Effects of Full-Day Kindergarten on Academic Achievement and Social Development
Author(s): Harris Cooper, Ashley Batts Allen, Erika A. Patall and Amy L. Dent
Source: *Review of Educational Research*, Vol. 80, No. 1 (March 2010), pp. 34-70
Published by: American Educational Research Association
Stable URL: http://www.jstor.org/stable/40658445
Accessed: 03/03/2014 14:48

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

American Educational Research Association is collaborating with JSTOR to digitize, preserve and extend access to *Review of Educational Research.*

http://www.jstor.org

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Review of Educational Research
March 2010, Vol. 80, No. 1, pp. 34–70
DOI: 10.3102/0034654309359185
© 2010 AERA. http://rer.aera.net

Effects of Full-Day Kindergarten on Academic Achievement and Social Development

Harris Cooper, Ashley Batts Allen, Erika A. Patall, and Amy L. Dent
Duke University

A meta-analysis found that attending full-day (or all-day) kindergarten had a positive association with academic achievement (compared to half-day kindergarten) equal to about one quarter standard deviation at the end of the kindergarten year. But the association disappeared by third grade. Reasons for this fade-out are discussed. Social development measures revealed mixed results. Evidence regarding child independence was inconclusive. Evidence was suggestive of a small positive association between full-day kindergarten and attendance and a more substantial positive association with the child's self-confidence and ability to work and play with others. However, children may not have as positive an attitude toward school in full-day versus half-day kindergarten and may experience more behavior problems. In general, the research on full-day kindergarten would benefit from future studies that allow strong causal inferences and that include more nonacademic outcomes. The authors suggest that full-day kindergarten should be available to all children but not necessarily universally prescribed.

KEYWORDS: achievement, elementary schools, meta-analysis.

> To have an atmosphere in which the children naturally may be comfortable together.
>
> To have the situations in school carried out as nearly as possible as they are in life outside of school.
>
> To establish a relationship between the child and the teacher in which he feels secure and knows the teacher is a guide and friend.
>
> To understand each child as a person and to further his development in light of his own personality that he may live as a cooperative person in a social group.
>
> To have the curriculum grow out of the interests the children bring to school from their home and community living.
>
> To help the child adjust to the routine necessary in a large school building.
>
> Mink (1937, p. 265)

Grace Mink was the training supervisor at the Michigan Normal College when its lab school provided full-day kindergarten (FDK) for 36 children from small

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

towns in the surrounding area. Having kindergarten in the school was not unique, but having FDK in 1937 was becoming increasingly rare. Similar to circumstances today, Mink suggested that the primary impetus for providing FDK was not to meet academic goals. Rather, she wrote that "the transportation problem makes it impossible for the kindergarten children to be taken home until the rest of the school is dismissed. Therefore, they must remain in school from eight-thirty A.M. until three-fifteen P.M." (p. 262). Still, Mink's objectives for her kindergarteners (notwithstanding her gendered language) could be widely adopted today by educators focused on the developmental needs of young children.

For a variety of reasons, a renewed national debate has emerged about whether it is in the best interests of children, their families, and the broader society to provide kindergarten for a full school day. In this article, we examine the research evidence on the impact of FDK compared to half-day kindergarten (HDK) on the academic, social, and personal development of children. We also look at FDK effects on parents and teachers. In addition to examining a wide variety of outcome measures, we attempt to determine through a synthesis of past research whether any effects of FDK are lasting and, if so, for how long. Finally, by also synthesizing research that compares HDK, FDK in which children attend school every school day, and FDK that meets on alternate school days, we hope to shed light on the issue of whether it is the number of hours that children attend kindergarten or how those hours are arranged that is the locus for any effects.

Before turning to the research, however, we first present a brief history of kindergarten and of its prevalence in the United States, including reasons for the renewed interest in FDK. We also provide a list of both the positive and negative effects that the proponents and opponents of FDK have offered in both the research and policy literatures. Then, we summarize some of the characteristics of FDK programs that might influence whether FDK has positive, negative, or no effects. These were used to help us identify possible moderators and mediators of the outcomes of empirical comparisons between FDK and HDK. Finally, we present the conclusions about FDK that others have drawn from examining the research literature.

A Brief History of Kindergarten

Culminating in Germany in the 1840s, the idea of kindergarten was already nearly 100 years old when Mink wrote about her full-day program (for detailed histories of kindergarten, see Beatty, 1995; Shapiro, 1983). The term *kindergarten* was coined by the German educator and philosopher of education Friedrich Froebel (1782–1852). Froebel was one in a group of thinkers about early childhood education, usually described as having started with Johann Pestolozzi (1746–1853) but including Johann Herbart (1776–1841), Maria Montessori (1870–1952), and John Dewey (1859–1952). All these thinkers observed children carefully, developed a variety of educational practices, and had enormous influence on early childhood education in Germany, Switzerland, the United Kingdom, and the United States.

