were recorded at four time points during the observations and averaged for analyses. Results from structural equation models showed that smaller child to teacher ratios (i.e. fewer children per teacher) significantly predicted levels of teacher cognitive stimulation, teacher detachment (inversely) and positive emotional classroom climate (NICHD 2002). While smaller ratios were reported to be better, the range of ratios that occurred was not indicated and we do not have a clear understanding of what makes a difference for teachers' practice.

Class Size and Children's Preschool Achievement

There is limited research on the relationship between class size in preschool and children's achievement. Of those studies that do exist, findings are mixed. Again, NICHD data is drawn upon for some of these studies, and findings again point to positive associations between class size and children's cognitive development in preschool. However, the distinction between group size and teacher-child ratio is often unclear, making implications for class size in preschool difficult to identify.

A handful of studies provide evidence for positive relationships between smaller classes and children's early learning skills. For instance, the National Day Care Study, mentioned previously as it related to ratios and classroom experiences, found a relationship between teacher-child ratio and preschoolers' developmental outcomes, but the effect was minimal (Ruopp et al., 1979). Specifically, 3-year olds that attended classes with high staff to child ratios (more adults per child) made more rapid gains on the PSI, an assessment of basic knowledge and school-related skills such as shapes, sizes, parts of the body, etc. No significant effects were found for 4-year olds. Group size, on the other hand, was shown to have consistently positive and significant effects on children's knowledge skills as well as vocabulary development at ages 3 and 4.

Class Size and Language Development

A study by Burchinal et al. (2000) also found positive associations between low child to teacher ratios (fewer children per adult) and children's development. The authors examined the contribution of ratio to the expressive and receptive language skills of 89

economically disadvantaged African-American children over time. Children entered the study between one and eleven months of age and were followed to age 3.

The ratio of children to adults in 3-year old classes ranged from 3:1 to 16:1. At the time, the American Academy of Pediatrics (AAP) recommended ratios of 5:1 or less for children between 31 and 35 months and 7:1 or less for children 36 months and older. Since many children in the 3-year old classes were not yet 36 months at the time of the study, the recommended ratios were averaged for the sake of analyses. Results from regression analyses showed that children's expressive and receptive language skills were significantly better when preschool classrooms met recommended teacher-child ratios (Burchinal, Roberts, Riggins, Seisel, Neebe, & Bryant, 2000).

Similarly, one set of analyses that utilized data from the NICHD study showed that children in classrooms that met more recommendations for regulatable features, such as class size, ratios, or teacher training and education, had better language comprehension and few behavior problems at age 3 (NICHD, 1999). In this case, recommended levels were 1:7 for teacher-child ratio and 14 for group size, for children 36 months of age and older. Teacher-child ratios alone were found to have significant effects only on children's behavior problems. Group size did not appear to have an impact on children's outcomes.

In addition, smaller teacher-child ratios (fewer children per teacher) were found to be positively related to children's cognitive and social competence at 54 months of age, by means of classroom processes (NICHD ECCRN, 2002). In other words, structural equation modeling was used to test the path from teacher-child ratio to child outcomes, as mediated by classroom processes. The mediated path was found to be significant,

suggesting that teacher-child ratios impact child outcomes through their effects on classroom processes. However, since structural equation models rely on correlational data, the causality cannot be inferred from these findings. Further, the effects of center-based care overall, regardless of ratio, were quite small compared to those found for maternal care.

Class Size and Cognitive Development

A number of large-scale, well-known preschool research studies, such as High/Scope and Chicago Child-Parent centers, have shown large educational benefits for economically disadvantaged children that attended programs with highly capable teachers and relatively small groups of children. Most of these studies showed increasingly large effects over long periods of time (e.g. Schwienhart, et al, 2005; Reynolds, Temple, Robertson, & Mann, 2002). However, class size specifically was not examined, thus they do not prove that small class size is effective.

In addition, the largest and most recent meta-analysis confirms that preschool education has significant lasting positive effects on cognitive abilities, school progress (e.g., less grade repetition, less special education placement, and increased high school graduation), and social behavior, and effects are larger when programs focus on small group learning and individualized one-on-one teaching (Camilli et al., 2010). Further, findings from the Effective Provision of Preschool Education (EPPE) project showed that the most effective preschool settings (in terms of their impact on children's cognitive outcomes) were those that encouraged 'sustained shared thinking,' or extended narratives between teachers and children, through 1:1 adult-child interactions (Siraj-Blatchford et

al., 2003). While these studies do not specifically examine the effects of class size,

findings point to the effectiveness of practices that have been associated with smaller

classes.

Minimal Class Size Effects

Finally, in some cases research has failed to show connections between class size

and children's achievement in preschool. For instance, as mentioned previously, the

NICHD study showed that group size was positively related to children's cognitive and

academic achievement at age 4 ½, but there was no evidence of positive associations with

staff-child ratios (NICHD ECCRN & Duncan, 2003). This is difficult to explain in light

of the positive relationships that were found to exist between staff-child ratios and

classroom quality (Phillipsen et al., 1997).

Further, in Dunn's (1993) exploratory study of 30 preschool caregivers, neither

staff-child ratio nor group size predicted children's cognitive outcomes. However, higher

staff-child ratios (more staff per child) were related to fewer behavior problems among

children. Group size ranged from 12 to 40, with an average of 21.47 and ratios ranged

from 1:9 to 1:20, with an average of 1:12.82.

<div align="center">

Summary

</div>

As evidenced in the review of literature just presented, there is a great deal of

research on class size in the early elementary grades. Many authors are starkly for or

against class size reduction, which is evidenced in the robust conclusions that are drawn

within the various articles on the topic. Regardless of which side one is on, research

evidence is provided as support for their position. Further, there is research evidence for

both sides that is of high quality, and advocates for one side typically do not ignore the conflicting evidence that exists. Because there are strong arguments on both sides of the class size debate, it is unclear why advocates are so strong in one direction versus another. It could have something to do with one's personal experiences or biases due to the high costs that are involved, but that is merely speculation.

A more plausible explanation for the different conclusions surrounding class size is the inconsistency in terms of how class size has been defined. For instance, the Tennessee STAR study compared children in K-3 that were randomly assigned to class sizes of 13-17 vs 22-25, and showed evidence of positive significant effects on reading and math in every grade for kids that attended small classes. Other studies have produced similar results but have defined small class size differently. For example, results from the NICHD (2004) study also showed positive effects of small classes on literacy scores in 1st grade, but in this case small classes were defined as those having 21 or fewer students.

There was also a meta-analysis completed by Glass & Smith (1979) that showed large effects on student achievement when class sizes were 15 or fewer. In contrast, Slavin (1990) also analyzed data used by Glass and Smith and found large effects for 1 on 1 teaching, but very small effects even for class sizes as small as 15 students. Shapson et al (1980) also found very little effects for class sizes of 16 compared to class sizes of 23, however there were differences when class sizes of 16 were compared to those as large as 30.

When we look at studies that focus more on children's classroom experiences in relation to class size (as opposed to their achievement), a similar trend emerges. For

instance, one study based in North Carolina found that class sizes of 14 (compared to 23) had more individual attention given to students and a greater degree of individualized communication among teachers and students in small classes. Another study conducted by NICHD ECCRN network (2004) found that for class sizes of fewer than 23 children, there was significantly more negative and disruptive behavior in classrooms. Classrooms containing more than 25 children were found to have more interactions among teachers and children, but lower overall instructional quality. Across all of these studies, smaller classes did not always lead to better experiences and outcomes for children.

While many individuals would argue that smaller classes are better in the early elementary grades, implications in fact remain unclear. Because of the inconsistencies that existed across the research in terms of defining a small class size, the question remains: how small is small enough? We can attempt to answer this question by drawing on the moderate amount of literature that exists for class sizes in the early elementary grades, but this is not the case for the preschool period. As was portrayed in the literature review in Chapter 2, research on class size in preschool is minimal, and especially so as it relates to student achievement.

A few studies have been conducted for this age, but none of them have used a rigorous design. Of those that have been conducted, 3 different studies that utilized NICHD data found effects on children's cognitive and academic achievement when teacher-child ratios were 1:7 (Burchinal et al, 2000; NICHD, 1999, 2002). Another study examining 360 classrooms across 4 states found large effects on quality scores when ratios decreased from 1:10 to 1:5. There was also the Nat'l Day Care Study in 1979,

which found effects on children when class sizes were smaller, for classes ranging from 12 to 24.

It is important to note that several of these studies are based on child care samples rather than what is currently referred to as *preschool* or *pre-K* which is characterized by a much different set of standards and goals. Formal preschool education throughout the country today is largely focused on advancing children's early skills and achievement while child care programs tend to focus on providing a place for children to socialize with little emphasis on academic achievement. This is not the case for all child care programs that exist today, but the potential difference is still important to point out. Despite the marked differences between child care and preschool education in terms of standards and goals, there are also many similarities that are inevitable. For instance, regardless of the goals, children are at the same developmental stage in life and they are grouped together in formal care with others of the same age outside of the home. Ultimately, implications may be different for each one, but since practices are aligned to some degree it is important to consider research that examines both types of care.

The general conclusions from small scale early education studies point to positive effects of small classes as they relate to the classroom environment and teacher practices. Further, a number of large-scale studies have shown positive effects on children's achievement as a result of attending programs that were characterized by small classes. However, since class size was neither randomly assigned nor studied specifically, and is just one of many variables that are studied in these experiments, it is difficult to understand its effectiveness. Additionally, a number of studies report of the importance of

smaller classes but they fail to identify what size is actually necessary to have an impact on children. In other words, they simply conclude that smaller is better.

The current context for preschool is much different now than it was back in 1979 and 1991, when the large-scale early education class size experiments took place. Preschool enrollment in the US has been growing over the past couple decades, and especially programs funded by state dollars (Barnett, Carolyn, Fitzgerald, and Squires, 2012). Further, the preschool landscape has evolved to one defined by standards ranging from teachers' education level to class size and teacher-child ratios. States and programs have adopted various standards for class size and teacher-child ratios. The majority of these programs implement class sizes of 20 and student-teacher ratios of 10:1. However, findings from STAR and the few that have been conducted for preschool age students (NICHD, 1999, 2002), suggest that class sizes smaller than 20 might be necessary for enhancing classroom practices and increasing children's achievement.

Additionally, it is possible that class size impacts diverse populations of children differentially in preschool because of the multitude of intersecting variables that come into play. It is also possible that class size differences work in several different ways to effect children's achievement, and that all of the dissimilarities that exist between teachers are more influential than perhaps class size alone. The current study will address the gap in our knowledge that exists around these possibilities using Bronfrenbrenner's Ecological Systems Theory as the guiding framework. Ecological Systems Theory places children's development in an ecological perspective, in which an individual's development is nested within interconnected systems (Bronfrenbrenner, 1989). For

instance, children influence classroom practices which influence family systems, and vice versa. All of these are nested within larger systems such as governmental policies and regulations (Marshall, 2004). The diagram below portrays the research framework for the current study as existing within the context of the Ecological Systems theoretical framework.

Figure 1. Theoretical Framework
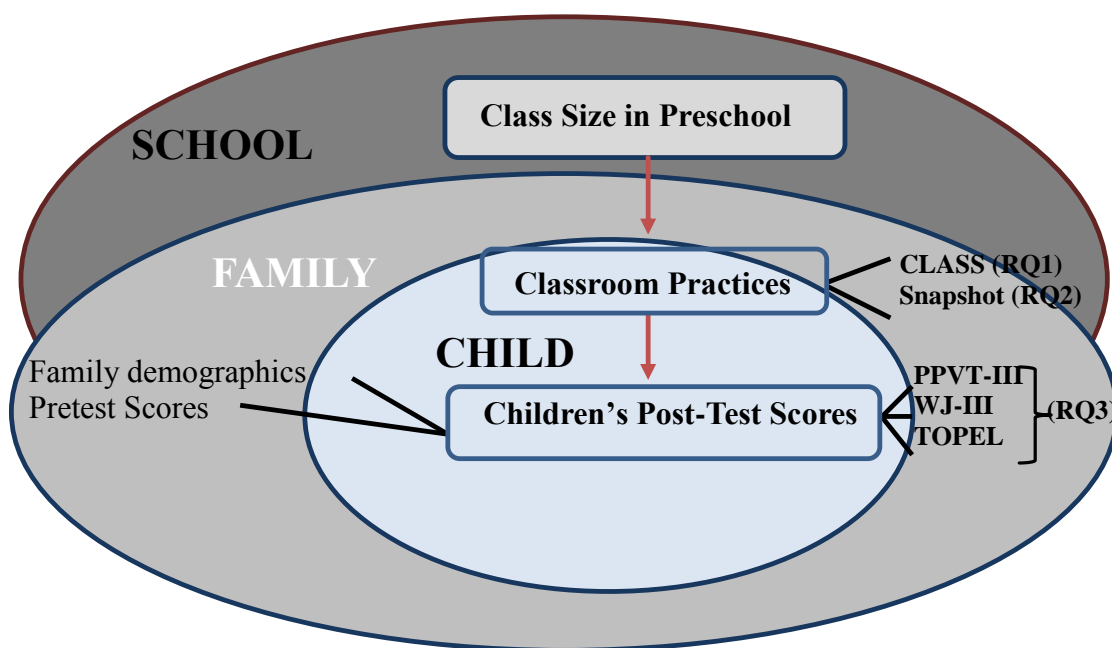
From this theoretical perspective, the child is placed in a circle in the center, and the family and school contexts are represented with two larger circles surrounding the child circle that intersect with one another.

The research framework encompassed by the larger circles shows that the class size that is implemented affects what children experience in the classroom, which affects their cognitive development (Mosteller, 1995; Word et al, 1990; Finn & Achilles, 1990 Nye, Hedges, and Konstopoulos, 2001, 2002; Molnar et al, 1999). This is based on the

rationale that different class sizes allow for differences in the interactions, activities, and content that occurs among teachers and peers in the classroom (Achilles et al, 1995; Phillipsen et al, 1997). As such, the research framework diagram shows class size leading to classroom practices, which leads to children's post-test scores.

Further, the child's circle encompasses child scores and classroom practices, which represents the notion that not only do classroom practices impact children, but children also impact what happens in the classroom. This is partially a product of their family characteristics and cognitive levels (pre-test scores) at the start of the school year. This entire framework sets the foundation for the method that will be presented next in Chapter 3.

CHAPTER 3

METHOD

Study Design

The current study will use quantitative methods to examine relationships between class size and teachers' practices and children's experiences and outcomes. The data for this study comes from a large 3-year examination conducted by the National Institute for Early Education Research (NIEER) at Rutgers University in the Chicago Public Schools. The purpose of the larger study was to understand the effects of three preschool inputs (class size, parental engagement, and a professional development initiative) on teachers' classroom practices and children's outcomes over time, ultimately for the sake of adopting policies that benefit children's learning and development. The current study addresses only one of these inputs, namely class size, by drawing upon data that was collected during the 2008-2009 school year.

In 2008, NIEER partnered with Metro Chicago Information Center (MCIC), a Chicago-based research institute, to accomplish the data collection for this study. A Chicago-based NIEER project coordinator worked closely with MCIC staff to carry out the study. A variety of data was collected during the 2008-2009 school year, only a portion of which will be utilized for the current study. Specifically, classroom observations were conducted at one time point during the year to provide a measure of teachers' emotional, organizational, and instructional interactions with children. In

addition, children's vocabulary, early literacy, and early math skills were assessed at the start and end of the school year, to provide a measure of their gains in these areas. A team of researchers hired by MCIC was trained and tested for inter-rater reliability on each measure prior to data collection.

All teachers who became part of the larger study in 2008-2009 were included in the current study. Only one input was implemented during this time, so the other two inputs do not pose the risk of having an impact. The sample includes only Preschool for All (PFA) classrooms, the state-funded preschool program in Chicago.

Research Questions

Three research questions guide the present study:

(1) Does preschool classroom quality vary by class size?

(2) Does the quantity of time children spend in particular types of preschool activities and interactions vary by class size? (See Table 4 for descriptions of activities and interactions).

(3) Do preschool children's cognitive outcomes vary by class size?

