



Full-day kindergarten and student achievement: A literature review

Andrea Lash
Soung Bae
Vanessa Barrat
Elizabeth Burr
Tony Fong

December 2008

A REL West Technical Assistance Response

This Technical Assistance Response was prepared under Contract ED-06-CO-0014, from the U.S. Department of Education, Institute of Education Sciences, by the Regional Educational Laboratory West, administered by WestEd. The content of the publication does not necessarily reflect the views or policies of IES or the U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. For more information about the national REL system, visit <http://ies.ed.gov/ncee/edlabs/>.

Abstract

The review of literature reported in this publication was conducted in response to a request by the Nevada Legislative Counsel Bureau for an update on recent research into full-day kindergarten. The review looked at research reported in the last decade (1998–2008) that examined the impact of full-day kindergarten, as compared with half-day programs, on student achievement in reading and mathematics. The intent was twofold: to examine and summarize the quality of the research designs of studies being reported and to summarize the findings of those studies whose research designs were determined to be of adequate quality. In addition to providing these summaries, this publication explains how the quality of a study's research design influences the strength of the study's findings.

The review began by identifying 299 unduplicated references of literature on full-day kindergarten published in the last decade. Of these, 21 passed an initial screening designed to identify research reports that examined academic outcomes of full-day kindergarten programs as compared with half-day programs. A systematic review of those 21 reports identified 11 that described research designs of sufficient rigor to be included in the summary of research findings, though none of them used randomized assignment for placing students in a kindergarten program, which would be necessary to establish a causal relationship between program type (i.e., full- or half-day) and academic outcomes.

The 11 reports represent only six studies because, of the 11 reports, 6 analyzed data from a single national survey. Because of their research designs, none of the six studies can provide strong evidence to judge the effect of full-day kindergarten. Their findings may suggest *associations* between students' participation in full-day kindergarten and achievement, but they cannot identify whether the link between program and achievement is caused by kindergarten type or is due to other factors that might distinguish students in full- and half-day programs (e.g., students' opportunities to learn outside of school).

The studies are consistent in suggesting a relationship between attendance in full-day kindergarten and higher levels of early reading skills as measured at the end of kindergarten. Studies that examined other reading outcomes at later points in time did not find the same association. Some study features that could produce different reading results for kindergarten and later grades include changes in: the study population after kindergarten (caused, for example, by students moving away); the outcomes that are measured (e.g., aspects of decoding for kindergarten students versus reading comprehension for students in higher grades); and students' post-kindergarten instructional experiences. The research does not provide sufficient information from which to draw conclusions about the relationship between mathematics achievement and attendance in full-day kindergarten.

Future studies could be strengthened by random assignment of students to programs. Even without randomization, studies could be improved if students were assigned in a manner that ensured academic equivalence of the two groups at the start of kindergarten. In addition to the question of programs' learning effects, other questions of interest that could be addressed in rigorous studies include: 1) What features of kindergarten instruction promote learning? 2) Are there advantages to developing early reading skills in kindergarten? 3) What changes may need to be made to the instruction in first grade if students enter better prepared from kindergarten?

Introduction

Since at least the 1960s, kindergarten in the United States has been offered as either a half-day or a full-day program (DeCicca, 2005). Over the last several decades, the availability of the full-day program has grown, with the proportion of kindergarteners attending full-day programs increasing from 10 percent in the 1970s (Puleo, 1988) to almost 60 percent as of 2008 (Education Commission of the States, 2008).

Policymakers have long debated the relative benefits of each type of program. On the one hand, the extra teaching time available in a full-day program may allow teachers to target instruction to meet the learning and behavioral needs of individual students, and at a pace that may be less stressful for young learners. The expectation is that the longer day can, thus, result in such positive student outcomes as greater achievement, more cooperative behavior, better preparation for first grade, and less chance of retention or placement in special education. On the other hand, compared to a full-day kindergarten, a half-day program serving the same number of children typically requires fewer teachers, less classroom space, and fewer financial resources. In addition, half-day programs may protect young students from the potential stress of having to be in school all day, or from being engaged in academics for long periods of time before students may be developmentally ready for it. Today, broad consensus about the need to improve student outcomes has intensified the kindergarten debate. Recognizing the importance of students' early experiences for their later education success (Shonkoff & Phillips, 2000), policymakers are looking closely at the benefits and costs of these early education programs.

In Nevada, where the state legislature has been considering adoption of full-day kindergarten, proponents and critics alike have presented research findings that support their respective positions. Because both proponents and opponents have been able to find research that seems to support their position, one suggestion has been to assume that the position with the greatest number of research citations on its "side" has greater credibility.¹ Another suggestion, based on the belief that political bias can influence research conclusions, is that research findings should be interpreted in light of the authors' organizational affiliations.² Yet neither of these suggestions gets to the heart of the matter: What is the quality of the research being discussed?

This Technical Assistance Response prepared at the request of the Nevada Legislative Counsel Bureau, is intended to serve as an aid for evaluating research evidence from recent (i.e., reported from 1998–2008) studies of full-day kindergarten. To that end, it first describes the features of research designs that help determine the strength of the evidence the research can provide. It then explains how REL West analysts identified and screened studies for inclusion in a summary of the recent research that examines the effect of attendance in full-day kindergarten, as compared with half-day kindergarten, on student achievement in reading and mathematics. Finally, based on the studies that were included, it summarizes their designs and findings.

The value of looking closely at research designs

A study's research design determines the quality and strength of the evidence it can provide and, thus, the confidence we can have in the study's conclusions. Rigorous designs may be compromised by poor implementation, such as the use of invalid tests to measure student learning. But even the best implementation cannot improve a study that has been poorly designed. Thus, in examining evidence from research studies we begin by examining their research designs.

A design feature of particular importance to studies of full- and half-day kindergarten programs and their respective student learning outcomes is how well the research design ensures equivalence of the students in the programs being studied. If a study is intended to compare student learning in different kindergarten programs,

¹ <http://elkodaily.com/articles/2007/05/16/opinion/editorial/opinion2.txt>

² <http://www.rgj.com/apps/pbcs.dll/article?AID=2007701210328>

ideally the students in the programs being compared should be equivalent at entry to kindergarten on all factors that could influence learning outcomes. Among such factors may be, for example, students' age, their preparation for kindergarten, and their opportunities to learn outside of school. If student populations from program to program are not comparable in all ways, their differences at the beginning of kindergarten could lead to differences in outcomes, which, in turn, could be interpreted incorrectly as evidence of differences in program effectiveness.

The challenge for researchers seeking to understand the effects of programs, such as full- or half-day kindergarten, is to design a study that does not allow a comparison of learning outcomes to be influenced by any factor other than the programs being studied. This requires assigning students to programs in a way that balances, or controls, all other possible influences. If students were randomly (i.e., by chance alone) assigned to their kindergarten program, there would be no reason to expect differences between the group of full-day students and the group of half-day students.³ In this case, any outcome differences that resulted between these groups of students could be attributed to program differences and not to differences that existed between the groups at the start of kindergarten. Thus, studies that use random assignment, and that also maintain quality in implementation of the study procedures, provide the strongest level of evidence to support a conclusion that the type of kindergarten a student attends leads to, or causes, differences in student learning. In this review, we refer to designs that use random assignment and, thus, can provide evidence of a *causal link* between programs and learning, as Level 1 (see figure 1).

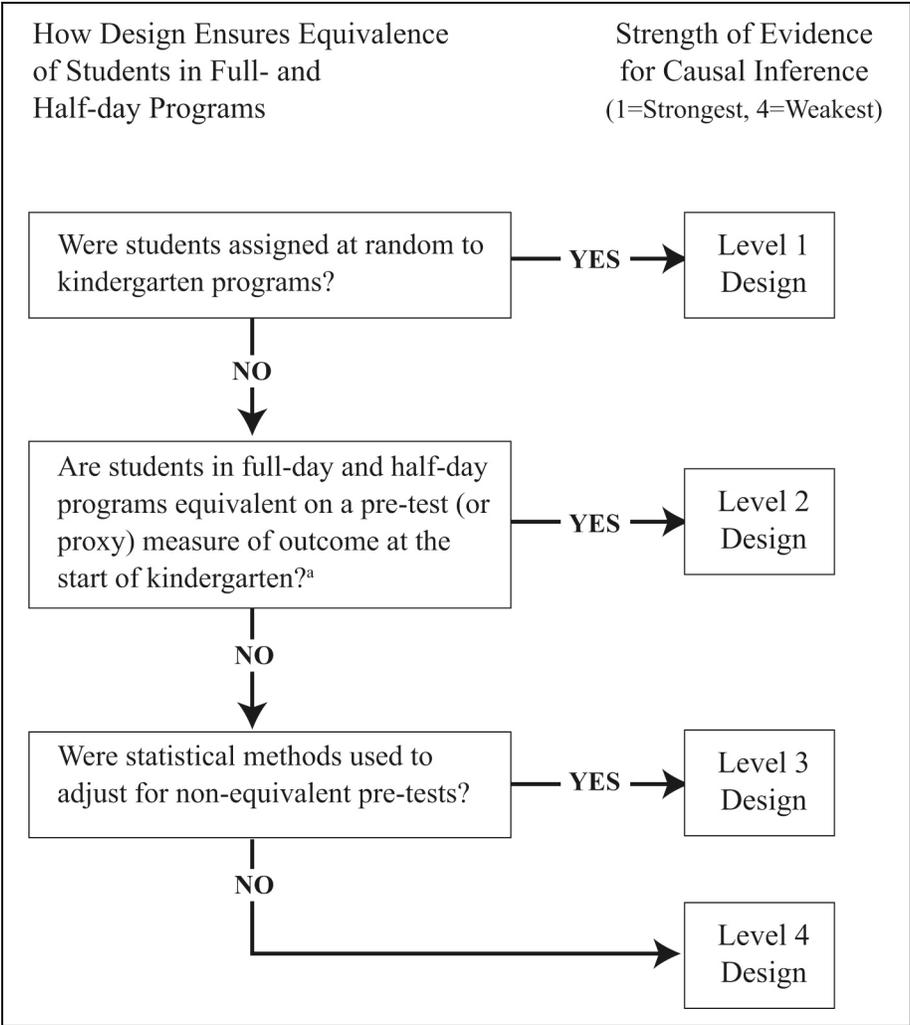
In studies using a Level 1 design, we can assume that any difference in student outcomes between programs has been caused by the programs, because all other variables that might have contributed to the variation were balanced by students' random assignment to programs. This is not the case for studies using a Level 2 or Level 3 design. Because neither of these designs uses random assignment, they can yield evidence that a *relationship* exists between program type and outcomes but cannot identify the *cause* of that relationship. Have differences in the kindergarten programs caused the different outcomes or, instead, are the outcomes caused by some difference in the students in each program (e.g., do the students in one program have more opportunities to learn outside of school)?

Levels 2 and 3 differ from each other in the degree to which they can support any conclusion of a relationship between kindergarten program and outcomes. This is because they differ in how well they ensure that, at the start of kindergarten, the program groups are equivalent in achievement level (though not necessarily equivalent in other factors that Level 1 studies would balance, such as students' opportunity to learn outside of school). To infer a relationship between kindergarten program and learning outcomes at the end of kindergarten by looking at outcome differences between students in full-day and half-day programs, the students being studied in each program would need to have started at the same academic level. Level 2 designs ensure this equivalence by matching each half-day student with a full-day student who has a similar score on a pre-test of achievement at the start of the study. Or they demonstrate equivalence by showing that the average pre-test scores for full-day and half-day students are the same. Level 3 designs have groups that are not equivalent at pre-test, but in analyzing outcomes the researchers use a statistical approach that reduces the effect of the pre-test difference.

Finally, studies using Level 4 designs do not have equivalent groups at pre-test and do not attempt to adjust outcomes for these pre-test differences. The results from Level 4 studies provide little evidence to support conclusions about program effects. In other words, their findings are weaker than those found in studies using a Level 3 design, which, in turn, are weaker than those using a Level 2 design, which, in turn, are weaker than those using a Level 1 design.

³ Random assignment creates groups that are "probabilistically" similar to each other. Because each student has an equal chance of being in full-day (or half-day) kindergarten, any differences between the groups will be due to chance alone. Factors that could affect outcomes, such as high motivation or greater opportunities to learn outside school, are unlikely to appear in one group more often than the other. Chances are small that large differences will occur between groups even on factors that are unknown to influence outcomes.

Figure 1. Classification of research designs according to their strength of evidence for causal inferences



^a For the purpose of this review, we judge groups to be equivalent prior to treatment if no significant differences are found in a statistical test that contrasts the mean pre-test (or proxy) measure of outcome for the two groups. For small samples (< 30) we also require that the standardized mean difference between groups be less than .50.