Participants

Schools

As mentioned, schools in the current study were chosen from the Chicago Public school system. In 2008-09, the first year of the study, the Chicago Public Schools (CPS) Office of Early Childhood Education served approximately 31,000 preschool children in over 600 classrooms. Eighty four percent of the city's students qualified for free or

reduced lunch, and the ethnic breakdown was 51 percent African-American, 38 percent

Hispanic, 8 percent Caucasian, and 3 percent Asian.

The initial study design stipulated that 10 schools would be randomly chosen for

participation in the study. However, the sample ended up being more purposefully

selected due to the small number of schools eligible to participate. Schools that did not

include preschool or were already participating in studies were removed from the sample

pool. In addition, schools where there was some type of upheaval (i.e. no principal, at risk

of being closed, etc.) were eliminated. From the schools that remained, the following

criteria were required:

(1) Schools must serve high-poverty (defined as free and reduced lunch rates greater than

85 percent) and working class families (defined as free and reduced lunch rates of 50 to

85 percent).

(2) Schools must have preschool populations that mainly speak English and Spanish as

their home language (due to the lack of standardized child assessment measures in

languages other than English and Spanish).

(3) Each school must have at least two preschool classrooms that utilize AM and PM

half-day (2.5 hour) schedules.

The first ten schools that were identified as meeting the criteria outlined above were

included in the current study.

Teachers/Classrooms

As mentioned, one requirement for participation in the study was that there were

at least 2 preschool classrooms in the school, each with an AM and a PM session. The

reason for this was so that one session could be capped at 15 students, and the other

session would remain at the regular CPS class size of 20 children. This would allow for

each teacher to have one reduced and one regular size class. Of the 10 schools that were

selected for participation in the study, 9 schools included 2 preschool classrooms and one

school included 4 preschool classrooms. This resulted in a total of 22 classrooms (44

sessions) to be included in the study. Within each classroom, the following criteria were

required:

(1) Classrooms cannot be blended, meaning that students are considered to be general

education.

(2) Teachers must implement the standard CPS pre-K program, which is the *Creative*

Curriculum.

(3) Teachers must have a minimum of a Bachelor's degree and Illinois Type 4 early

childhood certification.

(4) Each lead teacher must have an assistant teacher in the classroom.

(5) Each teacher must receive the same CPS-provided teacher professional development

and abide by CPS program standards.

As mentioned, class size was capped at 15 children in one session for each teacher

that participated in the study. The session that was capped was randomly assigned to be

either the AM or PM session. Initially, 11 AM sessions and 11 PM sessions were

assigned to be reduced in size. However, 6 teachers switched their initially assigned

reduced session from AM to PM, which resulted in a total of 5 reduced size AM sessions

and 17 reduced size PM sessions. It was unclear as to why teachers refused to cap their

AM sessions in several instances.

Children

Parents of children attending participating classrooms were informed of the study

at the start of the school year. At that time, teachers assisted researchers in collecting

parental consent forms at the school during drop-off and pick-up hours. Consent forms

were collected from the majority of parents in participating classrooms, and 12 children

were randomly selected from those that were returned. Several sessions did not have as

many as 12 consent forms returned, therefore children could not be randomly selected

and any children with returned consent forms were included. No consent forms were

collected from two PM sessions in the sample, which resulted in a final sample of

children from 42 classroom sessions. In effect, these classrooms will be included in

analyses of classroom practices, but not in analyses of child outcomes since there is no

match in one of the conditions.

The total sample in fall 2008 consisted of 414 children (223 in AM sessions; 191

in PM sessions), with an average of 9.41 children per session, ranging from 2 to 12. 188

children (39 AM; 149 PM) were in reduced class sizes and 226 children (184 AM; 42

PM) were in regular class sizes. The sample was 48.8% female and the ethnic breakdown

was 58.5% Hispanic, 38.6% Black, 1.4% White, 0.5% Asian, and 1.0% other. In addition,

51.4% spoke English as their primary home language, and the remaining percentage

spoke either primarily Spanish, both English and Spanish, or some other language at

home. The average age in months was 52.8. Comparisons between treatment and control

groups at baseline are presented in Table 1 below. Differences between groups were

analyzed and are reported later in the *Results* section in Chapter 4.

Table 1. Regular v. Reduced Class Size Sample Characteristics

	Full Sample (n=414)	Regular Class Size (n=226)	Reduced Class Size (n=188)
% Female	48.8	49.6	47.9
Ethnicity			
Asian	0.5	.9	0
Black	38.6	38.5	38.8
Hispanic	58.5	58.4	58.5
White	1.4	1.8	1.1
Other	1.0	.4	1.6
% Primary English speakers	51.4	49.6	53.7
Age in mos.	52.8	53.2	52.3
% in AM sessions	53.9	81.4	20.7
Income			
<15K	40.1	41.6	38.3
15K-25K	23.7	23.9	23.4
>25K-50K	14.0	14.2	13.8
>50K-75K	1.7	2.2	1.1
>75K	1.0	0	2.1
Missing	19.6	18.1	21.3
Maternal Education			
Less than HS	30.9	35.8	25.0
HS diploma	27.1	23.0	31.9
Some college	17.9	15.5	20.7
BA or more	8.5	9.7	6.9
Missing	15.7	15.9	15.4

All children were tracked over the course of the year for follow-up assessments in

the spring of 2009. 83 children were added to the sample at time 2 testing, but they were

not included in the current study due to not having pretest data. One classroom

(containing 24 children in the study) discontinued participation at follow-up and 36

children transferred to new schools. This resulted in a sample of 354 children from 40

sessions in 20 classrooms that were assessed in both fall 2008 (time 1) and spring 2009 (time 2). Attrition across regular and reduced class size groups is documented in Table 2 below.

Table 2: Fluctuation in Sample Size for Children in Reduced and Regular Size
   Classrooms at Time 1 & Time 2 of Assessments

	Reduced	Regular	Total
Total Assessed at Time 1	188	226	414
Added to sample @ T2	(59)	(24)	(83)
Child Transferred	-15	-21	-36
Classroom Closed	-12	-12	-24
Total Assessed at Time 2	161	193	354

Child Assessment Instruments and Procedures

To understand the effects of class size on children's development, children's receptive vocabulary, emergent literacy, and early math skills were assessed at the beginning (fall) and end (spring) of the preschool year. Fall assessments served as the pre-test and spring assessments served as the post-test. Bilingual (English and Spanish) and Spanish-speakers were assessed in both English and Spanish, with primary language ascertained from the classroom teacher. Assessments were conducted one-on-one in the child's school and were scheduled to avoid meals, nap, and outdoor play times. Each assessment lasted approximately 30-minutes. The following assessment measures were used.

Peabody Picture Vocabulary Test – Third Edition (PPVT-III) (Dunn & Dunn, 1997)

This is a 204-item test of receptive vocabulary in standard English. The *Test de Vocabulario en Imagenes Peabody (TVIP)* (Dunn, Lugo, Padilla, & Dunn, 1986) uses

125 translated items from the PPVT to assess receptive vocabulary acquisition of Spanish-speaking and bilingual students. The PPVT is predictive of general cognitive abilities and is a direct measure of vocabulary size. The rank order of item difficulties is highly correlated with the frequency with which words are used in spoken and written language. The test is adaptive (to avoid floor and ceiling problems), establishing a floor below which the child is assumed to know all the answers and a ceiling above which the child is assumed to know none of the answers. Reliability is good as judged by either split-half reliabilities or test-retest reliabilities. The TVIP measures growth in Spanish vocabulary for bilingual students and for monolingual Spanish speakers. The results of these tests are found to be strongly correlated to school success.

Test of Preschool Emergent Literacy (TOPEL) (Lonigan, et al., 2007)

The TOPEL was used with English-speakers and the Spanish version of the *Preschool Comprehensive Test of Phonological & Print Processing (Pre-CTOPPP)* (Lonigan, Wagner, Torgeson, & Rashotte, 2002) was used with Spanish-speakers because the assessments are identical and there is not yet a published Spanish version of the TOPEL. Both the Pre-CTOPPP and the TOPEL have three subtests, but only *Print Knowledge* subtest 1 was used in the current study. This subtest has 36 items and measures alphabet knowledge and early knowledge about written language conventions and form. The child is asked to identify letters and written words, point to specific letters, name specific letters, identify letters associated with specific sounds, and say sounds associated with specific letters.

Woodcock-Johnson Psycho-Educational Battery-Third Edition (WJ-III)

The Woodcock-Johnson and the *Bateria Psico-Educativa Revisada de Woodcock-Muñoz (WM-R)* (Woodcock & Johnson, 2001; Woodcock & Munoz-Sandoval, 2005) were used to assess children's early math skills. This assessment includes multiple subtests, but only the *Applied Problems* subtest was used in the current study. The English form of the subtests was normed on a stratified random sample of 6,359 English-speaking subjects in the United States. The Spanish form was normed on 3,911 primarily monolingual Spanish-speaking subjects from samples obtained both inside and outside the United States. Internal consistency reliabilities range from the high .70s to low .90s on both subtests for preschool-aged children. Correlations of the WJ-R and WM-R with other tests of cognitive ability and achievement are reported to range from .60 to .70.

All children were assessed first with the PPVT, regardless of home language, to get a sense of receptive vocabulary skills in English. The rationale for this was that instruction was taught primarily in English in these preschool classrooms, so it would be reasonable to assess gains in English over the course of the school year. Children who were identified as having Spanish or both English and Spanish as their primary home language were assessed with both the PPVT and TVIP, to get a sense of their vocabulary skills in both languages. There were a handful of children that could not be assessed with the PPVT because their English skills were not sufficient. The remaining 2 assessment measures (TOPEL and WJ) were administered in *either* English or Spanish, depending upon what the teacher identified as the child's best testing language. Ideally, all measures

would be administered in both languages for children identified as Spanish-speaking or bilingual but this was not feasible due to budgetary and time constraints.

Standard scores were calculated for the PPVT and TVIP, and whichever one the child scored best on was used for analyses. At fall pre-testing, 216 monolingual English-speakers (also including children who spoke a primary language other than English or Spanish) were administered the PPVT only. There were also 5 bilingual students who were only administered the PPVT. In addition, 190 children were administered both the PPVT and TVIP. Of those children administered the tool in both languages, 158 had better standard scores on the Spanish version and 32 had better standard scores on the English version. Of the 158 children that had better standard scores on the Spanish TVIP, 107 were administered the Spanish version of final two assessments and 51 were administered the English version. Only 2 out of 32 children with better standard scores on the English PPVT were administered the Spanish versions of the final two measures. This inconsistency in the language used for administration was the result of teacher judgment and will be addressed in subsequent analyses.

A similar pattern of language used for assessments was found at the time of post-testing (spring 2009). Specifically, 179 students were administered the PPVT only (139 monolingual English speakers, 37 bilingual students, and 3 with a home language other than English or Spanish). The remaining 175 students were administered both the PPVT and TVIP. Of those that were administered the tool in both languages, 135 had better standard scores on the Spanish version and 41 had better standard scores on the English version. Of the 135 children that had better standard scores on the Spanish TVIP, 95 were

administered the Spanish version of final two assessments and 40 were administered the

English version. 14 out of 41 children with better standard scores on the English PPVT

were administered the Spanish versions of the final two measures.

Family Demographic Data

At the beginning of the school year, mothers of children that were selected for

participation in the study were asked to complete a brief survey regarding their

educational level, employment status, and income. Data was available for 333 children

out of 354 (94.1%) who were assessed at both pre- and post-test.

Classroom Observation Instruments and Procedures

Several classroom observation instruments were used in the larger study to assess

the effects of class size on teachers' practices and children's experiences in the

classrooms. Results from two of these instruments will be analyzed in the current study,

in an attempt to understand how class size impacts the quality of the interactions that

occur between teachers and children (research question #1), and the quantity of activities

and interactions that occur among children and teachers (research question #2).

Classroom Assessment Scoring System (CLASS) (Pianta, LaParo and Hamre, 2005)

The CLASS is an observational system that assesses classroom practices in

preschool by measuring the interactions between students and adults. Observations

consist of 4 to 5, 20-minute cycles, followed by 10-minute coding periods. Scores (codes)

are assigned during various classroom activities, and then averaged across all cycles for

an overall quality score. Interactions are measured through 10 different dimensions,

which are divided into 3 larger domains.

The *emotional support* domain is measured through the use of 4 dimensions: Positive Climate, Negative Climate, Teacher Sensitivity, and Regard for Student Perspectives. The CLASS also measures Classroom Organization through 3 dimensions: Productivity, Behavior Management, and Instructional Learning Formats; and *Instructional Support* through 3 dimensions: concept development, Quality of Feedback, and Language Modeling. Each scale uses a 7-point Likert-type scale, for which a score of 1 or 2 indicates low range quality and a score of 6 or 7 indicate high range quality. Each dimension and domain is assigned a score during each 20-minute cycle. The number of children and adults in the classroom are also recorded during each 20-minute cycle. Table 3 below presents descriptions of each CLASS dimension.

Table 3. Descriptions of CLASS Dimensions

Domain	Dimension	Description
Emotional Support	Positive Climate	Reflects the emotional connection between teachers and children and among children, and the warmth, respect, and enjoyment communicated by verbal and nonverbal interactions.
	Negative Climate	Reflects the overall level of expressed negativity in the classroom. The frequency, quality, and intensity of teacher and peer negativity are key to this dimension
	Teacher Sensitivity	Encompasses the teacher's awareness of and responsiveness to students' academic and emotional needs.
	Regard for Student Perspectives	Captures the degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view and encourage student responsibility and autonomy.
Classroom Organization	Behavior Management	Encompasses the teacher's ability to provide clear behavior expectations and use effective methods to prevent and redirect misbehavior.
	Productivity	Considers how well the teacher manages instructional time and routines and provides activities for students so that they have the

		opportunity to be involved in learning activities.
	Instructional Learning Formats	Focuses on the ways in which teachers maximize students' interest, engagement, and abilities to learn from lessons and activities.
Instructional Support	Concept Development	Measures the teacher's use of instructional discussions and activities to promote students' higher-order thinking skills and cognition and the teacher's focus on understanding rather than on rote instruction.
	Quality of Feedback	Assesses the degree to which the teacher provides feedback that expands learning and understanding and encourages continued participation.
	Language Modeling	Captures the effectiveness and amount of teacher's use of language-stimulation and language-facilitation techniques.

Emergent Academics Snapshot

Data was also collected on the amount of time children spent in various activities and interactions through classroom observations that were coded with the *Emergent Academics Snapshot* (Ritchie, Howes, Kraft-Sayre, & Weiser, 2001). Observations consist of time sampled codes assigned to teacher and child behaviors, every 60 seconds (representing one cycle) over the course of the day. Four children were randomly selected from each classroom and each child was observed for 40 seconds, followed by 20 seconds of coding, for as many cycles as could be completed during one classroom session. Data for the current study includes a total of 4,115 observation cycles from all 44 classroom sessions.

Codes were divided into 5 subscales, including activity setting (i.e. whole group, free choice, transitions), peer interaction (simple social, cooperative pretend), child engagement (i.e. science, math, oral language development), teacher-child engagement (i.e. scaffolds, didactic), one-on-one teacher-child interactions (elaborated, routine).

Because children were randomly selected, the way that their time is spent is meant to

reflect the average experience of the children in the classroom. Table 4 below presents

the descriptions of the codes that were analyzed in the current study.