Those conducting reviews of research, like the authors of this publication, differ in the quality of design they require for studies to be included in the review. Some reviewers include only Level 1 studies; others may include non-randomized experiments from Levels 2 and 3 (see Higgins & Green, 2008; Shadish & Myers, 2004; What Works Clearinghouse guidelines⁴). Our intent was to exclude research using a Level 4 design and summarize results for Level 1 to Level 3 studies. As it turned out, however, no Level 1 or Level 2 studies were found. The implications of this, as the research summary will show, is that the current research on full- and half-day kindergarten does not (i.e., cannot) provide strong evidence to judge the effects of these programs on student learning.

How research reports were identified and reviewed for this publication

To identify relevant research, we searched the following social science databases: ERIC, Education Research Complete, Academic Search Premier, Applied Social Sciences Index and Abstracts, ProQuest Journals and Education: a Sage Full-Text Collection using the search terms “full day kindergarten,” “all day kindergarten,” and “extended day kindergarten.” In addition, we used web search engines (Google and Google Scholar) and searched organization websites for information (i.e., Education Commission of the States, Phi Delta Kappa International, RAND, and Public Policy Institute of California). With duplications removed, the search identified 298 references. We then examined the bibliographies of literature reviews and published research reports included in the 298 references to uncover additional references that may not have appeared in the database search. One additional report was found this way, bringing the total number of unduplicated references to 299.

THE FIRST STAGE OF REVIEW

Two readers read document abstracts from these 299 references to determine if they met the following six criteria for inclusion in the literature review:

- The intervention being studied was full-day kindergarten.
- The study population was kindergarten students.
- The publication date was in the period of 1998–2008.
- The publication type was a primary report of research on the impact of full-day kindergarten.⁵
- The outcome being examined was academic achievement.
- The study location was the United States.

In complete agreement with one another, the readers identified 21 of the original 299 references as meeting all of the criteria.

THE SECOND STAGE OF REVIEW

In the next stage of review, each of the 21 reports was read by two researchers with graduate-level training in quantitative research methods and research design. Researchers followed a detailed review protocol that focused on four aspects of each study:

- Study features, which gathers information about the study population, intervention, and outcomes in order to confirm that the study meets criteria for inclusion in the review;
- Research design, which identifies information needed to judge the rigor of the research plan;
- Study implementation problems, and other threats to study validity, which examines how the study guards against common sources of bias that can create misleading or erroneous results; and
- Statistical reporting, which describes the nature of the outcome variables and how the findings are reported.

⁴ For What Works Clearinghouse guidelines, visit <http://ies.ed.gov/ncee/wwc/references/iDocViewer/Doc.aspx?docId=2&toId=3>

⁵ For instance, news editorial and literature reviews and policy briefs were not included in the population of reports to be reviewed.

For each report, the reviewers summarized “study features” by indicating whether or not the study met criteria for inclusion in the review and, in doing so, they agreed on 20 of the 21 studies (95%). These reviewers summarized “research design” information by classifying the study design into one of the four levels described above and, in their rating of design, they agreed on 16 of the 21 studies (76%). Where the reviewers’ judgments were not identical, a third researcher, who did not have knowledge of the judgments, reviewed the report and provided a tie-breaking judgment. The detailed information gathered in the third and fourth sections — on sources of bias and on statistical reporting, respectively — did not lend itself to a simple summary rating. If information recorded by the two reviewers for those sections was not in agreement, the reviewers met to discuss and reconcile their differences. Appendix A provides descriptions of the study samples, programs studied, and achievement measures from all of the 21 reports.

The evidence available from recent studies of full-day and half-day kindergarten

SUMMARY OF RESEARCH RIGOR

Ultimately, 11 of the original 299 references met all the criteria and have been included in the research summary for this publication. Of the 21 reports that were identified for review after the first screening, 8 were excluded subsequently because they used Level 4 designs. A 9th report was excluded because, while it examined full-day kindergarten, it did not compare full-day and half-day programs; and a 10th report was excluded because it presented inconsistent findings that could not be resolved with the data provided.⁶ Collectively, the final 11 reports represent six different studies (see table 1), because 6 of them summarize analyses of the same study: the Early Childhood Longitudinal Study (ECLS), which followed a national sample of students from the kindergarten class of 1998–1999 into elementary school.⁷

Studies of full-day kindergarten whose results have been reported within the last decade cannot provide strong evidence with which to judge the effects of a full-day program on student achievement. This is because the studies have used weak research designs, with the highest level of design among them being Level 3 (see table 1). Unlike Level 1 studies, they did not control for factors other than the kindergarten program that can affect outcomes. Instead, the reports summarized here used statistical methods to obtain partial control for pre-test differences that existed between groups. Their findings suggest associations between attendance in full-day kindergarten and learning outcomes, but they cannot identify reasons for any observed associations.

SUMMARY OF THE RESEARCH RESULTS

All 11 reports included in the research summary examined reading outcomes (see table 1). Mathematics outcomes were studied in the 6 reports of the ECLS and in 2 other reports. All but 1 report examined kindergarten outcomes; outcomes for later grades were examined in the 6 reports of the ECLS and in 2 other reports. The 6 reports that analyzed data from the ECLS identified students in the sample who had attended full-day kindergarten and contrasted them with students who had attended half-day kindergarten. Each of the other 5 reports summarized an evaluation of a full-day program in a local school or district.

The findings of the 11 reports are summarized in table 2 for measures of reading-related outcomes and in table 3 for mathematics outcomes. For each study, tables 2 and 3 identify the outcomes measured and indicate if there was a difference between the average scores of full-day and half-day students. Differences are only identified if

⁶ Attempts to obtain the needed information directly from the authors through email and phone were not successful.

⁷ <http://nces.ed.gov/ECLS>

they are greater than would be expected by chance alone.⁸ They are noted by identifying the group (i.e., Full or Half) that scored higher. (Note that in each table, the 6 reports on the ECLS are summarized as one study.⁹)

With regard to reading, the research reports published in the last decade are consistent in suggesting a relationship between attendance in full-day kindergarten and higher levels of early reading skills (e.g., phonemic awareness, letter-sound correspondence). Studies that examined reading achievement at higher grades (e.g., reading comprehension) did not find the same association with the kindergarten program (see table 2).

Table 1. Reports identified for review, by those included in the summary and those excluded from the summary

Reports included in the research summary	Study location or data source	Outcomes ^a		Design level	
		Kindergarten	Grade 1–5		
1 Alban, Nielson, & Schatz (2003)	Suburban District in Maryland	R	R	3	
2 Baskett, Bryant, Rhoads, & White (2005)	Rural-Suburban District in Maine	R		3	
3 Hildebrand (2001)	District in Nebraska	R,M		3	
4 Wolgemuth, Cobb, Winokur, Leech, & Ellerby (2006)	Suburban School in Colorado	R,M	R,M	3	
5 Zvoch, Reynolds, & Parker (2008)	Large District in Southwest	R		3	
6 Cannon, Jacknowitz, & Painter (2006)	<i>Early Childhood Longitudinal Study</i>	R,M	R,M	3	
7 DeCicca (2007)	<i>Early Childhood Longitudinal Study</i>	R,M	R,M	3	
8 Le, Kirby, Barney, Sctodji, & Gershwin (2006)	<i>Early Childhood Longitudinal Study</i>		R,M	3	
9 Lee, Burkham, Ready, Honigman, & Meisels (2006)	<i>Early Childhood Longitudinal Study</i>	R,M		3	
10 Votruba-Drzal, Li-Grining, & Maldonado-Carreno (2008)	<i>Early Childhood Longitudinal Study</i>	R,M	R,M	3	
11 Walston & West (2004)	<i>Early Childhood Longitudinal Study</i>	R,M		3	
Reports excluded from the research summary		Reason for exclusion			
12 Colvin & Salkind (2005)	Contains inconsistent report of research results				
13 del Gaudio Weiss, & Offenberg (no date)	Level 4 Design				
14 Denton, West, & Walston (2003)	Level 4 Design				
15 Gullo (2000)	Level 4 Design				
16 Larson (2003)	Level 4 Design				
17 Plucker, Eaton, Rapp, Lim, Nowak, Hansen, & Bartleson (2004)	Level 4 Design				
18 Reynolds, Temple, Robertson, & Mann (2001)	Did not compare full-day to half-day kindergarten				
19 Saam & Nowak (2005)	Level 4 Design				
20 Stofflet (1998)	Level 4 Design				
21 Wang & Johnstone (1999)	Level 4 Design				

^a Cell entries indicate the outcomes examined in the report at the end of kindergarten and at the end of one or more other elementary grades. R=reading outcomes, M=math outcomes.

⁸ A difference between groups is considered to be greater than expected by chance alone if a statistical test shows it to be greater than would be expected 5 times out of 100; that is, that the statistical test was significant at $p \leq .05$.

⁹ When reports examine data from the same sample, as do the 6 reports on the Early Childhood Longitudinal Study, their findings are not independent sources of evidence to support conclusions about program effects, even if the studies examine different outcomes. This is because, if the study sample has an unusual feature that skews results in one report, for example, the chances are high that the unusual feature would bias results of other reports as well. If the research summary were to treat the reports as if they were independent of one another, and include all of them separately, any sample feature that skewed results for the individual reports could bias the whole summary of research. For this reason, we consider the 6 reports as a report of one study, and have summarized and represented them as such in tables 2 and 3. (See appendix B for the results of the separate reports of the Early Childhood Longitudinal Study.)

Table 2. Higher scoring program (full-day vs. half-day) in analyses of reading outcomes from Level 3 studies

Report No.	Outcome	Grade tested					
		Kindergarten	First	Second	Third	Fourth	Fifth
3	Early Reading Skills	Full					
5	Early Reading Skills (growth)	Full ^a					
2	Reading Level (growth)	X					
	Primary Literacy Skills (growth)	Full ^a					
1	Early Reading Skills	Full					
	Reading Achievement			X			
4	Oral Reading Fluency, end of year	Full	Full	X			
	Oral Reading Fluency, start of year		Full	X			
	Reading Achievement			X		X	
ECLS ^b	Reading	Full	X		X		Half

Note: Report number is from table 1. Cell entries name the higher-scoring program. An X indicates that the difference between programs was tested but was not greater than one would expect by chance alone. A shaded cell indicates the research report did not investigate differences at this grade level.

^a This finding compared student growth or gain in test score.

^b Information based on results reported in six separate studies of the Early Childhood Longitudinal Study data (see appendix B).

With regard to mathematics, the research conducted in the past decade does not provide sufficient information from which to draw conclusions about the relationship between student achievement and attendance in full-day kindergarten. Only three of the studies examined mathematics outcomes; they did not all examine them at the same grade level, and the findings do not replicate across studies (see table 3).

Table 3. Higher scoring program (full-day vs. half-day) in analyses of mathematics outcomes from Level 3 studies

Report No.	Outcome	Grade tested					
		Kindergarten	First	Second	Third	Fourth	Fifth
3	Early Math Knowledge	X					
4	Math Achievement			X	X	X	
ECLS ^a	Math	Full	X		X		Half

Note: Report number is from table 1. Cell entries name the higher-scoring program. An X indicates that the difference between programs was tested but was not greater than one would expect by chance alone. A shaded cell indicates the research report did not investigate differences at this grade level.

^a Information based on results reported in six separate studies of the Early Childhood Longitudinal Study data (see appendix B).

Examples of how study implementation can affect the quality of evidence

Findings in this summary for short-term and long-term reading outcomes may appear to be inconsistent. Results show positive relationships of full-day kindergarten to early reading skills measured at the end of kindergarten, but not to other reading outcomes measured in later grades (see table 2). Someone might suggest that this is because the relationship between full-day kindergarten and reading achievement changes with time, suggesting a loss or fading of the effects of full-day kindergarten. But many other factors can lead to these results, including the original weak designs that did not begin with equivalent groups of students in full-day and half-day programs. Before drawing conclusions from a comparison of two research findings, it is important to determine if the findings are based on comparable data. The way the design of a study was implemented can cause early and later data to lack comparability. Some questions to consider include:

1. *Are the results derived from the same student population?*

In longitudinal studies that follow students across grades, the population being studied can change over time as students transfer schools, miss a test administration, or are otherwise lost to follow-up. If the students who leave a study differ systematically from those who remain, the population change alone can produce different results for different grades. Additionally, smaller sample sizes make statistical analyses less sensitive to differences between groups and, thus, less likely to identify a group difference as being greater than expected by chance alone at the later grades.

There is some evidence that the populations changed over time in the studies included in this literature review. Wolgemuth, Cobb, Winokur, Leech, and Ellerby (2006) reported that one half of the students in their study who provided kindergarten data were not included in the second-grade achievement testing, and close to three fourths were not included in the fourth-grade testing. In the Early Childhood Longitudinal Study, close to 40 percent of the students had some missing data by the fifth grade and, thus, could not be included in the analyses at every grade level unless their data were estimated statistically (Le, Kirby, Barney, Setodji, & Gerswin, 2006).