Table 4. Snapshot Code Descriptions

Subscale	Code	Description
Activity	Basics	Toileting, standing in line, clean-up time, wait time between activities, waiting for materials to be passed out, transitional activities, i.e. moving out of whole group into the next activity.
	Meals/Snacks	Eating lunch, breakfast or snacks, or enjoying food that the class cooked during a cooking project.
	Whole Group	Child is engaged with the whole group in a teacher-initiated activity. The child's focus is on the teacher.
	Free Choice/Center	Child is engaged in free choice activities. During this time children are able to select what and where they would like to play or learn. It does not matter if the activity they have chosen is individual or in a small group. It does not matter if the activity is with or without the teacher.
	Individual Time	Child has been assigned to work individually with or without teachers, on worksheets, independent projects, computer work etc. This is coded when this is the activity setting for the whole class or for a small group in which the target child is involved
	Small Group	Child is engaged in small group activities that are teacher organized.
Peer Interaction	Solitary	Child is playing productively alone, with no eye gaze or mutual interest in objects with any peer. Do not code if the child is playing with an adult.
	Parallel Aware	Child and at least one peer are playing with the same type of objects and are within three feet of each other and have mutual awareness of one another.
	Simple Social	Child and at least one peer are engaged in play with social interaction. This type of play may be

		either verbal, or non-verbal (physical gestures or facial expressions).
	Complementary Reciprocal	Child and at least one peer engage in social play with turn taking structure and role reversal.
	Cooperative Pretend	Child and at least one peer are engaged in (at least) simple-social play that includes a script
	Complex Pretend	Child and at least one peer are engaged in: "cooperative pretend" play plus the children name roles.
Child Engagement	Read-to	Child is being read to by an adult.
	Preread/Read	Child is reading on her/his own or with peers, listening to a book on tape while looking at a book, involved in a sequencing activity, or involved in recognition of whole words.
	Letter/Sound	Child is practicing rhymes that help her/him recognize sounds, talking about sound-letter relationships, identifying letters, sounding out words or practicing vowel sounds.
	Oral Language Development	Child is involved in an activity or an interaction where a teacher is taking action to draw communication from the children to build expressive language or is actively listening to children speak, by allowing them to complete their thoughts.
	Computer	Child is using the computer for something other than writing.
	Writing	Child is writing, pretending to write, or using a keyboard/computer specifically for writing.
	Math	child is counting, identifying written numerals, matching numbers to pictures, making graphs, playing counting games, keeping track of how many days until a special event, etc.
	Science	Child is identifying and exploring natural phenomena in their environment, using science equipment, working with sand or water, or engaged in trial and error/experimentation.
	Social Studies	Child is talking, reading, or engaged in activities about their world (e.g. their neighborhood, their school, the farm, the community workers).
	Aesthetics	Child is engaged in art or music activities.
	Gross Motor	
	Fine Motor	Child is utilizing pincer grasp (e.g. stringing beads, building with Legos, cutting, using crayons and markers or paint brushes, pencils or pens, etc.)
Adult Interaction	No Response	Teacher does not respond to child's verbal or physical bid.

	Routine	Teacher interacts with target child during routine caregiving (i.e. passes out materials) but does not verbally interact with the child.
	Minimal	Teacher verbally responds to target child with a few words.
	Simple	Teacher responds to target child with short sentences.
	Elaborated	Teacher engages in physical contact (high fives, hugs or holds child), engages in reciprocal conversation that validates a child's feelings or demonstrates teacher interest in what the child is saying.
Teacher-Child Engagement	Literate	Target child is engaged in read-to or letter/sound engagement with teacher involved.
	Scaffolds	Teacher shows an awareness of an individual child's needs and responds in a manner that supports and expands the child's learning.
	Didactic	Teacher engages child in rote teaching (i.e. teacher gives instructions, models, demonstrates).
	Second Language	Teacher is speaking a language other than English.
	Facilitate Peer	Teacher attempts to facilitate child's peer interactions.

Data collector training and reliability. All of the assessments and observations previously described were conducted by data collectors trained by the NIEER Project Coordinator and hired by MCIC. Data collectors were required to have at least a B.A. in education or psychology. Bilingual data collectors were hired to conduct the necessary child assessments in Spanish.

Child assessors were trained by NIEER's Project Coordinator on each child assessment in September 2008, just prior to the start of data collection. The training lasted 2 days, and the second day each data collector was shadow scored to ensure they could conduct the assessment with 100 percent accuracy. The Project Coordinator conducted quality checks of incoming data throughout the data collection period, and

provided feedback to data collectors as necessary. A one-day refresher training on all assessment measures took place in late spring, just prior to the final round of child assessments. Training was used only as a refresher, since all data collectors were returning from the fall, and were therefore already trained to reliability.

Classroom observers were trained in January 2009, just prior to conducting observations. Training in administering the observation protocol (CLASS) was provided by the NIEER Project Coordinator, who was previously trained by the authors of both the CLASS and the *Emergent Academics Snapshot* was developed. Training on the CLASS lasted 3-days, and the third day was used to ensure reliability through the use of video observations. Each observer was required to observe and score 5 videos in 80% agreement with pre-assigned master scores. If 80% reliability was not achieved after 3 attempts, an observer was not included in data collection. Training on the Snapshot lasted 3 days, after which each observer was required to observe and score 80-minutes of video footage in 75% agreement with pre-assigned Snapshot master codes on video. The CLASS and Emergent Academics Snapshot were used simultaneously in the classroom during structured observations. All classrooms were observed over the course of 3-months. The NIEER project coordinator shadowed each observer after every 10 observations, to minimize scoring drift.

Scores from all child assessments and classroom observations were entered into SPSS for data analyses. Standard scores were calculated, based on each child's age in months, for assessments that allowed for standardization. Average scores for each

CLASS dimension and domain were also entered, as well as the frequency of Snapshot

codes that were collected for each classroom.

Analysis Strategy

Each question was answered with the specific analysis methods detailed below.

(1) Does preschool classroom quality vary by class size?

In order to answer the first question a dummy variable for class size (0 = regular; 1 =

reduced) was entered. This variable served as the independent variable in paired samples

t-test comparisons, since the sample of teachers having reduced sized classes is the same

as the sample having regular size classes. Dependent variables include overall CLASS

dimension and domain average scores. The claim that average CLASS scores for reduced

classes are equal to average CLASS scores for regular classes was tested using the

following null hypothesis:

H_o: $X_{reduced} = X_{regular}$, and X represents the mean CLASS scores for each group.

To test whether there were differences in CLASS scores by class size, the following
model was used:

$$t = \frac{\Sigma d}{\sqrt{\frac{n(\Sigma d^2) - (\Sigma d)^2}{n-1}}}$$

The top of the formula is the sum of the differences (i.e. the sum of d). The bottom of the

formula reads as: The square root of the following: n times the sum of the differences

squared minus the sum of the squared differences, all over n-1.

(1) The sum of the squared differences: Σd^2 indicates to take each difference in turn,

square it, and add up all those squared numbers.

(2) The sum of the differences squared: $(\sum d)^2$ means add up all the differences and square the result.

The paired t-test was calculated to take into account the fact that pairs of subjects (in this case, two classes for each teacher) go together. It is based on the differences between the CLASS scores for each pair of classrooms - that is one subtracted from the other. This difference is notated as d. The formula for the paired t-test uses just d and n (the number of values in the data), and nothing else. The way these two values affect the value of t are as follows:

(1) As the average of the differences gets bigger, t gets bigger;

(2) As the variation in the differences gets bigger, t gets smaller;

(3) As the number of values gets bigger, t gets bigger.

The mean difference is what is being tested against 0. If the value of t that is obtained for a given score exceeds the critical value of t for that score (and sample size), then the null hypothesis can be rejected because the difference in group scores is larger than that due to chance alone, and it can be said that classroom quality varies by class size. Results from analyses are presented in Table 5 in the subsequent Chapter.

(2) Does the quantity of time children spend in particular types of preschool activities and interactions vary by class size? (See Table 4 for descriptions of activities and interactions).

In order to answer the second question, overall percentages of time were first calculated for each code identified in the Emergent Academics Snapshot. Percentages of time were then converted into minutes, based on a 2.5-hour preschool day. A dummy

variable for class size (0 = regular; 1 = reduced) was then entered. Again, this served as the independent variable in a paired samples t-test comparison. Dependent variables include the amount of time spent in each type of Snapshot coded activity and interaction. The claim that the quantity of time spent in each Snapshot code is equal for reduced and regular classes was then tested using the following null hypothesis: H_o: $X_{reduced} = X_{regular}$, and X represents the quantity of time (in minutes) for each Snapshot code.

In order to determine whether there were statistically significant differences between reduced and regular size classes in terms of the amount of time children spent in various types of activities and interactions, the t equation indicated above was used again. In this case, however, the differences (d) are defined as the differences in the amount of time (in minutes) spent in activities. For each snapshot code, if the value of t that was obtained for a given percentage exceeds the critical value of t for that score (and sample size), then the null hypothesis can be rejected because the difference in group amounts of time is larger than that due to chance alone, and it can be said that the amount of time spent in different activities/interactions vary by class size. Results from paired t-test analyses are presented in Table 6 in Chapter 4.

Since paired t-tests do not account for differences between groups that could potentially inflate or mask the impact of class size, chi square tests were used as a secondary analysis to gauge for differences between groups in terms of children's age, home language, gender, ethnicity, household income, and maternal education level. Any of these variables that differed significantly between groups were controlled for in a set of regression analyses. Along with significantly different child-level characteristics

between groups, these regression analyses also controlled for whether or not teachers switched their assigned reduced class size session from that which was randomly assigned. Since teacher assignment was an important component of the study design, teachers that switched present a concern for the overall findings and it is important to consider any potential effects of switching.

(3) Do preschool children's cognitive outcomes vary by class size?

The final research question addresses the effect of class size on children's cognitive outcomes. The claim that cognitive outcome scores do not vary by class size was tested with the following null hypothesis: H_o: $B=0$, which says that outcome scores are not associated with class size. Data were collected at both the child and teacher level, with multiple children having the same teacher. Therefore, a multi-level regression model was used, which allows one to estimate the association between variables assessed at different levels by taking into account that all children with a given teacher have shared a similar preschool experience at the teacher level.

Within each teacher, children experienced either a regular or reduced class size, which served as a level-2 predictor in the model. Data was collected from J teachers, with a different number of pupils N_j for each teacher. On the pupil level, we have the dependent variable Y_{ij} (outcome score) and a set of explanatory variables X_{ij} (e.g., pupil income level, pretest score, gender, ethnicity). The following regression equation was used to predict the dependent variable Y from the explanatory variable X: $Y_{ij} = B_{0j} + B_{1j}P_{ij} + B_{2j}X_{ij} + e_{ij}$.

In this equation, Y_{ij} are the scores of pupil I with teacher j, P_{ij} is the pretest score of pupil I with teacher j, and X_{ij} are the explanatory characteristics of pupil i (i=1…j) with teacher j (j=1…j). There are several explanatory pupil characteristics that will each have their own coefficient results; they are lumped together to simplify the equation. This model specifies that the different teachers are characterized by different regression equations; each teacher has its own intercept B_{0j} and slopes B_{ij} and B_{2j}. Essentially, this means that the relationship between class size as well as any other moderator variables in the model, and child outcome scores can vary for each teacher. Because there is just one teacher-level explanatory variable Z_j (reduced class size), the model for the Bs becomes:

(1) $B_{0j} = B_{00} + B_{01}Z_j + u_{0j}$,

(2) $B_{1j} = B_{10} + B_{11}Z_j + u_{1j}$,

(3) $B_{2j} = B_{20} + B_{21}Z_j + u_{2j}$.

In the first of the equations listed above, B_{00} and B_{01} are the intercept and slope of the regression equation used to predict B_{0j} from Z_j, and u_{0j} is the residual error term in the equation for B_{0j}. Thus, if B_{01} is positive and significant, the null hypothesis can be rejected and it can be concluded that outcome scores in reduced size classes are higher than in regular size classes.

In the second of the equations listed above, B_{10} and B_{11} are the intercept and slope to predict B_{1j} from Z_j, and u_{1j} is the residual error term in the equation for B_{1j}. Thus, if B_{11} is positive and significant, it can be concluded that the effect of pretest scores is stronger in reduced class sizes. In other words, the pretest score would act as a moderator

variable for the relationship between class size and outcome scores; this relationship varies according to the value of the moderator value.

In the third of the equations listed above, B_{20} and B_{21} are the intercept and slope to predict B_{2j} from Z_j, and u_{2j} is the residual error term in the equation for B_{2j}. Thus, if B_{21} is positive and significant, it can be concluded that the effect of pupil characteristics (e.g. income, pretest, gender, ethnicity, etc.) is stronger in reduced class sizes. In other words, the pupil characteristics would act as moderator variables for the relationship between class size and outcome scores; this relationship varies according to the value of the moderator value.

The equation below shows the model as one single equation after substituting (and rearranging) the teacher-level equations into the pupil-level equation:

$$Y_{ij} = B_{00} + B_{01}Z_j + B_{10}\,P_{ij} + B_{11}Z_j\,P_{ij} + B_{20}\,X_{ij} + B_{21}Z_j\,X_{ij} + u_{0j} + u_{1j}\,P_{ij} + u_{2j}\,X_{ij} + e_{ij}.$$

This equation allows for a better visual of the interaction between class size and pupil characteristics as well as the interaction between class size and pretest score. The importance of these interactions was explained in the preceding paragraphs in terms of the potential for these moderating variables to affect the relationship between class size and child outcome scores. As is also shown in this equation, the regression coefficients no longer carry a subscript j for teachers; in the combined equation they refer to the average value of the regression across all teachers. Results from these analyses are presented in Table 10 in the next Chapter.

CHAPTER 4

RESULTS

Prior to addressing the three primary research questions for this study, the average

numbers of children in regular versus reduced class sizes were calculated. The purpose

for this was to get a handle on the actual numbers of children in attendance, since

subsequent analyses will focus on assigned class sizes. In other words, this study seeks to

examine effects on classrooms and children that exist due to a class size assignment of 20

versus a class size assignment of 15.   Based on the attendance data collected at the time

of each classroom observation, it was found that the average class size for a "reduced"

class was 12.61 students. The average class size for a "regular" class size was 16.23

students. Therefore, typical daily class sizes were indeed smaller, on average, than the

assigned enrollment, and the difference in class size was just about 3.5 students on

average compared to a difference of 5 in enrollment as assigned. However, the

attendance rates" are similar 84% in the smaller classes and 81% in the larger classes.

The difference of 3.6 students per preschool classroom is still substantial and is nearly as

large a percentage difference from actual class size as that between 15 and 20. It is

important to keep these actual class sizes in mind when interpreting the results.

Classroom Quality

The Classroom Assessment Scoring System (CLASS) provided ratings, from 1 to

7, of the interactions that occurred between teachers and children. Each score is meant to

reflect the average experience of the children in the classroom. Overall, classrooms in

85

both groups scored nearly a full standard deviation higher on instructional support and about half a standard deviation lower on emotional support than the average in a large study of pre-K quality in 11 states (Mashburn et al, 2008). This might have had something to do with the actual sizes of classes that were found, which were quite small and might allow for higher quality interactions in the area of Instructional Support. However, no statistically significant differences were found between smaller and larger classes as determined by paired-sample t-tests conducted on all dimensions of the CLASS and on the overall score. Results are presented in Table 5 below. On average, mean scores on measures of classroom quality were nearly identical for regular and reduced size classes. The largest difference measured was just 0.25 (effect size = mean difference/control group standard deviation), which favored the regular class size group.

Overall, the average level of quality in this sample of classrooms varied depending upon the dimension that was being observed. Average scores were highest on the negative climate dimension, indicating that teachers expressed very little negativity in the classroom. These scores were considered high quality, indicating very little room for improvement in either regular or reduced size classes. No other dimensions scored in the high quality range and the next highest scores were for positive climate and behavior management, both scoring at the high end of the mid-range level of quality. There is room for improvement, but teachers were found to be displaying consistently positive language and communication and behavioral strategies with children. This was the case for both regular and reduced size classes.

Table 5. Average CLASS Scores for Regular v. Reduced Class Sizes

	Regular Class Size Mean Score (N=22) (s.d.)	Reduced Class Size Mean Score (N=22) (s.d.)	t	p-value
Positive Climate	5.39 (1.09)	5.16 (1.00)	0.83	0.42
Negative Climate (inverted)	6.63 (0.56)	6.69 (0.37)	0.74	0.47
Teacher Sensitivity	4.58 (1.42)	4.23 (1.47)	0.76	0.45
Regard for Student Perspectives	4.84 (1.23)	4.76 (1.01)	0.25	0.81
Behavior Management	5.43 (1.14)	5.36 (1.32)	0.25	0.81
Productivity	5.14 (1.25)	4.87 (1.18)	0.71	0.49
Instructional Learning Formats	4.56 (1.16)	4.46 (1.21)	0.32	0.75
Concept Development	2.80 (1.30)	2.59 (1.18)	0.55	0.59
Quality of Feedback	2.72 (0.85)	2.85 (0.95)	-0.53	0.60
Language Modeling	2.96 (1.14)	2.79 (1.11)	0.48	0.64
CLASS overall average score	4.50 (0.67)	4.38 (0.76)	0.62	0.54
Emotional Support Domain	5.36 (0.85)	5.21 (0.80)	0.64	0.53
Classroom Organization Domain	5.04 (0.98)	4.90 (1.13)	0.49	0.63
Instructional Support Domain	2.83 (0.92)	2.74 (0.96)	0.30	0.77

Mid-range scores were also found for the *teacher sensitivity, regard for student perspectives, productivity*, and *instructional learning formats* dimensions. This indicates that positive behaviors in each of these areas occurred, but they were not always consistent and at times infrequent during the school day.