2. *Are the results summarizing the same outcomes?*

Studies may use the same label for outcomes — “reading achievement,” for example. But what is measured as reading achievement may not be the same on all occasions. The studies of full-day kindergarten examined early reading skills at the end of kindergarten. These skills include student knowledge of letters of the alphabet, letter-sound correspondence, phonemic awareness, and oral reading fluency (i.e., the number of words a student can read correctly in one minute from a grade-appropriate text), all of which provide information about students’ decoding skills. But tests of reading achievement administered in studies in later grades assessed students’ comprehension of grade-appropriate text. While early reading skills and reading comprehension are related and may develop sequentially, they are not the same. Knowing that students are skilled decoders does not indicate they will be skilled in comprehension since other abilities, in addition to decoding, influence students’ understanding of text.

3. *Are students’ instructional experiences comparable?*

Studies of kindergarten effectiveness that measure student outcomes years after the students finish kindergarten assume that the post-kindergarten education experiences are equally effective for all students. Yet, this is unlikely to be the case. Students’ skill levels relative to their peers can change as they move up the grade levels depending on the nature of the instruction they receive. Some teachers focus their instruction on lower-ability students in an attempt to develop their skill levels. Others may individualize instruction and provide all students with tasks tailored to their needs. And still others may emphasize further development of the higher-ability students’ skills. Changes that occur in students’ standing relative to their peers as a result of the instruction they receive after

kindergarten can affect the relationships observed between full-day kindergarten and student outcomes that are measured in later grades.

Looking ahead

To obtain evidence of any *effect* of full- versus half-day kindergarten on student achievement (rather than obtaining evidence that there is a *relationship* between program and achievement) requires research studies that randomly assign students to kindergarten programs. Such studies could be conducted as pilots or try-outs of new programs. Even without randomization, studies could be improved if students were assigned to full-day and half-day kindergarten in a manner that ensures the academic equivalence of the two groups at the start of kindergarten. In addition to the question of programs' learning effects, other questions of interest could be addressed in rigorous studies to provide information for policymakers and practitioners, such as: 1) What features of kindergarten instruction promote learning? 2) Are there advantages to developing early reading skills in kindergarten? 3) What changes may need to be made to the instruction in first grade if students enter better prepared from kindergarten?

References

- Alban, T., Nielson, J., & Schatz, C. (2003). *Evaluation of the longitudinal impact of comprehensive early childhood initiatives on student academic achievement*. Rockville, MD: Office of Shared Accountability.
- Baskett, R., Bryant, K., Rhoads, K., & White, W. (2005). Half-day to full-day kindergarten: An analysis of educational change scores and demonstration of an educational research collaboration. *Early Child Development and Care, 175*(5), 419–430.
- The Campbell Collaboration, <http://nces.ed.gov/ecls/kindergarten.asp>
- Cannon, J. S., Jacknowitz, A., & Painter, G. (2006). Is full better than half? Examining the longitudinal effects of full-day kindergarten attendance. *Journal of Policy Analysis and Management, 25*(2), 299.
- Colvin, M. R., & Salkind, N. J. (2005). Factors promoting academic gains in kindergarten: Parental attitude and length of school day. *Journal of Early Childhood Education & Family Review, 12*(3; 3), 5–13.
- DeCicca, P. (2007). Does full-day kindergarten matter? Evidence from the first two years of schooling. *Economics of Education Review, 26*(1), 67–82.
- del Gaudio Weiss, A., & Offenberg, R. (n.d.). *Differential impact of type of kindergarten experience on academic achievement and cost benefit through grade 4: An examination of four cohorts in a large, urban school district*. Philadelphia, PA: School District of Philadelphia.
- Denton, K., West, J., & Walston, J. (2003). *Reading — Young children's achievement and classroom experiences*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Available at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2003070>. Last accessed on December 16, 2008.
- Education Commission of the States. (2008). *State statutes regarding kindergarten: Policies concerning district offering of and student attendance in full- and half-day kindergarten programs*. Available at <http://www.ecs.org/ecsmain.asp?page=/html/issue.asp?issueID=77>. Last accessed on December 16, 2008.
- Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly, 12*(4), 459–80.
- Gullo, D. F. (2000). The long-term educational effects of half-day vs. full-day kindergarten. *Early Child Development and Care, 160*, 17–24.
- Higgins J. P. T., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions* [Version 5.0.1, updated September 2008]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.
- Hildebrand, C. (2001). *Effects of three kindergarten schedules on achievement and classroom behavior*. Phi Delta Kappa Center for Evaluation, Development, and Research.
- Larson, J. (2003). *Reducing the school performance gap among socio-economically diverse schools: Comparing full-day and half-day kindergarten programs*. Rockville, MD: Office of Shared Accountability.
- Le, V., Kirby, S. N., Barney, H., Setodji, C. M., & Gershwin, D. (2006). *School readiness, full-day kindergarten, and student achievement: An empirical investigation*. Santa Monica, CA: RAND Corporation.
- Lee, V. E., Burkam, D. T., Ready, D. D., Honigman, J., & Meisels, S. J. (2006). Full-day versus half-day kindergarten: In which program do children learn more? *American Journal of Education, 112*(2), 163–208.
- Plucker, J. A., Eaton, J. J., Rapp, K. E., Lim, W., Nowak, J., Hansen, J. A., et al. (2004). *The effects of full day versus half day kindergarten: Review and analysis of national and Indiana data*. Bloomington, IN: Center for

Evaluation and Education Policy. Available at <http://ceep.indiana.edu/projects/project.php?id=76&category=3>. Last accessed December 16, 2008.

- Puleo, V. T. (1988). A review and critique of research on full-day kindergarten. *Elementary School Journal*, 88(4; 4), 427–439.
- Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (2001). Long-term effects of an early childhood intervention on educational achievement and juvenile arrest: A 15-year follow-up of low-income children in public schools. *Journal of the American Medical Association*, 285(18), 2339–2346.
- Saam, J., & Nowak, J. A. (2005). The effects of full-day versus half-day kindergarten on the achievement of students with Low/Moderate income status. *Journal of Research in Childhood Education*, 20(1), 27.
- Shadish, W., & Myers, D. (2004). *Campbell Collaboration research design policy brief*. Campbell Collaboration. Available at http://camp.ostfold.net/artman2/uploads/1/Research_Design_Policy_Brief.pdf. Last accessed on December 16, 2008.
- Shonkoff, J. P., & Phillips, D. A. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington DC: National Academy Press.
- Stofflet, F. P. (1998). *Anchorage school district full-day kindergarten study: A follow-up of the kindergarten classes of 1987–88, 1988–89, and 1989–90*. Anchorage, AK: Anchorage School District.
- Villegas, M. (2005). *Full-day kindergarten: Expanding learning opportunities*. San Francisco: WestEd. Available at http://www.wested.org/online_pubs/po-05-01.pdf. Last accessed on December 16, 2008.
- Votruba-Drzal, E., Li-Grining, C. P., & Maldonado-Carreño, C. (2008). A developmental perspective on full-versus part-day kindergarten and children's academic trajectories through fifth grade. *Child Development*, 79(4), 957–978.
- 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* (NCES 2004-078). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Wang, Y. L., & Johnstone, W. G. (1999). Evaluation of a full-day kindergarten program. *ERS Spectrum*, 17(2), 27.
- Wolgemuth, J. R., Cobb, R. B., Winokur, M. A., Leech, N., & Ellerby, D. (2006). Comparing longitudinal academic achievement of full-day and half-day kindergarten students. *Journal of Educational Research*, 99(5), 260–269.
- Zvoch, K., Reynolds, R. E., & Parker, R. P. (2008). Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference? *Early Childhood Research Quarterly*, 23(1), 94–107.