Low-range scores were found for all of the instructional support domains: *concept development, quality of feedback*, and *language modeling*. Low range scores are

characterized by rarely occurring behaviors in each of these areas. The lowest scores were found in the area of *concept development*, which translates to a lack of instructional discussions and activities used by teachers to promote students' higher-order thinking skills.

Quantity of Time Spent in Activities and Interactions

Results from Snapshot analyses indicated more pronounced differences between small and large class size groups on this measure. For these analyses all interactions and activities that accounted for less than 1% of the time were deleted from analyses. All remaining percentages were converted into minutes based on a 2.5-hour program total time, to provide an estimate of the number of minutes that children spent in each type of activity and interaction. Table 6 below presents means (in minutes), ranges, and standard deviations by group for each activity coded from the Snapshot.

Table 6. Time Spent in *Emergent Academics Snapshot* Activities and Interactions

Subscale and Code	Regular Class size Minutes Spent (N=22)		Reduced Class Size Minutes Spent (N=22)	
	Mean	Std. Dev	Mean	Std. Dev
Activity				
Basics	28.36	16.43	27.82	14.68
Meals	8.68	13.98	6.41	12.25
Whole Group	52.75	20.41	53.52	23.10
Free Choice	42.04	20.99	40.04	19.04
Individual	3.87	4.91	6.61	11.35
Small Group	13.08	16.43	14.17	14.54
Peer Interactions				
Solitary	4.92	5.13	7.70	10.23
Parallel Aware	99.47	26.58	88.51	28.92
Simple Social	34.77	20.98	40.21	22.78
Complimentary Reciprocal	4.33	6.08	3.73	4.40
Cooperative Pretend	2.43	4.04	3.63	6.11
Child Engagement				

Read to	12.48	6.46	14.28	9.97
Pre-read	13.64	11.07	15.80	11.19
Letter-Sound	8.00	5.32	9.17	9.08
Oral Language Development	4.62	8.07	6.33	5.98
Writing	6.32	6.69	5.22	5.57
Math	10.60	6.14	10.27	7.18
Science	15.07	14.76	17.67	12.41
Social Studies	27.36	18.33	24.50	10.02
Aesthetics	25.10	12.48	25.45	19.37
Fine Motor	14.13	7.85	13.53	9.57
Computer	3.79	5.56	4.94	6.38
No Child Engagement	52.99	16.68	56.70	40.33
Teacher Child Engagement				
Literate	17.95	7.74	21.67	14.39
Scaffold	9.75	11.98	13.38	13.27
Didactic	25.76	19.10	37.94+	22.03
Second Language	16.58	22.34	17.41	19.16
Facilitate Peer Interaction	2.99	5.02	1.91	2.83
No Teacher-Child Engagement	97.22	20.75	90.16	23.01
Adult 1-1 Interaction				
Routine	3.89	4.00	4.36	5.99
Minimal	7.36	5.20	6.22	4.03
Simple	4.78	3.94	5.67	3.54
Elaborated	3.23	3.08	6.29+	5.92
No Adult 1-1 Interaction	130.76	8.08	127.69	8.84

+$p<.10$

Teacher-Child Interactions

The Snapshot teacher-child interaction subscale provides information on the types and quantity of interactions that the teacher has with one or more children in the classroom. Multiple codes can be assigned within a given cycle. In a 150 minute session, children spent on average nearly 100-minutes *not* engaged in any type of interaction with teachers, regardless of class size. No statistically significant differences were found between groups on any measure of average time spent in specific activities at the conventional .05 level. However, in two instances there was suggestive evidence at .10

level: didactic interactions approached significance in favor of reduced class sizes as did

elaborated interactions (discussed later on below).

The difference on didactic interactions suggests that in general there was more

communication between teachers and children in reduced class sizes, in the form of

teachers giving information or instructions, or clarifying specific things. Figure 1 below

displays the time spent in teacher-child interactions for children in regular versus reduced

size classrooms.

Figure 2. Time Spent in Teacher-Child Interactions for Regular v. Reduced
Class Sizes. This figure illustrates differences between class sizes in the
time children spent in classroom activities and interactions.

Child Engagement

Child engagement codes capture children's engagement in learning activities. The

target child can be passively or actively engaged in all codes with the exception of gross

and fine motor. As a result, multiple codes can be assigned within a given cycle. No

statistically significant differences were found. Overall, time spent in no type of child

engagement was greater than any particular child engagement code. This indicates that there is room for improvement in terms of teachers' initiating and sustaining children's classroom engagement, regardless of class size. Figure 2 below displays the time spent engaged in different content for children in regular versus reduced size classrooms.

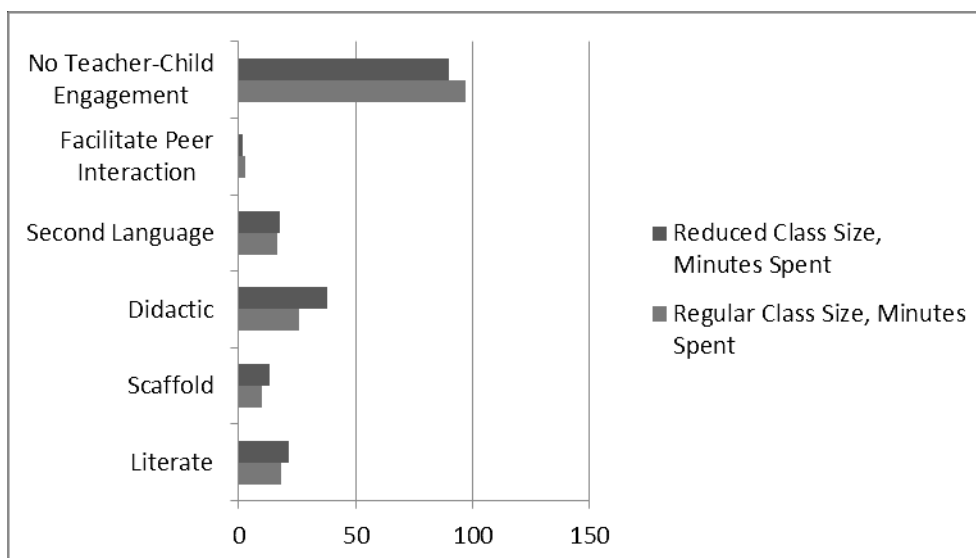Figure 3. Time Spent in Child Engagement Content for Regular v. Reduced Class Sizes. This figure illustrates differences between class sizes in the time children spent engaged in various content.

Adult Interaction

This section reflects the level of complexity of the teacher's 1:1 interaction with the target child. Only one code can be assigned in a given cycle, and in cases where multiple 1-1 interactions occurred in a given cycle, only the most complex interaction was coded. Differences in elaborated interactions between groups approached statistical significance at the .05 level (p=.052). This suggests that there were more 1-1 extended back and forth conversations between teachers and children in smaller classes. There is much room for improvement in terms of the amount of 1-1 interactions, regardless of class size. For instance, on average over 100-minutes of classroom time was

characterized by no 1-1 teacher-child interactions. Figure 3 below displays the time

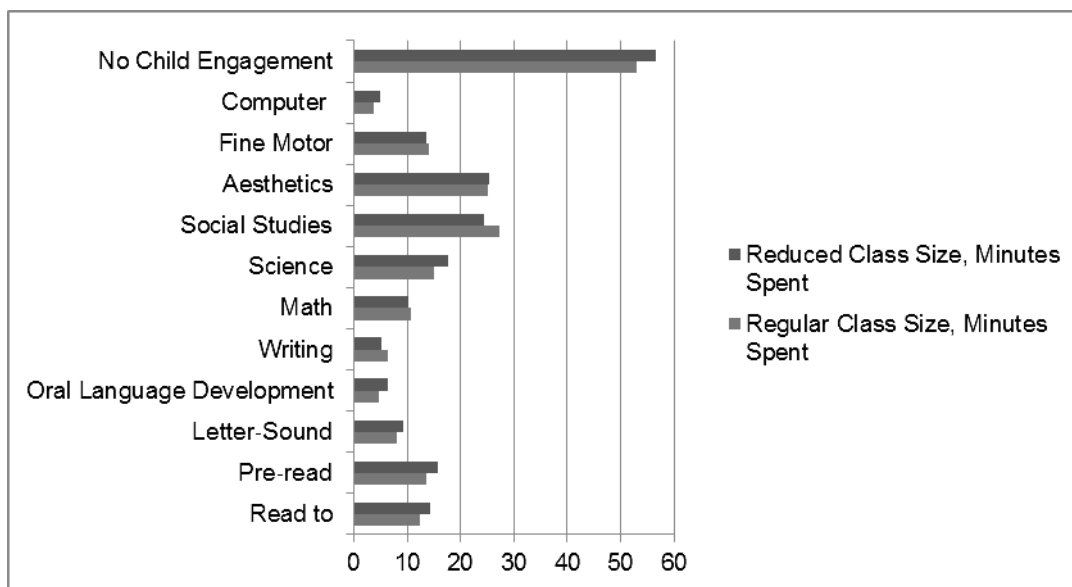children spent engaged in 1-1 interactions with adults for regular versus reduced size

classrooms.

Figure 4. Time Spent in 1-1 Adult Interactions for Regular v. Reduced
Class Sizes. This figure illustrates differences between class sizes in the
time children spent in 1-1 interactions with classroom staff.

Activity Setting

This set of codes captures the activity that the teacher has prepared for the

children *or* for the target child if the activity is different from the rest of the group. Only

one code can be assigned within a given cycle. Overall, children spent just over 1/3 of the

day in whole groups, and just under 1/3 of the day in free choice. Children generally

spent their time in the same way, regardless of class size. Figure 4 below displays the

time children spent different activity settings for regular versus reduced size classrooms.

Figure 5. Time Spent in Activity Settings for Regular v. Reduced
Class Sizes. This figure illustrates differences between class sizes in the
time children spent in varying activity settings.

Peer Interactions

This section refers to the complexity of the child's social interactions with other

children. The categories are listed from the simplest to the most complex. Only one code

can be assigned in a given cycle, and in cases where multiple peer interactions occurred

in a given cycle, only the most complex interaction was coded. Overall, a very small

amount of time was spent in complex peer interactions and these percentages were almost

identical for both groups. There was a greater amount of time spent in parallel play in

regular size classes and a greater amount of simple social peer interactions in reduced

size classes, but these differences were not statistically significant. Differences between

groups in terms of the amount of time not spent in any type of peer interactions were not

analyzed since this only occurred for about 4.8 minutes across the entire sample. Figure 5

below displays the time children spent different types of peer interactions for regular
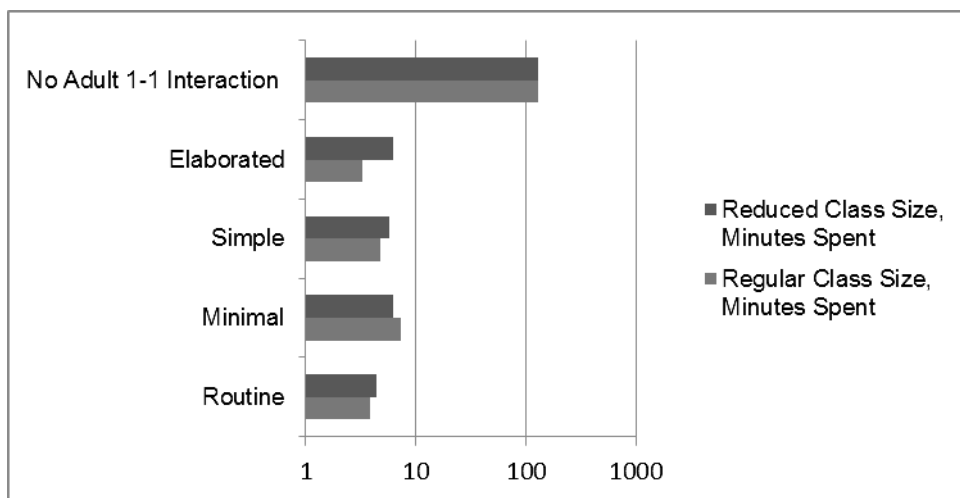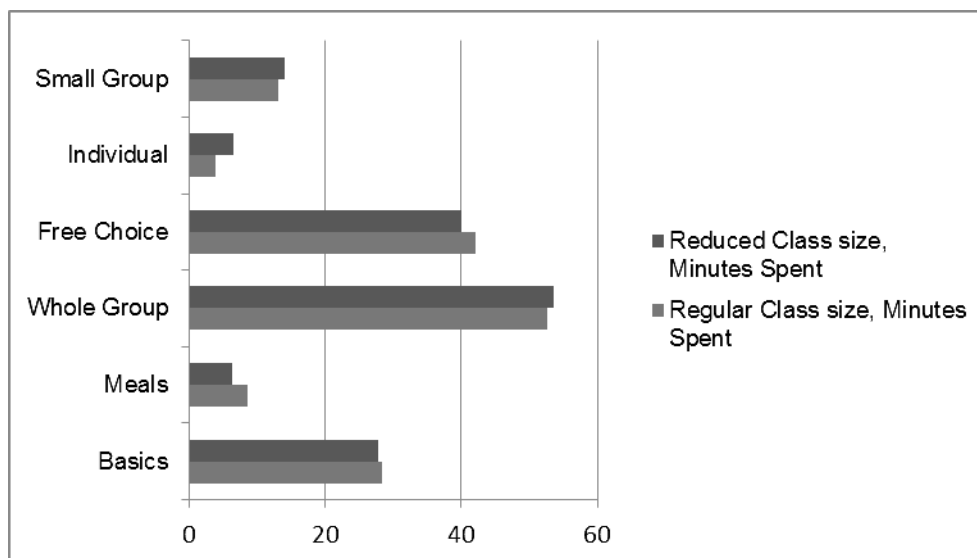
versus reduced size classrooms.
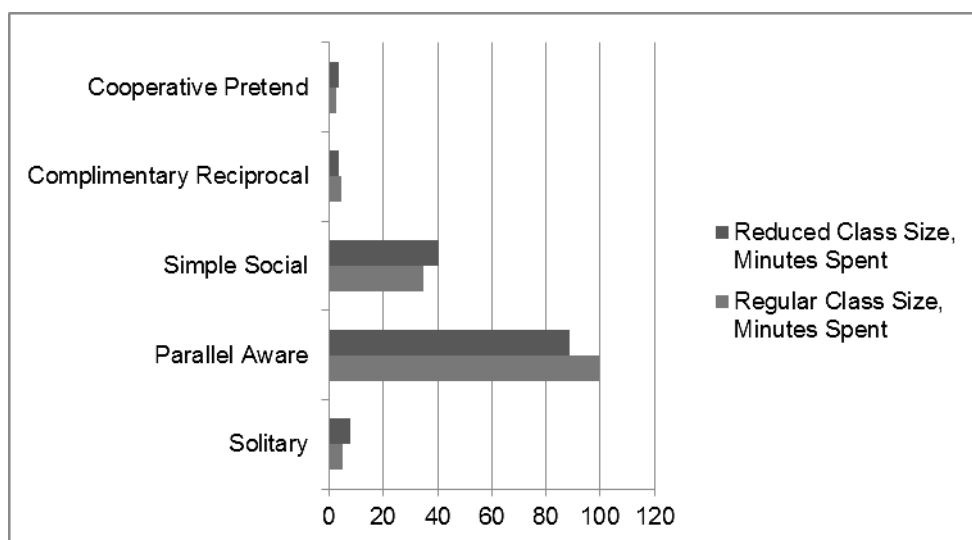
Figure 6. Time Spent in Peer Interactions for Regular v. Reduced
Class Sizes. This figure illustrates differences between class sizes in the
time children spent in different types of peer interactions.