Appendix A

Table A1: Study population and brief description for 21 reports identified for review

Table A2: Study samples and outcome variables for 21 reports identified for review

Table A1. Study population and description for 21 reports identified for review

Study	Population of study	Brief description (when provided)
Alban & Schatz (2003)	Students enrolled in the 34 neediest schools (as defined by proportion of English language learner students and those eligible for free and reduced-price lunch) in a suburban Maryland school district (Montgomery County Public Schools)	In 2001, Montgomery County Public Schools launched its Kindergarten Initiative, which includes FDK, reduced class size, revised kindergarten curriculum, ongoing professional development through the summer and school year, and increased communication between parents and school staff. All elements were in place in kindergarten in 2001 except FDK, which was phased into 56 schools sequenced from lower to higher socioeconomic status and educational need. FDK implementation began with 17 schools in Phase 1 (identified as having the highest levels of need), 17 more schools in Phase 2 (identified as next most in need), and 22 schools in Phase 3. Second-grade students who attended FDK were compared to students in the same schools in three previous consecutive academic years and to students from the same cohort who attended Phase 2 schools.
Baskett, Bryant, Rhoads, & White (2005)	228 students from 6 elementary schools in the Auburn School District, a rural-suburban district in south-central Maine, with an average free and reduced-price lunch rate of 43 percent	In a post-hoc design, the study compares the effects of FDK vs. HDK. Students were compared in 2 consecutive academic years, not simultaneously. The first group was in the last year of HDK and the second group was in the first year of FDK. The length of the half-day classes was 3 hours per day, and the average length of the full-day classes was 5 hours and 50 minutes.
Cannon, Jacknowitz, & Painter (2006)	ECLS-K: Students included in this analysis are those who have ECLS-K parent interview data for all 4 rounds of data collection (fall and spring of K, spring 1st grade, and spring 3rd grade); were in a regular K class (one-year K class primarily for 5 year olds prior to 1st grade); and were in the same type of K program throughout the year	Longitudinal study of national Kindergarten cohort of 1998–99
Colvin & Salkind (2005)	Kindergarteners in a suburban Midwestern school district serving a mostly white, middle-class population	
DeCicca (2007)	ECLS-K: Sample restricted to public school children who were first-time kindergarteners and who did not switch schools during the K year	Longitudinal study of national Kindergarten cohort of 1998–99
del Gaudio Weiss, & Offenberg (no date)	Students in a large, urban, public school district (Philadelphia) serving a high percentage of low-income students, excluding special education students and students who withdrew from school between K and 3rd grade	Beginning in 1995–96, the district changed from HDK to FDK. This study compares students who attended the same schools over 4 consecutive years, to examine the relationship between type of K and subsequent achievement. The first cohort includes children who were in HDK or no K the previous year. The other two cohorts attended FDK or no K.
Denton, West, & Walston (2003)	ECLS-K: Study focuses on students' reading skills in their first 2 years of school	

Study	Population of study	Brief description (when provided)
Gullo (2000)	Kindergarten students in a large Midwestern school district who attended either full- or half-day K and were continuously enrolled through the 2nd grade, excluding children who were identified prior to K as having special needs	HDK and FDK were offered to all students in the district. Students in FDK and HDK were combined and subsequently randomly assigned into 1st and 2nd grade classrooms, according to district policy.
Hildebrand (2001)	Students from 3 schools in a central Nebraska district	The school district applies a common curriculum. School principals and the superintendent made K assignments to each participating school based on their perception of children's needs. Five veteran K teachers were involved in the study. The HDK teacher taught in both morning and afternoon. Two teachers taught FDK in adjoining classrooms.
Larson (2003)	Kindergarten students continuously enrolled in the same school from K to 2nd grade in Maryland's Montgomery County Public Schools. Students in the first FDK group of 17 schools (those with the highest levels of educational need, based on proportion of English language learner students and those eligible for free and reduced-price lunch) were compared in 2nd grade to the previous cohorts who experienced HDK.	In 2001, Montgomery County Public Schools launched its Kindergarten Initiative, which includes FDK, reduced class size, revised K curriculum, ongoing professional development through the summer and school year, and increased communication between parents and school staff. All elements were in place in kindergarten in 2001 except FDK, which was phased into 56 schools sequenced from lower to higher SES and educational need. FDK implementation began with 17 schools identified as having the highest levels of need (FDK), 17 more schools in Phase 2 identified as next most in need (HDK1), and 22 schools in Phase 3 (HDK2). The remaining group of 63 schools (HDK3) provided HDK throughout recent years. In this design, 2nd grade students that attended FDK were compared to students in 3 previous consecutive academic years attending the same FDK schools and to students from the same cohort who did not attend FDK.
Le, Kirby, Barney, Setodji, & Gershwin (2006)	ECLS-K: Includes data from students attending public and private schools	Longitudinal study of national Kindergarten cohort of 1998–99
Lee, Burkam, Ready, Honigman, & Meisels (2006)	ECLS-K: Includes student data from 504 public schools that offered FDK and HDK as schoolwide programs, from the first 2 waves of the ECLS-K, in the fall and spring of K. Children selected for this study are ones who remained in the same public school over the year of K, had complete cognitive test scores (fall and spring), and had a non-missing weight value. Schools selected for this study are those schools that offered either FDK or HDK (but not both), included at least 5 children from the initial student sample, and included grades other than preK and K.	Longitudinal study of national Kindergarten cohort of 1998–99
Plucker, Eaton, Rapp, Lim, Nowak, Hansen, et al. (2004)	Students in 2 districts in Indiana: a large, urban district and a rural district	Urban study: The district implemented FDK solely in Title I schools primarily to lessen the achievement gap between students.
Reynolds, Temple, Robertson, & Mann (2001)	Low-income, mostly black children born in 1980 and enrolled in alternative early childhood programs in 25 sites in Chicago, IL	This study follows up, 15 years later, on a nonrandomized, matched group cohort of 1,539 low-income, mostly black children enrolled in alternative early childhood programs, including the Chicago Child-Parent Center.