As just discussed, paired sample t-tests were conducted to analyze the general

impact of reduced class size on classroom practices. However, these tests do not account

for differences between groups that could potentially inflate or mask the impact of class

size. As a secondary analysis, a set of regression analyses that controlled for several

classroom composition variables were conducted to test for associations between class

size and classroom quality. Classroom composition variables included maternal education

level, household income, ethnicity, average age in months, home language, and gender.

These variables were first explored to test for differences between regular and reduced

class size groups.

Results from Chi square tests showed that there was a significant relationship

between class size and maternal education, X^2 (4, N=342) = 11.54, p<.05. There was also

a significant relationship between class size and household income, X^2 (4, N=342) =

12.01, p<.05. Despite the differences in maternal education and income that were found

between groups, there was no consistent pattern that clearly differentiated the two groups. For instance, while the regular class size group had more mothers with less than a high school diploma, this group also had more mothers with a BA or more as their highest level of education. In the same respect, regular class size families were more likely to have household incomes in the 50 to 75K income range, while reduced class size families while reduced class size families were more likely to have household incomes of more than 75K. Further, there were no other classroom composition variables that differed by class size.

In addition to the concern related to classroom composition variables, another concern that arose was the number of teachers that switched their assigned reduced class size session from that which was randomly assigned. As mentioned previously, the initial research design stipulated that11 AM sessions and 11 PM sessions would be randomly chosen to be reduced in size. However, 6 teachers switched their initially assigned reduced session from AM to PM, which resulted in a total of 5 reduced size AM sessions and 17 reduced size PM sessions. It was unclear as to why teachers refused to cap their AM sessions in several instances. Therefore, the sample for the teachers that switched was compared to the sample that kept their initial reduced class size assignment on all of the variables previously identified. Comparisons between groups are presented in Table 7 below.

Table 7. Initially Assigned Session v. Switched Session Sample Characteristics

	Teachers Maintained Random Assignment (n=289)	Teachers Switched Session Assignment (n=125)
% Female	48.8	48.8
Ethnicity		
Asian	.3	.3
Black	45.7***	22.4
Hispanic	50.9	76.0***
White	2.1	--
Other	1.0	.8
% Primary English speakers	62.3***	26.4
Age in mos.	53.0	52.4
Income		
<15K	38.1	44.8
15K-25K	22.5	26.4
>25K-50K	15.9	9.6
>50K-75K	1.7	1.6
>75K	1.4	0
Missing	20.4	17.6
Maternal Education		
Less than HS	29.1	35.2
HS diploma	28.7	23.2
Some college	20.4	12.0
BA or more	6.6	12.8
Missing	15.2	16.8
Pretest Scores		
PPVT-III	81.9	85.4*
WJ-III	92.5	93.7
TOPEL/PreCTOPP	11.6	10.0

*p<.05 **p<.01 ***p<.001

There were some significant differences in terms of the characteristics between these two groups. Specifically, the group that attended sessions where teachers switched their original was characterized by a much larger percentage of Hispanic children and a much smaller percentage of Black children. Not surprisingly, the 'switch' group also

included a much smaller percentage of primary English speakers. There were minimal differences between groups in terms of income, education, gender, and age.

The teacher 'switch' variable and all of the classroom composition variables just discussed (ethnicity, age, maternal education, household income, gender, and home language), as well as session (AM/PM) were then controlled for in a set of regression analyses. Even after controlling for all of these variables, a similar picture emerged for regular versus reduced class sizes. There were no significant effects of reduced class size on classroom quality (CLASS scores) or the quantity of time spent in various types of activities and interactions (Snapshot minutes).

Despite the general consistency in classroom quality and time spent on activities and interactions, regardless of class size, it is possible that differences existed that were not detectable with the two measures that were used in the current study. This notion could help to explain the gains in children's outcomes at the end of preschool that were found to be associated with class size. Results from class size in relation to child outcome gains are presented next.

Child Outcomes

Prior to conducting analyses of child outcomes, analyses were conducted to compare the group of children that remained in the study over the course of the school year to the group that did not remain due to attrition. These groups were compared in terms of demographic variables and pre-test scores. There were significant differences between the attrition and non-attrition group in terms of gender, ethnicity, income, maternal education, and missing information on each of the latter two variables (see

Table 8 below). More specifically, the group of children that left the study after testing time 1 was characterized by a smaller percentage of females and Hispanic children, and a greater percentage of Black children.

A great deal of the attrition between time points in this study was due to one classroom closing between time 1 and time 2 testing. Therefore, the difference in ethnicity breakdown between attrition and non-attrition groups is not surprising, given that the demographic characteristics of classrooms within the Chicago Public Schools are determined largely by geographic location. This particular classroom was populated primarily by Black children. Pre-test scores also differed between the two groups, however, the non-attrition group scored higher in some areas at pretest while the attrition group scored higher in other areas. Since pretest scores were controlled for (as will be discussed below), this classroom was not included in analyses of child outcomes.

There were also large differences between the attrition and non-attrition groups in terms of income and maternal education. The most notable difference, however, lies in the percentage of missing information. Specifically, there was a much greater percentage of missing information for the group of children that did not remain in the study after time 1 testing. This is an interesting finding given that these children were also more mobile, which may represent a unique population within the Chicago Public Schools that should be studied differently or perhaps more closely.

In general, since there were such prominent differences between the group of children that remained in the study as compared to the group that left the study after time 1 testing, there are some concerns about how generalizable findings will be to this

particular population. Thus, it is important to address any unforeseen implications this might have in relation to the overall results. For instance, the effects of smaller classes on children's cognitive outcomes that will be presented and discussed next might not be generalizable to the CPS preschool population as a whole.

Even so, when we look closely at the two groups (attrition v. non-attrition) and how the reduced and regular class size groups compare within each of these groups, a more credible picture emerges. For the group of children that remained in the study over the course of both time points, demographic characteristics and pretest scores were similar regardless of whether children were assigned to a regular or reduced class size. This was also the case for the group of children that left the study after the first test point. Therefore, while there were a few areas with differences, the reduced and regular class size groups were similar enough that any effects found in relation to class size can be reasonably attributed to class size for this particular sample of children.

Table 8. Attrition v. Non-attrition Sample Characteristics

	Study Sample at Time 1 and Time 2 Testing			Sample that Left Program after Time 1 Testing		
	Full Sample (N=354)	Reduced Class Size (N=161)	Regular Class Size (N=193)	Full Sample (N=60)	Reduced Class Size (N=27)	Regular Class Size (N=33)
Gender						
% Female	50.6	49.1	51.9	40.4	44.4	36.7
Ethnicity						
Asian	.6	0	1.1	0	0	0
Black	32.5	33.5	31.5	68.4***	70.4	66.7
Hispanic	64.6***	64.0	65.2	28.1	25.9	30.0
White	1.2	0.6	1.7	3.5	3.7	3.3
Other	1.2	1.9	0.6	0	0	0
Age in Months	52.7	52.09	53.20	53.6	53.8	53.5
% Primary English Speakers	45.6	48.4	43.1	82.5***	85.2	80.0
Income						
<15K	41.5	39.8	43.1	33.3	29.6	36.7
15K-25K	26.3	24.2	28.2	12.3	18.5	6.7
25K-50K	15.8	16.1	15.5	3.5	0	6.7
50K-75K	1.5	0	2.8*	3.5	7.4	0
>75K	.9	1.9	0	1.8	3.7	0
Missing	14.0	18.0	10.5	45.6	40.7	50.0
Maternal Education						
Less than HS	34.8*	27.3	41.4**	10.5	11.1	10
HS diploma	27.2	33.5**	21.5	22.8	22.2	23.3
Some college	18.1	19.9	16.6	21.1*	25.9	16.7
BA or more	9.4	7.5	11.0	3.5	3.7	3.3
Missing	10.5	11.8	9.4	42.1	37.0	46.7
Pretest Scores						
PPVT-III	83.12	82.52	83.64	81.75	80.67	82.76
TOPEL/PreCTOPP	10.47	9.83	11.04	14.88**	12.67	16.87
WJ-III	92.47	91.46	93.38	95.26	96.26	94.37

*p<.05 **p<.01 ***p<.001

Child outcome measures were next analyzed using a multi-level regression model,

given students were clustered within teachers, and each teacher had both a regular and

reduced size class. These analyses controlled for age, pretest score, home language, gender, ethnicity, household income, and maternal education level. Quality scores, session, and 'teacher switch' were also included in supplementary analyses to check whether these variables altered the results. None of these variables alone were associated with child outcomes and they did not change the overall results. Analyses of each child outcome measure were conducted for the full sample, and separately for Hispanic students and non-Hispanic students.

Gain scores are also presented, to provide a simple picture of differences between groups prior to controlling for multiple variables. Children's vocabulary (PPVT-III), early literacy (TOPEL), and early math skills (WJ-10) were assessed at the beginning and end of their preschool year. Only those teachers with both regular and reduced class sizes were included in these analyses. The two classrooms that did not have consent forms for any children in the PM sessions were removed from child analyses, resulting in a sample size of 342 children from 19 teachers (38 sessions).

PPVT/TVIP

As mentioned previously, bilingual students were administered both the PPVT and the TVIP, to get a sense of their receptive vocabulary skills in both languages. Standard scores were calculated for both and whichever measure had the best standard score was used for the dependent variable. In general, a mean score of 100 on the PPVT represents the 50th percentile, and a standard deviation of 15 indicates that 67% of the national population would attain standard scores between 85 and 115. Further, 95% of the population would attain standard scores between 70 and 130. The sample in the current

study appears to attain scores similar to the national population, and specifically the national population of children that attend Head Start.

The mean scores for the sample were 83.12 in Fall 2008 and 86.55 in Spring 2009, which are slightly lower than those reported for the Fall 2009 (85.4) and Spring 2010 (88.9) FACES (Family and Child Experiences Survey) data that presents the most recent Head Start figures (OPRE Report 2011-37b; Aikens et al., 2012). The average gain on the PPVT/TVIP was 3.6 points. This compares to 3.6 points for Head Start nationally in 2009, again the most recent reported Head Start figure. Children in reduced class sizes gained an average of 4.74 points, while those in regular class sizes gained an average of 2.62 points. Results are presented in Table 9 below.

Results from multi-level regression analyses showed that students in reduced class sizes scored non-significantly higher than students in regular class sizes in receptive vocabulary, after controlling for age, pretest score, home language, gender, ethnicity, household income, and maternal education level. Pre-test score, gender, and income level were found to be significantly associated with children's vocabulary at the end of the school year. Effect sizes were also calculated (using Cohen's *d*) to provide an estimate of the size of the effect of class size reduction on children's learning. Effect sizes help to put the relationship between class size and outcome scores in perspective, rather than simply looking at whether the relationship could be due to chance. They allow us to quantify the differences that exist between regular and reduced class size groups in the current sample, without attending to sample size or the true relationship that exists in the population. The effect size (coeff./control group std. dev.) on PPVT/TVIP scores was

found to be 0.12, indicating that children in reduced class sizes scored approximately one eighth of a standard deviation above children in regular class sizes.

When data was analyzed separately for Hispanic students compared to others in the sample, results were similar. Receptive vocabulary scores were non-significantly higher for students in reduced class sizes. Further, effect sizes were 0.13 for both Hispanic children and non-Hispanic children when they were calculated separately for each group, thus they were quite comparable to the overall sample. Results are presented in Table 10 below.

TOPEL/PRE-CTOPPP

Children were administered the TOPEL in either English or Spanish, depending upon what the teacher designated as the child's best testing language. Regardless of which language was used for administration, raw scores were combined into one variable because the Spanish version is a direct translation of the English version. The mean scores for the sample were 10.47 in Fall 2008 and 16.91 in Spring 2009, for an average gain of 6.41 points over the course of one school year. Children in reduced class sizes gained an average of 7.36 points, while those in regular class sizes gained an average of 5.56 points. This difference between groups was statistically significant ($p<.05$). Results are presented in Table 9 below.

Results from multi-level analyses showed that students in reduced class sizes scored significantly higher than students in regular class sizes in early literacy skills, after controlling for age, pretest score, home language, gender, ethnicity, household income, and maternal education level. Pre-test, age, and gender were also found to be

significantly associated with children's early literacy scores at post-test. In addition, the effect on TOPEL scores was found to be 0.20, or one fifth of a standard deviation gain for children in reduced class sizes. Results are presented in Table 10 below.

Again, TOPEL scores were analyzed separately for Hispanic students compared to others in the sample. Results were similar to those that were found for the full sample in that TOPEL scores were higher for students that attended reduced class sizes. This was the case for both the Hispanic and non-Hispanic groups. Differences were statistically significant at the p<.10 level rather than the p<.05 level for both groups, which is likely due to the reduced sample sizes. Effect sizes were slightly larger for the non-Hispanic group (.18) compared to the Hispanic group (0.20), and both were comparable to those found for the overall sample.

To put these effect sizes in context, they can be compared to a recent meta-analysis which analyzed the average effect size associated with preschool participation across multiple studies (Camilli et al., 2010). Results showed that the average unweighted effect size on children's cognitive outcomes for treatment/control conditions was .231, which is similar to that found in the current study. In the meta-analysis, cognitive outcomes include intelligence and cognitive/reading outcomes combined into one domain.

WJ-III/WM-R

Children were administered the Applied Problems subtest of the Woodcock-Johnson-III in English or of the Woodcock-Munoz Revised version in Spanish, depending upon what the teacher designated as the child's best testing language. The

mean standard score for the sample was 92.47 in Fall 2008 and 93.38 in Spring 2009. Again, these means are similar to those reported for the 2009-10 Head Start FACES sample. It is difficult to make direct comparisons because National Head Start means are reported separately for children that were administered the assessment in English (90.0 in Fall 2009 and 92.4 in Spring 2010) versus those administered the assessment in Spanish (82.3 in Fall 2009 and 84.2 in Spring 2010). However, it is clear that the scores in this sample are somewhat higher than scores would be from the national Head Start sample if English and Spanish language scores were reported together.

The average gain for the sample was 0.91 standard score points, which is lower than national Head Start gains of 2.4 points for children administered the assessment in English and 1.9 points for those administered the assessment in Spanish. However, the average gain for children in reduced class sizes was 1.86, compared to a gain of only .04 standard score points for children that attended regular class sizes. Results are presented in Table 9 below.

Findings from multi-level analyses revealed that children in reduced class sizes scored an average of 0.88 standard score points higher than children in regular class sizes, after controlling for age, pretest score, home language, gender, ethnicity, household income, and maternal education level. However, this difference was not statistically significant. The estimated effect of class size on WJ-10 scores was found to be 0.07. Results are presented in Table 10 below.

When scores were analyzed separately for Hispanic students compared to other ethnicities, again no statistically significant differences were found. The estimated effect

of class size on WJ-10 scores was found to be 0.13 for the Hispanic group and 0.01 for

the non-Hispanic group. In this case, Hispanic students appeared to benefit more from

reduced class size than their non-Hispanic peers.