Study	Population of study	Brief description (when provided)
Saam & Nowak (2005)	Students enrolled in either FDK or HDK in 3 Midwestern school districts. The analysis focuses in a second part on students eligible for free or reduced-priced lunch.	Three K options were available in these districts: FDK, HDK in the morning, and HDK in the afternoon. Comparisons are made across the three K options. All FDK school sites were Title I schools.
Stofflet (1998)	All students who enrolled in K for the first time in the Anchorage (AK) school district in 1987–88, 1988–89, or 1989–90	The 1987–88 school year was the last year in which there was no FDK. In 1989–90, 9 schools offered FDK, all serving relatively low-income, high-mobility, low-achieving students (Title I schools). The 2nd year of FDK in these schools was 1989–90.
Votruba-Drzal, Li-Grining, & Maldonado-Carreño (2008)	ECLS-K: This sample consists of first-time kindergartners who remained in the same type of K program (full or part day) throughout the K year, who were in K for at least 4 days a week, and who had at least one valid observation in reading and in math from K through 5th grade. The sample represents 78 percent of the entire ECLS-K sample.	Longitudinal study of national Kindergarten cohort of 1998–99
Walston & West (2004)	ECLS-K: The sample was drawn from children who attended public K programs for the first time in 1998–99, did not change teachers during the school year, and who were assessed in English. Specifically, children who repeated K, attended a transitional or 2-year K program, and those in a multi-graded or ungraded K class were excluded from the sample.	Longitudinal study of national Kindergarten cohort of 1998–99
Wang & Johnstone (1999)	English-speaking and bilingual K students in Texas' Irving Independent School District who stayed in the same program throughout the 1996–97 school year	In 1995–96, the Irving School District piloted a FDK program. This study discusses the 2nd year evaluation of that program, which was comprised of 4 parts: investigating oral language development, emergent reading skills, early mathematics concepts, and appropriate behaviors in students in FDK as compared to those in HDK. Similar curriculum and instruction was assumed.
Wolgemuth, Cobb, Winokur, Leech, & Ellerby (2006)	Middle- to upper-class K students at one elementary school in a Colorado city (in the Poudre School District). The majority of students was Caucasian (91%) and did not receive monetary assistance with lunch (89%). Students who switched between FDK and HDK programs during the academic year were excluded from the sample.	FDK was offered for 6.5 hours per day, and HDK for 3 hours. Both programs included similar curricula, and all K teachers met weekly to discuss and align the curriculum. When participants entered K, school personnel (K teacher or principal) assessed them for letter-name and sound knowledge, oral reading fluency, and initial math ability. Reading achievement was assessed in K, 1st, and 2nd grade, and reading and math achievement levels were assessed in 2nd grade.
Zvoch, Reynolds, & Parker (2008)	Kindergarten students in economically disadvantaged school contexts in a large southwestern school district. Students who stayed in school for the full K year were included in the study.	All students were exposed to the same curriculum, Harcourt Trophies. The DIBELS was administered 3 times over the K year to K students attending schools on 9-month schedules and 4 times to those attending year-round schools.

Note: ECLS-K = Early Childhood Longitudinal Study – Kindergarten Class; K= kindergarten; FDK = full-day kindergarten program; HDK = half-day kindergarten program.

Table A2. Descriptions of study samples and outcome variables for 21 reports identified for review

Study	Sample size & characteristics for FDK	Sample size & characteristics for comparison treatment group	Achievement variables studied	Corresponding measures / tests used
Alban & Schatz (2003)	1,158 Montgomery County Public Schools students with Grade 2 Comprehensive Test of Basic Skills reading scores who were in K in 2000–01 in a Phase 1 school (FDK)	Total of 937 MCPS students with Grade 2 Comprehensive Test of Basic Skills reading scores who were in K in 2000–01 in a Phase 2 school (HDK). These students were most similar to those entering Phase 1 schools (FDK) in terms of their foundational skills at K entry and their continuous enrollment in one school.	Foundational skills and reading skills	Letter Identification Assessment Tool; Concepts About Print Tool; Hearing and Recording Sounds Assessment Tool; Word Recognition Assessment Tool; Running Record Assessment Tool, TerraNova Comprehensive Test of Basic Skills
Baskett, Bryant, Rhoads, & White (2005)	119 FDK students randomly selected from the year's total enrollment of 276 children in the all-day class	109 students from the HDK class in a convenience sample determined by records availability	Reading and literacy skills, grades	Brigance screening; Informal Reading Inventory; Observational Survey; report cards
Cannon, Jackowitz, & Painter (2006)	4,487 FDK students from a nationally representative database	4,053 HDK students from a nationally representative database	Math and reading achievement	Item Response Theory-adjusted scores from the ECLS-K assessments
Colvin & Salkind (2005)	19 HDK students in the same district	School achievement	Math, vocabulary, and word analysis subscales of the Iowa Tests of Basic Skills	19 HDK students in the same district
DeCicca (2007)	Total of 8,164 nationally representative students overall (including FDK and HDK): 5,559 are white, 1,445 are black, and 1,160 are Hispanic. An estimated 53 percent of the students were in FDK.	Total of 8,164 nationally representative students overall (including FDK and HDK): 5,559 are white, 1,445 are black, and 1,160 are Hispanic. An estimated 47% of the students were in HDK.	Math and reading achievement	Item Response Theory-adjusted scores from the ECLS-K assessments
del Gaudio Weiss, & Offenberg (no date)	Total of 10,622 students; breakdowns by K status were not provided	Total of 10, 622 students; breakdowns by K status were not provided.	Standardized achievement test scores (reading, math, science), grades	Stanford Achievement Tests, 9th Edition; composite normal curve equivalent achievement scores in reading, math, and science; report cards
Denton, West, & Walston (2003)	Not identified	Not identified	Reading skills and achievement, including basic skills, vocabulary and comprehension	ECLS-K assessments (individually administered, adapted, and untimed cognitive assessments of skills)
Gullo (2000)	730 FDK students	244 HDK students	Reading and math achievement	Reading and math subjects of the Iowa Test of Basic Skills