Table 9. Gain Scores for Reduced versus Regular Class Sizes

	Total Sample			Regular Class Size			Reduced Class Size		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Child Outcomes									
PPVT standard score									
Fall	337	83.12	13.35	181	83.64	13.74	156	82.53	12.89
Spring	342	86.55	13.82	181	86.25	14.03	161	86.89	13.62
Fall – Spring Gain	337	3.60	12.04	181	2.62	11.41	156	4.74	12.69
Woodcock Johnson standard score									
Fall	340	92.47	10.77	179	93.38	11.16	161	91.46	10.26
Spring	342	93.38	10.96	181	93.42	11.38	161	93.33	10.50
Fall – Spring Gain	340	0.91	9.60	179	0.04	10.39	161	1.87	8.58
TOPEL									
Fall	341	10.47	8.55	180	11.04	8.86	161	9.83	8.16
Spring	341	16.91	10.23	181	16.62	10.17	160	17.24	10.32
Fall – Spring Gain	340	6.41	7.34	180	5.56	6.94	160	7.36*	7.68
Child Characteristics									
Age in mos. (at pre-test)	342	52.68	6.43	181	53.20	6.07	161	52.09	6.79
Child's gender (%)									
Male	169	49.4%		87	48.1%		82	50.9%	
Race (%)									
Asian	2	0.6		2	1.1		--	--	
Black	111	32.5		57	31.5		54	33.5	
Hispanic	221	64.6		118	65.2		103	64.0	
White	4	1.2		3	1.7		1	0.6	
Other	4	1.2		1	0.6		3	1.9	
Home Language (%)									
English only	156	45.6		78	43.1		78	48.4	
Spanish*	183	53.5		103	56.9		80	49.7	
Other	3	0.9		--	--		3	1.9	
Income (%)									
Less than 15K	142	41.5		78	43.1		64	39.8	
>15K to 25K	90	26.3		51	28.2		39	24.2	
>25K to 50K	54	15.8		28	15.5		26	16.1	
>50K to 75K	5	1.5		5	2.8		--	--	
>75K	3	0.9		--	--		3	1.9*	
Missing	48	14.0		19	10.5		29	18.0	
Maternal Education (%)									
Less than hs	119	34.8		75	41.4*		44	27.3	
HS diploma	93	27.2		39	21.5		54	33.5*	
Some College	62	18.1		30	16.6		32	19.9	
BA or more	32	9.4		20	11.0		12	7.5	
Missing	36	10.5		17	9.4		19	11.8	

 *p<.05, **p<.01, ***p<.001

*Note: Children with primary home language 'Spanish' or 'English and Spanish' are included in this category.

Table 10. Multilevel Results: Reduced Class Size/Child Outcome Scores

	WJ Standard Score	PPVT/TVIP Standard Score	TOPEL Raw Score
Reduced Class Size	0.88	1.74	1.99**
	(0.92)	(1.17)	(0.74)
Effect Size	0.07	0.12	0.20
Controls:			
Pre-test score	0.57***	0.60***	0.77***
	(.04)	(.05)	(.05)
Age in months	-0.18*	0.06	0.27***
	(.07)	(.09)	(.07)
Gender	0.73	3.56*	2.49**
	(0.92)	(1.17)	(.74)
Black	-8.95**	-4.89	-3.58
	(3.33)	(4.16)	(2.80)
Hispanic	-9.12**	-3.30	-3.22
	(3.51)	(4.36)	(2.95)
Other	-10.29	-12.17	-2.82
	(5.92)	(8.56)	(4.71)
Home Language English	0.53	2.50	0.63
	(1.48)	(1.86)	(1.27)
Income	1.37	1.61*	-0.13
	(0.62)	(0.79)	(0.50)
Income Missing	1.18	1.95	-1.31
	(1.95)	(2.47)	(1.59)
Education	0.63	0.26	0.76
	(0.56)	(0.71)	(0.46)
Education Missing	3.24	5.19	0.17
	(2.17)	(2.75)	(1.78)
N	340	337	340

*p<.05, **p<.01, ***p<.001. Standard error in parenthesis.

CHAPTER 5

DISCUSSION

The current study randomly assigned teachers' am or pm preschool sessions to

class sizes of either 15 students (reduced) or 20 students (regular). Thus, each teacher had

one session of each class size and the class size comparison controls for all teacher

characteristics. In actuality, reduced size classrooms averaged 12.61 students and regular

size classrooms averaged 16.23 students, based on attendance data that was collected at

the time of the observations. The bulk of prior research on class size would consider both

of these small classes. Even so, this study finds that there are differences in how much

children learn in the domains of vocabulary and literacy, with children learning more if

they attend classrooms with fewer students. Measures of classroom quality and activities

revealed few indications of differences in teacher behavior that would explain the

differences in child outcomes. It may be that since each teacher prepared for one small

and one large class daily this constrained how much they changed their practices for the

smaller class. The findings also suggest the importance for quality of factors not

captured by measures of teacher child interaction such as the CLASS and other measures

of classroom activity.

Effects on Classroom Life

The two measures of classroom practices that were used for this study were

chosen very specifically. The CLASS was chosen based on the understanding it would be

responsive to differences in the types of interactions between teachers and children in reduced versus regular size classes. In other words, it measures the quality of interactions that occur between teachers and children. In contrast, the Snapshot gets at the quantity of interactions between teachers and children, as well as the content and settings that children partake in on their own and with their peers. It is assumed that both of these are very important to children's learning, but only on the second was there any evidence of differences by class size.

As with any measure of classroom quality, there is the possibility for scores to be affected simply by having observers in the classroom. The presence of an additional person disrupts the natural classroom life, regardless of how unobtrusive the individual might be (Blease, 1983). While there is not much that can be done about this given that the nature of these measures require that someone observe within the context of the classroom, it is important to point out nonetheless.

The quality of classrooms for the full study sample averaged in the mid-range, as rated by the CLASS. This means that on average there were positive behaviors in the areas that are defined in the CLASS, but they were not always consistent and at times infrequent during the school day. Average scores were similar regardless of class size.

The *emotional support* domain scored highest, averaging mid to high scores, indicating that teachers were relatively good at providing emotional support and that some classrooms had very little room for improvement in particular areas. The *classroom organization* domain was characterized by mid-range scores, which indicates that on average behaviors defined in this domain were not always consistent and at times

infrequent during the school day. The *instructional support* domain was the lowest-scoring area, with desirable behaviors relatively rarely occurring for these types of interactions. This pattern of behaviors, with emotional support scoring highest, followed by classroom organization and then instructional support, was the same for both regular and reduced size classes, and scores for both class sizes were nearly identical. This is the typical trend for samples that are studied with the CLASS (e.g. Pianta et al, 2005; LaParo et al, 2009).

The Snapshot breaks down interactions into five different categories, or subscales. These include (1) interactions that take place between teachers and children, (2) one-to-one interactions between teachers and children, (3) learning content that children engage in either on their own, with peers, or with their teachers, (4) types of groupings of children's activities, and (5) peer interactions among children. The only suggestions of any differences between regular and reduced class sizes were found for the first two of these subscales.

There was some evidence that the percentage of time that children and teachers were involved in 1-to-1 elaborated conversations was notably greater in reduced class sizes. This makes sense as the smaller number of children should have allowed teachers to interact with individual children more frequently. While the average amount of time that was spent was small (only 3.1 minutes in regular classes and 6.3 minutes in reduced class sizes), this amounts to 9.3 and 18.9 hours of individual attention for each child in reduced classes and regular classes, respectively, over the course of a 180 day school year. While this is a notable difference across a school year, another way to look at it is

for a given week, in which case it amounts to only about 15-minutes and 30-minutes per week for regular and reduced class sizes respectively, and clearly there is room for improvement regardless of class size.

These results are important not only in light of the preschool experiences they presented for children, but also because of the potential they create for impacting children's achievement. For instance, a recent meta-analysis confirms that preschool education has significant lasting positive effects on cognitive abilities, school progress (e.g., less grade repetition, less special education placement, and increased high school graduation), and social behavior, and effects are larger when programs focus on small group learning and individualized one-on-one teaching (Camilli et al., 2010). Additionally, findings from the Effective Provision of Preschool Education (EPPE) project showed that the most effective preschool settings (in terms of their impact on children's cognitive outcomes) were those that encouraged 'sustained shared thinking,' or extended narratives between teachers and children, through 1-to-1 adult-child interactions (Siraj-Blatchford et al., 2003).

There were also a greater percentage of didactic interactions that occurred between teachers and children in reduced size classes, and differences again approached significance. Didactic interactions are coded in the teacher-child interaction subscale of the Snapshot, thus again differences were related to what occurs between teachers and children rather than what children engage in or the degree to which they interact with one another.

Despite the differences in the amount of 1-1 interactions and didactic interactions for regular and reduced class sizes, the average child in the sample spent nearly 2/3 of the day not engaged in any type of teacher-child interactions and over 85% of the day not engaged in direct 1-1 interactions with their teachers. These percentages were not found to be significantly different between the two class size groups.

In addition, children spent over 1/3 of the day not engaged in any type of learning activities that were observed. They also spent the majority of their time in parallel play, rather than partaking in direct interactions with one another. The setting in which children spent the majority of their time was in a whole group, followed by free choice time. Again, no significant differences between regular and reduced class sizes were found for these subscales.

While there were differences found in two areas of the *Emergent Academic Snapshot*, there were numerous areas of the measure where no differences between class sizes were detected. Given the large number of analyses that were conducted, it is possible that the significant findings that emerged were coincidental and it is difficult to have a great deal of confidence in their meaning.

There is room to boost the amount of interactions that occur in the classroom regardless of class size. It is possible that reduced class sizes could produce more pronounced differences if teacher preparation or professional development efforts were aimed at helping teachers to increase the number and duration of interactions they have with the children in their classrooms. In other words, because differences were found, albeit small, strategies that help teachers to increase the number of individual interactions

in the classroom might increase the effects of small classes on teacher practices and result in even larger impacts on children's learning than were found.

Currently, the field lacks information about what is taught in professional development programs and how that instruction is delivered, and the mechanisms through which it translates into classroom practice. In fact, federal funding initiatives in early childhood conclude that a careful, systematic program of research is needed to understand the impact of early childhood teacher professional development programs on teacher quality, classroom quality, and child outcomes (U.S. Department of Education, IES, 2005). Research that does exist indicates that professional development can in fact lead to improved teacher behavior and instructional practice, and enhanced child outcomes (Whitaker, Kinzie, Kraft-Sayre, Mashburn, & Pianta, 2007; Raver et al., 2008; Layzer, L.I., et al., 2007). However, the benefits of particular professional development strategies are unclear, and whether or not such strategies target the frequency of teacher-child interactions is not known (Ramey & Ramey, 2008).

In contrast, professional development research related to the CLASS is more straightforward. One particular study found clear associations between a web-based system of professional development research and the quality of preschool teachers' classroom interactions (Pianta et al., 2008). Specifically, teachers in a "consultancy group" that were assigned to receive on-line consultation and feedback targeted to their interactions showed significantly greater increases in ratings of the quality of interactions than did those only receiving access to a website with video clips. Score increases for the consultancy group of teachers ranged from one half to one point on the CLASS after just

one year of the intervention. Consequently, if teachers are given skills to increase the quality of their interactions, smaller classes might be more effective.

Perhaps more important than focusing on increasing the time spent in teacher-child interactions or the quality of those interactions is helping teachers think through techniques to take advantage of smaller classes. Given the lack of differences that were found overall for classroom interactions between the different class sizes, it is clear that reducing class size does not automatically result in differences in classroom practices, and training teachers on skills to maximize the benefits of smaller classes might have led to more pronounced differences.

On the other hand, it is also possible that differences in classroom practices were not found simply because the average class sizes were quite small for both the "regular" and "reduced" class size groups. By most standards, both class sizes that were studied would be considered small and perhaps quality would have varied more if the "regular" size classes were as large as they were assigned to be. However, even though class sizes in this sample were on average smaller than they were assigned to be, there were a few classrooms that included the maximum 20 children when they were observed. Teachers in these classrooms did not score lowest on the CLASS or partake in the least amount of Snapshot interactions; therefore, it is difficult to put much confidence in this line of reasoning.

Additionally, it is important to recognize that the measures of classroom quality that were employed may not be capturing some classroom dynamics that are beneficial to children's achievement. In light of this, the inclusion of alternative measures of

classroom quality would be beneficial for future research. This might require the use of measures that are more closely aligned with curriculum models that are utilized, or even the development of new measures that are based on qualitative investigations.

For instance, future research that qualitatively examines class size differences might allow for a better understanding of the mechanisms through which class size affects children's development. This type of research could incorporate teacher perspectives regarding potential benefits of smaller classes and their methods for maximizing those benefits in the classroom. Transactional models of classroom dynamics would also be a useful framework for future research on class size, as transactional models would allow for a more thorough understanding of how class size might affect different classrooms differentially because of the unique composition that characterizes each classroom.

In general, results from the current study show that reducing class size does not inevitably increase classroom quality but it may increase the quantity of interactions in the classroom. However, differences between class sizes were minimal and since there were significant and substantive effects of class size on children's learning, findings suggest that class size reduction affected children in some other way that was not measured by the CLASS or Snapshot.

Effects on Student Achievement

Children's outcomes as they relate to class size were examined in several ways. The simplest analyses compared gains of children that attended reduced class sizes to those that attended regular class sizes. Results from these analyses showed that children

in reduced class sizes gained modestly more on every assessment measure over the course of one school year. These findings were them compared to results from FACES data (for measures that were consistent across studies), which provide annual reports on the national Head Start population. This comparison was made because of the similarities between the current sample and the national Head Start population in terms of household income, maternal education, etc.

In general, gains were about the same for the current sample and the National Head Start sample from the 2009-10 school year, and they were slightly larger for the reduced class size sample. This indicates that smaller classes may contribute to greater gains for children in preschool after just one year. It is important to note, however, that children in the small class sample began the school year with slightly lower average scores on every assessment measure. While these differences between groups at the start of the school year were not significant, they were large enough such that mean scores at the end of the school about year were the same for the reduced and regular size samples, despite larger gains for the reduced size group.

In addition to gain scores favoring reduced size classes, results from multilevel analyses also showed evidence that reduced classes are associated with greater cognitive gains for children. Specifically, attending a reduced size classroom was associated with an increase of 1.74 standard score points on the PPVT/TVIP and 1.99 raw score points on the TOPEL at the conclusion of the preschool year. TOPEL increases were found to be statistically significant and PPVT increases approached significance when controlling for pre-test, ethnicity, age, gender, income and education. There were no statistically

significant differences between regular and reduced class size groups for increases in early math skills.

While PPVT increases did not prove to be significantly associated with reduced class size, the increase is noteworthy in light of other research that has been conducted on classroom quality and children's learning gains. For instance, a study by Mashburn et al. (2008) indicates that a 1-point gain on the CLASS instructional support domain, which is more than a full standard deviation, would be required to produce a .69 gain on the PPVT. Similar associations were found with *Instructional Support* quality and PPVT gains in this study. Therefore, to get the size gain that was found from class size would require a 2.5 point gain on the instructional support domain. This would require teachers in the current study to score well above a 5, on average, to produce gains similar to those found from class size. As mentioned previously, professional development efforts related to the CLASS have shown improved scores of one half to one point after one year, in response to PD that was offered (Pianta et al., 2008). In effect, quite intensive efforts would be needed to achieve a 2.5 point gain in *Instructional Support* quality.

Effect sizes were also calculated as a way to provide an estimate of the strength of the relationship between class size and outcome scores. The estimated effect of a reduced class size on children's receptive vocabulary was 0.12, or one eighth of a standard deviation. Further, the estimated effect of a reduced class size on children's early literacy skills was 0.20, or one fifth of a standard deviation. These effect sizes coincide with a jump in children's achievement, roughly from the 50th to the 52nd percentile in receptive vocabulary and from the 50th to the 53rd percentile in early literacy.

Implications and Limitations

The current study is unique from previous research on class size in a few ways. First, there was the possibility to compare the effects of reduced and regular class sizes among the same group of teachers, as each teacher had both a regular and reduced size class. This allowed controlling for a host of variables that might differ between teachers, which could potentially mask the effects of class size otherwise. Further, the session to which a reduced class size was assigned (AM versus PM) was chosen randomly at the start of the school year, which allowed for the random assignment of children to class sizes. In other words, parents chose to enroll their children into either an AM or PM session without any knowledge of whether the class would be reduced in size.

The fact that teachers switched the assigned reduced session from AM to PM in a number of cases introduces a limitation to this study, as there were indeed differences between the group that switched from random assignment and the group that maintained random assignment. Specifically, the classrooms that maintained random assignment had a greater percentage of African American children and primary English speakers compared to the group that switched, which had more Hispanic Spanish-speakers. This could simply be a product of the concentrations of particular ethnicities in different locations in the Chicago Public Schools, since a teacher's choice to switch did not appear to have any effect on classroom practices. Even so, it is important to acknowledge this difference as a potential bias in the interpretation of findings as it may be related to some influential variable that was not measured in this study.