Study	Sample size & characteristics for FDK	Sample size & characteristics for comparison treatment group	Achievement variables studied	Corresponding measures / tests used
Hildebrand (2001)	47 children (29 males and 18 females) in FDK	44 HDK students (24 males and 22 females)	Reading and math achievement, writing skills	Two standardized assessments: Test of Early Reading Ability-2 and the Test of Early Math Ability Concepts of Writing, an informal assessment tool
Larson (2003)	Continuously enrolled 2nd graders (896) who had participated in FDK in 2000–01	Three successive cohorts of 2nd graders (with approximately 850 students in each cohort), enrolled continuously since K in the 17 schools that implemented FDK in 2000–01, and 4 cohorts of students who were continuously enrolled in schools that had HDK in 2000–01. Students in this second group of schools were separated in 3 groups (HDK 1, 2, and 3), sequenced from lower to higher school SES, and there were approximately 700 students in each of the 4 cohorts enrolled in HDK1 schools, approximately 1,000 students in HDK2 schools, and approximately 3,000 students in HDK3 schools.	Reading, language, and math achievement	5 subtests of the TerraNova Comprehensive Test of Basic Skills: Reading, Language, Language Mechanics, Mathematics, Mathematic Computation
Le, Kirby, Barney, Setodji, & Gershwin (2006)	Total of 7,897 nationally representative, first-time K students overall (combined FDK and HDK) in 1998–99 whose English language proficiency was sufficient to take the full battery of cognitive tests, and whose parents had complete interview data. An estimated 53 percent of the students were enrolled in an FDK program.	Total of 7,897 nationally representative, first-time K students overall (combined FDK and HDK) in 1998–99 whose English language proficiency was sufficient to take the full battery of cognitive tests, and whose parents had complete interview data. An estimated 47 percent of the students were enrolled in an HDK program.	Math and reading achievement	ECLS-K assessments (individually administered, adapted, and untimed cognitive assessments of skills)
Lee, Burkam, Ready, Honigman, & Meisels (2006)	8,455 students in 504 public schools overall (including FDK and treatment); 56 percent were in FDK	8,455 students in 504 public schools overall (including FDK and treatment); 44 percent were in HDK	Literacy, math, and general knowledge	ECLS-K assessments (individually administered, adapted, and untimed cognitive assessments of skills)
Plucker, Eaton, Rapp, Lim, Nowak, Hansen, et al. (2004)	Urban study: FDK students in Title I schools; 1,886 students overall (including FDK and HDK); Rural study: 92 FDK students	Urban study: HDK students in the same district or school; 1,886 students overall (including FDK and HDK); Rural study: 70 HDK students	Academic achievement	Indiana Statewide Testing for Educational Progress-Plus

Study	Sample size & characteristics for FDK	Sample size & characteristics for comparison treatment group	Achievement variables studied	Corresponding measures / tests used
Reynolds, Temple, Robertson, & Mann (2001)	Study and control groups did not compare FDK to HDK	Study and control groups did not compare FDK to HDK	Rates of high school completion and school dropout by age 20; grade retention and special education placement by age 18	School records
Saam & Nowak (2005)	1,046 3rd graders who had attended FDK	1,986 3rd graders who had attended HDK, either morning or afternoon	Language arts and math achievement	Indiana Statewide Testing for Educational Progress-Plus
Stofflet (1998)	519 students for the 1988–89 cohort and 592 students for the 1989–90 cohort	8,951 students, in 3 cohorts, who attended HDK schools and 521 students who attended the 9 FDK schools the year before FDK was implemented	Reading, language arts, and math achievement	Iowa Tests of Basic Skills for years prior to 1995–96 and California Achievement Test after 1995–96
Votruba-Drzal, Li-Grining, & Maldonado-Carreño (2008)	7,574 first-time K students who remained in FDK throughout the K year	6,202 first-time K students who remained in HDK throughout the K year	Math and reading achievement	ECLS-K assessments (individually administered, adapted, and untimed cognitive assessments of skills), including recalibrated 5th grade IRT scores
Walston & West (2004)	4,515 FDK students for the reading achievement analysis; 3,722 FDK students for the math achievement analysis	3,547 HDK students for the reading achievement analysis; 3,046 HDK students for the math achievement analysis	Reading, language arts, and math gains	ECLS-K assessments (individually administered, adapted, and untimed cognitive assessments of skills)
Wang & Johnstone (1999)	Math Proficiency: 412 FDK students (regular), 80 FDK students (bilingual); Emergent Literacy: 415 FDK students (regular), 84 FDK students (bilingual); IPT Oral Language: 412 FDK students (regular), 80 FDK students (bilingual)	Math Proficiency: 143 HDK students (regular), 45 HDK students (bilingual); Emergent Literacy: 135 HDK students (regular), 101 HDK students (bilingual); IDEA Proficiency Test Oral Language: 148 HDK students (regular), 30 HDK students (bilingual)	Oral language development, emergent reading skills, and early mathematics concepts	IDEA Proficiency Test Oral Language Assessment; the concepts and reasoning section of the Woodcock-McGrew Werder Mini-Battery of Achievement; and "An Observation Survey of Early Literacy Achievement"
Wolgemuth, Cobb, Winokur, Leech, & Ellerby (2006)	206 FDK students	283 HDK students (morning and afternoon)	Math and reading abilities	Assessments at K entry; One-minute Reading Assessments; and Northwest Evaluation Association standardized tests
Zvoch, Reynolds, & Parker (2008)	228 FDK students	215 HDK students	Literacy skills	Dynamic Indicators of Basic Early Literacy Skills

Note: ECLS-K = Early Childhood Longitudinal Study – Kindergarten Class; K = kindergarten; FDK = full-day kindergarten program; HDK = half-day kindergarten program.

Appendix B

Table B1: Summary of reports on the Early Childhood Longitudinal Study

Table B1. Higher scoring program (full- vs. half-day) in analyses of reading and math scores from the Early Childhood Longitudinal Study

Report	Student sample ^a	Grade tested							
		Kindergarten		First		Third		Fifth	
		Reading	Math	Reading	Math	Reading	Math	Reading	Math
6	All	Full	Full	X	X	X	X		
	All	Full ^b	Full ^b	X	X	X	X		
7	Black	X	Full	X	X				
	Hispanic	Full	Full	X	X				
	White	Full	Full	X-	Full				
11	All	Full ^b	Full ^b						
9	All	Full	Full						
10	All	Full	Full					Half	Half
8								---	Half

Note: Cell entries name the higher-scoring program. An *X* indicates that the difference between programs was tested but was not greater than one would expect by chance alone. A shaded cell indicates the research report did not investigate differences at this grade level.

^a Results are for the full sample only except where the report did not provide full-sample results.

^b This finding compared student growth or gain in test score.