In addition, class sizes of 15 were compared to class sizes of 20, which in much of the class size research would both be defined as small classes. Further, each class was staffed by two adults at all times, allowing for ratios of 1:7.5 and 1:10, respectively. Despite both class sizes being considered small, it was found that a reduction from 20 to 15 in a preschool setting has large implications for children's learning and development. Specifically, children who attended reduced class sizes experienced more one-one interactions with teachers and they scored significantly higher on a measure of early literacy after just one year of preschool. Further, there were not consistent differential effects of class size for different groups of students (i.e. by ethnicity or income level).

It is also important to point out that the majority of children served in these classrooms were Latino. Nearly 60 percent of the students in the sample were identified as Hispanic and about 50 percent were identified as primary Spanish-speakers. Bilingual data collectors observed in classrooms where a large percentage of Spanish-speakers were identified but only a small percentage of classroom discourse occurred in Spanish even in these classrooms. This is not surprising given that only 5 out of the 22 lead classroom teachers spoke Spanish. Several assistant teachers also spoke Spanish (the exact number is not known because this data was not formally collected) but the majority of classroom discourse was found to occur between lead teachers and the children.

When models were estimated separately for the Hispanic group of children compared to the non-Hispanic group, effect sizes were essentially unchanged. The exception to this was for the Applied Problems subtest of the Woodcock-Johnson.

Specifically, the estimated effect of a reduced class size on children's early math skills was larger for the Hispanic group (.13) compared to the non-Hispanic group (.01).

Research that has examined the degree of Spanish-speaking in preschool classrooms has shown positive associations with more Spanish usage and learning gains for English language learners (Barnett et al., 2007; August & Shanahan, 2008). Perhaps if Spanish was spoken to a greater degree in smaller classrooms that contained large numbers of Spanish-speaking children, smaller class sizes might have benefitted achievement levels even more for this particular population.

It is also important to note the actual sizes of classes that were observed in the current study. While the aim was to compare class sizes of 15 (reduced) to class sizes of 20 (regular), it was found that the average class size for a "reduced" class was 12.61 students. The average class size for a "regular" class size was 16.23 students. Therefore, typical daily class sizes were indeed smaller, on average, than the assigned enrollment, and the difference in class size was 3.6 students on average compared to a difference of 5 in enrollment as assigned. However, the "attendance rates" are similar: 84% in the smaller classes and 81% in the larger classes. The difference of 3.6 students per preschool classroom is still substantial and is nearly as large a percentage difference from actual class size as that between 15 and 20.

The actual class sizes that were observed highlight the need to reduce absence rates, and to gain a better understanding of the nature of absence rates. For instance, is the 20 percent absence rate that was found due to several children that were very frequently absent or were absences spread evenly among children? Chronic absenteeism has been

defined as students who miss school more than 10 percent of the time (Romero & Lee, 2007). Therefore, if absences were evenly spread among students in this study, the average student would be characterized as chronically absent. However, this line of research is focused on elementary students and absent rates in preschool must be viewed differently since preschool is not mandatory and children are essentially learning to go to school while their parents are learning to send their children to school. Even so, this does not mean that high absence rates are not problematic. In fact, research shows that absence rates among Chicago Public School preschool students have been linked with lower achievement levels (Ehrlich et al., 2013).

In essence, frequent absences are an issue that cannot be ignored, as class size or any other intervention is unlikely to be as effective as it could be if students are frequently absent. A large body of research has linked higher absence rates with lower test scores and higher dropout rates later in life (Lamdin, 1996; Alexander et al., 1997). In light of this, it is possible that reduced absence rates might have led to more pronounced class size effects.

In addition to the effects that high absence rates can have on children, the actual class sizes that were observed in this study should also be acknowledged because of the potential differences they present for children. For instance, a class size of 16.23 preschool students is much different than that of 20 preschool students, and a class size of 12.61 students is much different than that of 15 students. A class size of 16.23 is quite small and if children who attended actual class sizes of 15 were compared to children

who attended actual class sizes of 20, effects of reduced class size might have proven to be much larger.

Another way to look at this is that if absence rates decrease, smaller classes might be even more important. Future research that examines absence rates over time and the reasons behind them across different locations could have implications for class size policies. In fact, a policy that is adopted in a location with historically high absence rates might not be ideal for a location that tends to have good attendance.

Of further note, all teachers in the current study used the *Creative Curriculum*, which maintains a strong literacy focus. This might have something to do with the larger effects of class size found for literacy, as opposed to math. Perhaps if the curriculum was more focused on math there might have been greater effects on children's early math skills attributable to class size.

In fact, research on math-focused curriculums has shown that they can have effects on children's mathematics achievement after one preschool year (Clements and Sarama, 2008; Clements et al., 2011). In each of these studies, a math-focused curriculum called *Building Blocks* was implemented in several classrooms and compared to a group of 'control' classrooms that implemented alternative curriculum models. One of the curriculum models used in the control condition was in fact the *Creative Curriculum*, which emphasized math implementation through play with provided materials and teacher scaffolding. However, this was done only in small groups for about 10-minutes per week. In contrast, the *Building Blocks* curriculum emphasized teaching strategies that shared a core of interwoven, research-based learning trajectories with activities that were

formatively evaluated. This was done for nearly two hours per week in small groups and large groups as well as on the computer and through engagement with families. Given the minimal focus placed on math within the *Creative Curriculum*, it makes sense that class size did not have an effect on children's mathematics gains. It is indeed possible that class size could have large effects on mathematics achievement if coupled with an intensive preschool mathematics curriculum, such as *Building Blocks*.

We must also acknowledge that the context and schools involved play a role in any intervention that is studied. Research has shown that the context of class size reduction can affect the degree to which it improves student achievement. For instance, class size reforms in Florida and California were met with a number of challenges such as unqualified teachers and a lack of classroom space to accommodate children (Ready, 2008). Class size reduction cannot be viewed outside of the constraints it produces. Smaller classes require additional classroom space, which is feasible in some schools but not others. Thus, we must consider potentially negative impacts of smaller classes if it is implemented in such a way that teachers are forced to teach in inadequate spaces. In this study, the space was the same, and this should be an important consideration when implementing class size reduction.

Graue (2009) conducted a qualitative study of the implementation of smaller classes, and findings point to several associated constraints that must be considered. For instance, teachers expressed the difficulty with cramming 30 children and two teachers into a very small space and their preference for continuing with 20 or even 25 children and one teacher in the same space. They also described the classroom arrangement as a

'teaching assistant' if the space was sufficient and it was set up in a way that helped them to smoothly carry out managerial tasks. In Graue's (2009) study, larger classes were not favored by teachers in and of themselves but the constraints produced by smaller classes were challenging enough that larger classes were viewed as the lesser impediment to optimal teaching.

It is also important to consider the need for additional teachers when class size reductions are implemented. Placing additional teachers in the classroom simply to reduce class sizes runs the risk of staffing classrooms with underqualified or unprepared teachers. Reducing class sizes without any attention to teacher qualifications and performance might reduce the positive effects on children's achievement levels, given the associations that have been found between teacher qualifications and children's learning (Berk, 1985; Burchinal et al., 2000). However, the implementation of smaller classes has been found to reduce teacher stress levels, which is important for teacher consistency and retention (Hattie, 2005). In effect, it is possible that class size reduction might boost student achievement more over time as teachers remain in their classrooms. Therefore, studies that estimate the effects of class size on student achievement levels only in the year immediately after class size reduction or that estimate it for the same teacher (such as this study) may underestimate the effects of class size reduction because they do not capture the differential effects on teacher stress levels and turnover.

The unique design of this study allowed for each teacher to have one session of each class size and, therefore, the class size comparison controls for all teacher characteristics. However, one potential limitation of this approach relates to teachers'

preparation for classroom instruction. If teachers have one class of each size, they have to prepare for both and perhaps teach the same way for both. In contrast, if teachers only teach a small class it is possible that they might change their practices more.

The current study provides evidence that a difference of five children in a preschool classroom's assigned enrollment with two teachers is associated with at best modestly different experiences for children as assessed by typical measures of classroom quality and activities. In contrast, there were substantial differences in children's achievement levels after one year. The notion that there were differences in child outcomes irrespective of differences in CLASS scores has broad implications for relying just on process measures or professional development to improve programs or regulate quality. Essentially, this shows that structure can matter in ways not captured by the CLASS, and therefore claims that the CLASS alone can be used as a guide to quality are not sufficient.

These findings with respect to children's learning coincide with findings from the majority of the class size research, including Tennessee's STAR project, Wisconsin's SAGE project, and California's CSRP, which has largely found that smaller class size leads to increases in student achievement, helps to close the minority-majority achievement gap, and has several other long lasting benefits. The current study is unique because it speaks specifically the preschool period, for which there is very limited highly rigorous research on class size.

Despite the contribution to the class size literature that this study marks, it demonstrates the need for more and different kinds of research, as several questions

remain. For instance, to what extent is one year in a reduced class size enough to benefit student achievement over time, and to what extent will gains erode after returning to a class of regular size? How do the benefits of class size reductions in preschool compare with those in early elementary grades, or even high school? Do the gains in children's achievement offset the extra costs that they require? How much would it cost to produce the same additional gain by other means, for example by paying teachers more, providing professional development, or starting at age 3? The current study simply examines classroom life and student achievement as they relate to class size, and while smaller classes appear to benefit student's achievement, no attention has been given to the overall benefits in relation to costs. Future research that examines costs versus benefits in preschool would be useful for making more clear policy recommendations.

Additionally, this study only examined the effects of class size over one year of time, and we do not know whether the benefits that were shown after preschool persist or diminish over time. Further, one year of preschool is a small amount of time in a child's life and it is possible that influences from the preschool year take time to manifest. While this study provides evidence that it is possible to detect modest positive effects of small classes after just one school year, it would be valuable to examine these effects over time in future research. It is possible that enhanced achievement in preschool could accumulate over time, leading to even greater impacts on children's achievement later in life. It is also possible that effects after one year of preschool could fade out over time, and that the early elementary experiences that children have are capable of reducing any gaps that might have resulted from differences in preschool class sizes. Even so, this

study provides evidence that children are better equipped with skills to help them succeed in kindergarten if they attend smaller classes in preschool.

We also must consider teacher training related to class size. For instance, if teachers are not given training on how to work better with smaller classes it may not result in any differences in what children experience or how much they achieve. Research has shown that teachers use various tactics in smaller classes, such as tag team teaching versus co-teaching and teachers are not often trained to implement strategies that take advantage of smaller class sizes (Graue, 2009). This notion makes it very difficult to generalize any of the findings from this study, but the point is vital nonetheless.

Finally, the question still remains: is there an optimal class size for maximizing student achievement? The smaller class sizes in the STAR project were 13 to 17 students while the larger class sizes were 22 to 25 students. While the SAGE project used smaller class sizes of 15 or less and a variety of larger ones, the California CSRP used class sizes of less than 20 as well as various class sizes of more than 20. Although smaller class size has been found to result in increases in student achievement, a "small" class size is not well defined in the literature. One advantage of the current study is the clear definition of class size: 20 students represents a regular class size and 15 students represents a reduced class size. This detail is useful for providing straightforward policy implications.

Future research could also focus on whether smaller class size affects different groups differently, and the extent to which different groups such as students who are poor performers come from low income families, members of a minority group, or English language learners might require different class sizes to maximize their achievement.

Additional research could also be conducted to determine the impact that different state teacher education and certification requirements have on the ability of teachers in different states to influence student achievement. It would also be helpful to learn what unique teaching methods are best suited for small as well as large class sizes and how to train teachers to use these methods effectively in the classroom.

Finally, an important limitation of this dissertation is that children's achievement level was the key construct of interest, and this represents just one facet of their overall development. Aspects of development such as social and emotional development, which are just as important for school and life success, might be even more strongly related to class size. Future research that seeks to examine class size in relation to other constructs of children's overall development would be very useful.

REFERENCE LIST

Achilles, C.M., Kiser-Kling, K., Aust, A., & Owen, J. (1995, April). A study of reduced class size in primary grades of a fully Chapter 1-eligible school: Success Starts Small. Paper presented at the annual meeting of the American Educational Research Association, San Francisco: CA.

Achilles, C. M., Nye, B.A., & Zaharias, J.B. (1995, April). Policy use of research results: Tennessee's Project Challenge. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Aikens, N., Molduddin, E., Xue, Y., Tarullo, L., and West, J. (2012). Data Tables for child Outcomes and Classroom Quality in FACES 2009 Report. OPRE Report 2011-37b. Washington, DC: Office of Planning, Research, Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education, 70,* 87–107.

Aud, S., Hussar, W., Johnson, F., Kena, G., Roth, E., Manning, E., Wang, X, and Zhang, J. (2012). The Condition of Eucation 2012 (NCES 2012-045). U.S. Department of Education, National Center for Education Statistics. Washington, DC. Retrieved October 27, 2013 from http://nces.ed.gov/pubsearch.

August, D., and Shanahan, T., *Developing Reading and Writing in Second-Language Learners: Lessons from the Report of the National Literacy Panel on Language-Minority Children and Youth* (New York: Routledge, 2008).

Barnett, W.S., Carolan, M.E., Fitzgerald, J., & Squires, J.H. (2011). *The state of preschool 2011: State preschool yearbook*. New Brunswick, NJ: National Institute for Early Education Research.

Barnett, W.S., Yarosz, D., Thomas, J., Jung, K., and Blanco, D. Two-Way and Monolingual English Immersion in Preshool Education: An Experimental Comparison. *Early Childhood Research Quarterly*, 22(3), 277-293.

Barnett, W.S., Schulman, K., and Shore, R. Class Size: What's the Best Fit? (2004) National Institute for Early Education Research. Website: HYPERLINK http://nieer.org/resources/policybriefs/9.pdf

Belfield, CR, Nores, M, Barnett, SW, and L Schweinhart. 2006. Cost-benefit analysis of

the High/Scope Perry Preschool Program using age 40 follow-up data. *Journal of Human Resources*, 41, 283–301.

Berk, L. (1985). Relationship of caregiver education to child-oriented attitudes, of satisfaction, and behaviors toward children. *Child Care Quarterly, 14* (2), 103-129.

Blatchford, P., Bassett, P., & Goldstein, H. (October, 2004). Are class-size differences related to pupils' educational progress and classroom processes? Findings from the Institute of Education class-size study of children aged 5-7 years. *Educational Administration Abstracts,* 39, 4.

Blatchford, P., Edmonds, S., & Martin, C. (January 01, 2003). Class size, pupil attentiveness and peer relations. *The British Journal of Educational Psychology, 73,* 15-36.

Blatchford, P., Goldstein, H., & Mortimore, P. (1998). Research on class size effects: A critique of methods and a way forward. *International Journal of Educational Research, 29*, 691–710.

Blatchford, P., Moriarty, V, Edmonds, S., and Martin, C. (2002). Relationships Between Class Size and Teaching: A Multimethod Analysis of English Infant Schools. *American Educational Research Journal*, 39 (1), 101-132.

Blease, D. (1983). Observer Effects on Teachers and Pupils in Classroom Research. *Educational Review*, 35 (3), 213-217.

Boat, Dinnebel, & Bay (2010). Individualizing Instruction in Preschool Classrooms. *Dimensions of Early Childhood*, 38 (1), 3-11.

Bourke, S. (1986). How Smaller is Better: Some Relationships between Class Size, Teaching Practices, and Student Achievement. *American Educational Research Journal*, 23 (4), 558-571.

Bronfenbrenner, U. 1989. Ecological systems theory. *Annals of Child Development*. Vol. 6, 187-249.

Burchinal, M., Roberts, J., Rhodus, R., Zeisel, S., Neebe, E., & Bryant, D (2000). Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally. *Child Development*, 71 (2), 339-357.

Cahan, Emily D. Past Caring: A History of U.S. Preschool Care and Education for the Poor, 1820–1965, New York: National Center for Children in Poverty, 1989.

Camilli, G., Vargas, S., Ryan, S., & Barnett, W.S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record, 112*(3), 579-620.

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, 6, 42 – 57.

Cawley, R. R., Burrows, M. J., & Goodstein, H. A. (1966). An assessment of intelligence, psycholinguistic abilities and learning aptitudes among preschool children. Storrs, CT: University of Connecticut.

Clarke-Stewart, K.A. Gruber, C.P. & Fitzgerald, L.M. (1994). Children at home and in day care. Hillsdale, NJ: Erlbaum.

Clements, D., & Sarama, J. (2008). Experimental Evaluation of the Effects of a Research-Based Preschool Mathematics Curriculum. *American Educational Research Journal*, 45(2), 443-494.

Clements, D., & Sarama, J., Spitler, M., Lange, A., & Wolfe, C. (2011). Mathematics Learned by Young Children in an Intervention Based on Learning Trajectories: A Large-Scale Cluster Randomized Trial. *Journal for Research in Mathematics Education*, 42 (2), 127-166.

Condry, S. (1983). History and background of preschool intervention programs and the Consortium for Longitudinal Studies. In The Consortium for Longitudinal Studies (Eds.), *As the twig is bent. Lasting effects of preschool programs* (pp. 1–31). Hillsdale, NJ: Erlbaum.

Dunn, L. (1993). Proximal and Distal Features of Day Care Quality and Children's Development." *Early Childhood Research Quarterly* 8: 167–192.

Dunn, L. M., & Dunn, L. M. (1997). Examiner's manual for the PPVT-III Peabody Picture Vocabulary Test. AGS.

Dunn, L.M., Padilla, Lugo & Dunn, L.M. (1986). Test de Vocabulario en Imágenes Peabody (TVIP). Circle Pines, MN: AGS Publishing.

Ehrlich, S., Gwynne, J., Pareja, A., & Allensworth, E. (2013). *Preschool Attendance in Chicago Public Schools: Relationships with Learning Outcomes and Reasons for Absences*. Chicago, IL: The University of Chicago Consortium on Chicago School Research.

Elliott, M. (1998). School finance and opportunities to learn: Does money well spent enhance students' achievement? *Sociology of Education*, 72, 223-245.

Evertson, C., and Folger, J. *Small Class, Large Class: What Do Teachers Do Differently?,* Paper presented at the annual meeting of the American Educational Research Association, San Francisco, March 1989.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard journal on legislation*, 28, 465-498.

Ferguson, R.F., & Ladd, H.F. (1996). How and why money matters: An analysis of Alabama schools. In H.F. Ladd (Ed.), Holding schools accountable: Performance-based reform in education (pp. 256-298). Washington, DC: Brookings Institution.

Finn, J.D., Gerber, S., Achilles, C. & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record*, 103(2), 145-183.

Finn, J.D., and C.M. Achilles. 1990. "Answers and Questions about Class Size: A Statewide Experiment." *American Educational Research Journal*, 27:557-77.

Frede, E.C. (1998). Preschool program quality in programs for children in poverty (pp. 77-98). In W.S. Barnett & S.S. Boocock (Eds.) Early care and education for children in poverty. Albany: SUNY Press.

Gerber, S., Finn, J., Achilles, C., and Boyd-Zaharias, J. (2001). Teacher Aides and Students' Academic Achievement. *Educational Evaluation and Policy Analysis*, 23(2): 123–143

Glass, G.V., Cahen, L.S. Smith, M.L., and Filby, N.N. School class size: Research and policy. Beverly Hills, CA: Sage, 1982

Glass, G. & Smith, M. (1979). Meta-analysis of research on the relationship of class size and achievement. *Educational Evaluation and Policy Analysis*, 1 (1), 2-16.

Glass, M., & Smith, G. Meta-Analysis of Research on Class Size and Its Relationship to Attitudes and Instruction. *American Educational Research Journal*, Vol. 17, No. 4 (1980), 419-433.

Goldstein, H., Yang, M., Omar, R., Turner, R., and Thompson, S. Meta-Analysis Using Multilevel Models with an Application to the Study of Class Size Effects, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 49, No. 3 (2000), 399-412.

Gormley, W., Gayer, T., Phillips, D. & Dawson, B. (2005). The Effects of Universal

Pre-K on Cognitive Development. *Developmental Psychology,* 41, 872-84.

Graue, E., Hatch, K., Rao, K., and Oen, D (2007). The Wisdom of Class Size Reduction, *American Educational Research Journal*, 44, 670-700.

Grissmer, D. W., S. N. Kirby, M. Berends, S. Williamson*, 1994. Student Achievement and the Changing American Family*, Santa Monica, CA: RAND.

Hanushek, E.A. 1999. Some Findings from an Independent Investigation of the Tennessee STAR Experiment and from other Investigations of Class Size Effects. *Educational Evaluation and Policy Analysis*. 21: 143-63.

Hanushek, E.A., Rivkin, S.G. (1997). Understanding the twentieth-century growth in U.S. school spending. *Journal of Human Resources,* 32 (1), 35–68.

Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale-Revised*. New York, NY: Teachers College Press.

Hattie, J. (2005). The paradox of reducing class size and improving learning outcomes. *International Journal of Educational Research, 43*, 387–425.

Howes, C. (1983). Caregiver behavior in center and family day care. *Journal of Applied Developmental Psychology*, 4: 99–107.

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

Hoyles, D. (1998). Constructions of pupil absence in the British education service. *Child and Family Social Work, 3,* 99–111.

Konstantopoulos, S. (2008). Do Small Classes Reduce the Achievement Gap Between Low and High Achievers?"Evidence from Project STAR." *Elementary School Journal,* 108: 275-291.

Konstantopoulos, S. (2011). How Consistent are Class Size Effects? *Evaluation Review,* 35 (1) 71-92.

Krueger, A.B. (1999). Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics*. 114: 497-532.

Lamdin, D. J. (1996). Evidence of student attendance as an independent variable in Education production functions. *The Journal of Educational Research, 89,* 155–162.

LaParo, K.M., Hamre, B. K., Locasale-Crouch, J., Pianta, R. C., et al., (2009). Quality in kindergarten classrooms: observational evidence for the need to increase children's learning opportunities in early education classrooms. *Early Education and Development*, 20, 657-692.

Layzer, J.I., Layzer, C.J., Goodson, B.D., & Price, C. (2007). Subsidy Strategies: Findings from Project Upgrade in Miami-Dade County. Washington, DC. Prepared for U.S. Department of Health and Human Services, Administration for Children and Families, OPRE and CCB. Washington, DC.

LoCasale-Crouch, J., Konold, T., Pianta, R., Howes, C., Burchinal, M., Bryant, D., et al. (2007). Observed classroom quality profiles in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. *Early Childhood Research Quarterly*, *22*, 3–17.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2002). Preschool comprehensive test of phonological and print processing. Florida State University.

Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). TOPEL: Test of preschool early literacy. Austin, TX: Pro-Ed.

Magnuson, Katherine, Marcia Meyers, Christopher Ruhm, & Jane Waldfogel. 2004. Inequality in Preschool Education and School Readiness. *American Educational Research Journal*. 41(1): 115-157.

Marshall, N. (2004). The Quality of Early Child Care and Children's Development. *Current Directions in Psychological Science*, 13, 165-168.

McCartney, K., Scarr, S., Rocheleau, A., Phillips, D., Abbott-Shim, M., Eisenberg, M., Keefe, N., Rosenthal, S., & Ruh, J. (1997). Teacher-child interaction and child-care auspices as predictors of social outcomes in infants, toddlers, and preschoolers. *Merrill-Palmer Quarterly*, 43, 426–450

Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A., & Ehrle, K. (1999). Evaluating the SAGE Program: A Pilot Program in Targeted Pupil-Teacher Reduction in Wisconsin. *Educational Evaluation and Policy Analysis, 21,* 2, 165-177.

Mosteller, F. The Tennessee Study of Class Size in the Early School Grades (1995). The Future of Children, 5 (2), 113-127.

Munton, T., Mooney, A., Moss, P., Petrie, P., Clark, A., & Woolner, J. (2002). Review of international research on the relationship between ratios, staff qualifications and

training, group size, and the quality of provision in early years and child care settings. London: Thomas Coram Research Unit, Institute of Education, University of London.

National Research Council (2001). Eager to Learn: Educating Our Preschoolers. Committee on Early Childhood Pedagogy. Barbara T. Bowman, M. Suzanne Donovan, and M. Susan Burns, editors. Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

NICHD Early Child Care Research Network (1999). Child outcomes when child care center classes meet recommended standards for quality. *American Journal of Public Health*, 89, 1072-1077.

NICHD Early Child Care Research Network (2000). Characteristics and quality of child care for toddlers and preschoolers. *Applied Developmental Sciences*, 4, 116-135.

NICHD Early Child Care Research Network (2002). Child Care Structure – Process – Outcome: Direct and Indirect Effects of Child-Care Quality on Young Children's Development. *Psychological Science*, 13 (3), 199-206.

NICHD Early Child Care Research Network (2004). Does Class Size in First Grade Relate to Children's Academic and Social Performance or Observed Classroom Processes? *Developmental Psychology*, 40 (5), 651-664.

NICHD Early Child Care Research Network and Duncan, G. (2003). Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development, *Child Development*, 74(5), 1454-1475.

Nye, Hedges, and Konstantopoulos (2000). The Effects of Small Classes on Academic Achievement: The Results of the Tennessee Class Size Experiment, *American Educational Research Journal*, 37(1), 123-151.

Nye, Hedges, and Konstantopoulos (2000a). Do the Disadvantaged Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment. *American Journal of Education*, 109 (1), 1-26.

Nye, Hedges, and Konstantopoulos (2001). Are Effects of Small Classes Cumulative? Evidence from a Tennessee Experiment. *The Journal of Educational Research*, 94 (6), 336-345.

Nye, Hedges, and Konstantopoulos (2002). Do Low-Achieving Students Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis*, 24 (3), 201-217.

Oden, S., Schweinhart, L. J., Weikart, D. P., Marcus, S. M., & Xie, Y. (2000). *Into adulthood: A study of the effects of Head Start*. Ypsilanti, MI: High/Scope Press.

Phillipsen, L. C., Burchinal, M. R., Howes, C. & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, 12, 281–303.

Phillips,D.,Mekos,D.,Scarr,S.,McCartney,K.,&Abbott-Shim (2000).Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, 15(4), 475-496.

Pianta, R., La Paro, K. M., & Hamre, B. K. (2008). *Classroom Assessment Scoring System (CLASS)*. Baltimore, MD: Brookes, Inc.

Pianta, R., Mashburn, A., Downer, J., Hamre, B., & Justice, L. (2008). Effects of web-mediated profesional development resources on teacher-child interactions in pre-kindergarten classrooms. *Early Childhood Research Quarterly*, 23 (4), 431-451.

Pianta, R., La Paro, K., Payne, C., Cox, M., Bradley, R. (2002). The Relation of Kindergarten Classroom Environment to Teacher , Family, and School Characteristics and Child Outcomes. *The Elementary School Journal*, 102 (3), 225-238.

Pianta, R.C., Howes, C., Burchinal, M., Byrant, D., Clifford, R., Early, C., et al. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, 9(3), 144-159.

Pierson, Donald E., and Sperber, Robert. The Brookline Early Education Project: An Issue for National Educational Policy. *Childhood Education*, January 1974.

Ramey, S.L., and Ramey, C.T. (2008). Establishing a science of professional development for early education programs: The knowledge application information systems theory of professional development. In L.M. Justice and C. Vukelich (Eds.). *Achieving excellence in preschool literacy instruction*. New York, NY: Guilford Press.

Raver, C.C., Jones, S.M., LiGrining, C.P., & Metzger, M. (2008). Improving Preschool Classroom Processes: Preliminary Findings from a Randomized Trial Implemented in Head Start Settings. *Early Childhood Research Quarterly*, 23, 10-26.

Ready, D., "Class-Size Reduction: Policy, Politics, and Implications for Equity," Education Week, April 9, 2008.

Reynolds, Arthur J. (2000). *Success in Early Intervention: The Chicago Child-Parent Centers,* Lincoln, Nebraska.: University of Nebraska Press, 2000.

Reynolds, A, Temple, J., Robertson, D., & Mann, E. (2002). Age 21 cost-benefit analysis of the Title I Chicago Child-Parent Centers. *Educational Evaluation and Policy Analysis*, 24, 267-303.

Ritchie, S., Weiser, B, Kraft-Sayre, M. & Howes, C. M. & Weiser, B. (2001). Emergent Academics Snapshot Scale. Los Angeles: UCLA (Unpublished Instrument).

Romero, M. & Lee, Y. (2007). *A National Portrait of Chronic Absenteeism in the Early Grades*. New York: National Center for Children in Poverty.

Ruopp, R., Travers, J., Glantz, F., & Coelen, C. (1979). Children at the Center: Final Report of the National Day Care Study. Cambridge, MA: Abt Associates.

Scarr, S., Eisenberg, M., & Deater- Deckard, K. (1994). Measurement of quality in child care centers. *Early Childhood Research Quarterly*, 9, 131–151.

Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R., & Nores, M. (2005). Lifetime effects: The High/Scope Perry Preschool study through age 40 (High/Scope Educatinal Research Foundation Monograph No. 14). Ypsilanti, MI: High/Scope Press.

Siraj-Blatchford, I., Sylva, K., Taggart, B., Sammons, P., & Melhuish, E. (2003) *Intensive case studies of practice across the foundation stage* (EPPE Tech. Paper 10). London: DfES/University of London, Institute of Education.

Slavin, R. (1990) Class size and student achievement: is smaller better? Contemporary Education, 62, 6-12.

Shapson, S., Wright, E., Eason, G., & Fitzgerald, J. (1980). An Experimental Study of the Effects of Class Size. *American Educational Research Journal*, 17 (2), 141-152.

Stecher, Brian , Bohrnstedt, George, Kirst, Michael, McRobbie, Joan, and Williams, Trish (May, 2001). Class-Size Reduction in California: A Story of Hope, Promise, and Unintended Consequences. *The Phi Delta Kappan ,*82, No. 9 (May, 2001), pp. 670-674.

"The Brookline Early Education Project: A 25-Year Follow-up Study of a Family-Centered Early Health and Development Intervention." *Pediatrics*, 2005.

Vandell, D.L. & Wolfe, B. (2002). Child care quality: Does it matter and does it need to be improved? Institute for Research on Poverty, University of Wisconsin-Madison.

Volling, B. L., & Feagans, L.V. (1995). Infant day care and children's social competence. *Infant Behavior and Development*, 18, 177–188.

Vandell, D.L. & Wolfe, B. (2002). Howes, C., Phillips, D.A., & Whitebook, M. (1992). Thresholds of Quality: Implications for the social development of children in center-based child care. *Child Development*, 63, 449–460.

Walston, J. & West, J. (2004). Full-day and half-day kindergarten in the United States: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99. Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Westinghouse Learning Corporation and Ohio University. (1969). *The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development* (Vols. 1 and 2, Report to the Office of Economic Opportunity). Athens: Author.

Whitaker, S., Kinzie, M., Kraft-Sayre, Mashburn, A., & Pianta, R.C. (2007). Use and evaluation of web-based professional development for teacher of at-risk preschool children. *Educational Technology & Society*, 9 (4), 194-204.

Whitehurst, G.J., "Class Size: What Research Says and What it Means for State Policy," 2011.

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. A. (2001). Woodcock–Johnson III (WJ-III). Itasca, IL: Riverside.

Woodcock, R. W., Munox-Sandoval, A. F., McGrew, K. S., Mather, N. (2005). Batería III Woodcock-Muñoz. Rolling Meadows, IL: Riverside Publishing.

Word, E., Johnston, J., Bain, H.P., et al. *Student/Teacher Achievement Ratio (STAR): Tennessee's K–3 class size study.* Final summary report 1985–1990. Nashville: Tennessee Department of Education, 1990.

Zigler, E., & Berman, W. (1983). Discerning the future of early childhood intervention. *American Psychologist*, 38, 894 – 906.

Zinth, K. (2009). "State Policies focusing on Class-size Reduction." Education Commission of the States (ECS).

VITA

Jessica Francis was born and raised in Point Pleasant, New Jersey. Before attending Erikson Institute for her doctorate, she attended the University of Virginia, where she earned a Bachelor of Arts in Economics and a Master of Education.

While at Erikson, Francis was awarded the James B. Harris Leadership scholarship. She also consulted on several research projects conducted by Erikson Institute and the Chicago Public Schools.

Currently, Francis is a Research Coordinator for the National Institute for Early Education Research at Rutgers University, where she manages several research studies and has served as Principal Investigator for one study. She lives in Brick Township, New Jersey.