Froebel, as well as the other early childhood reformers of the time, divided early education into three stages—infancy, early childhood, and childhood. He was most interested in the years between ages 4 and 6, encompassing the transition of education from the family to the school. Froebel's kindergarten, or "child garden," was

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

for children of this age. The essential elements of Froebel's kindergarten, according to Shapiro (1983), were,

> First, the kindergarten was to be an institution where the child could congregate with his peers outside the restraints of the family and school. At the same time, the protective gardenlike atmosphere of the kindergarten [Froebel actually envisioned kindergarten taking place in gardens (Brosterman, 1997)] would guard the child against the corrupting influence of society and the dangers of nature. In the child garden, the mental, physical and social faculties of the child could be cultivated, unfolded and ripened. In the end *kindergarten* signified for Froebel more than an institution—it was an approach to early child training. (p. 22)

Froebel viewed children as possessing an innate desire for creative expression that would flower if left to its own devises (in the proper environment). Froebel's romantic vision of childhood and early childhood education was in stark contrast to other contemporaneous conceptions of what young children could or should do and learn. On one hand was the notion of the child as innocent and full of important potential. On the other hand was the notion of the child as fundamentally sinful and needing to be trained into good behavior and thought. For example, Calvinism, a common strain of Protestantism at the time, saw children as born in ignorance and open to being led to Satan if not properly and carefully educated in the precepts of Christianity.

The philosophical roots of the first kindergartens in the United States remained firmly in the Romantic tradition. The first was opened in Wisconsin in 1856 and was operated by Margaret Meyer Schurz, who was trained by Froebel. Thus, these kindergartens not only adhered to Froebel's philosophy but also were conducted using the German language. The first English-speaking kindergarten in the United States was opened in Boston in 1860 but still used Froebel's method. Not surprisingly, the Transcendentalists of New England found a kindred spirit in Froebel. As the movement expanded, the content of kindergarten activities became more homegrown.

In the 1870s kindergartens were established for the children of immigrants and the urban poor. These typically were privately funded by charity organizations. The first kindergarten associated with public schools was established about the same time in St Louis, Missouri. Because of cost considerations, these first public kindergartens often operated on a half-day schedule.

In the late 1800s the kindergarten movement in the United States took a more pragmatic turn. Exemplary of this change was John Dewey's direction of kindergarten at his lab school from 1896 to 1903 (Beatty, 1995). Dewey reacted to some of the interpretations of Froebel's belief that kindergarten children should be left to their own devises. He did so in the larger context of having seen Montessori classrooms and his uniquely American valuing of independence and self-sufficiency. Dewey felt there was a middle ground between free play and formal instruction. He also felt that the play children engaged in should be grounded in real activities performed at home, which directly built on the work of Montessori. So rather than engaging in the abstract play that was a hallmark of Froebel's approach, Dewey's students played at washing clothes, weaving rugs, and cooking their own lunches (Weber, 1984).

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

During World War I and the period of a huge influx of immigrants from Europe to urban centers, many of the private kindergartens in the United States were incorporated into the public school system. These programs took on the added task of teaching the children of immigrants to speak English and introducing them to American culture.

During World War II, fewer kindergarten teachers were available and the half-day model became the most prevalent form of program delivery. And although Froebel's philosophy still was the dominant approach (as Mink's objectives attest), the emphasis on imparting academic skills to kindergarteners continued to grow. This push for the introduction of academics into kindergarten was further impelled in the 1950s during the cold war, when concern about global competition with our ideological adversaries led to a national desire for the acceleration of academic knowledge acquisition throughout the school years.

The 1960s and 1970s were times of significant expansion of state- and community-funded kindergartens (Elicker & Mathur, 1997). These largely remained half-day programs and maintained the focus on play, socialization, and the transition to more formal learning that would still be recognizable to Mink. But, Elicker and Mathur (1997) report, by the 1990s kindergarten had undergone a significant transformation:

> Curriculum goals had become more academic and skill-oriented. A typical kindergarten morning (or afternoon) was packed with tightly-scheduled reading and writing instruction, mathematics lessons, and other structured activities. . . . Play and socialization in many programs had taken a back seat to preparation for an increasingly rigorous first grade curriculum. (p. 460)

Another development during the 1990s was the reemergence of the FDK program.

In sum then, in many ways, conflicting views of the child as an agent in a process of discovery or an empty vessel to be filled by more or less well-intentioned influences still infuses the modern-day debate over kindergarten, what its goals should be, and whether it is best for children to attend kindergarten for an entire school day. Certainly, however, with the passage of time the context and content of the debate has changed dramatically.

A Contemporary Definition of Kindergarten

Because finding a place for gardens or Satan in our definition would be highly restrictive, for purposes of this article the definition of *kindergarten* is less infused with any particular philosophical perspective than would please the Romantics or the Calvinists. To mesh with today's common understanding, we define *kindergarten* as (a) a formal program (b) offered in a school or school-like setting (c) during the year prior to entering first grade and (d) lasting one school year. *Formal program* means that the kindergarten activities are developed, organized, and supervised by at least one adult, typically at least one of whom has been credentialed to provide kindergarten instruction. *School-like setting* means that the kindergarten is attended by groups of children typically in a building that serves older school children, but some kindergarten programs are free standing. Children attending kindergarten are typically between the ages of 4 and 6.[1]

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

FDK in the United States

As Mink's description suggests, early in the history of kindergarten in the United States FDK was not rare and was especially popular in rural areas. But by 1977 only about one in four kindergarteners attended all day. By 2003, this figure had increased to two in three children in FDK (Child Trends Data Bank, 2003). These percentages were even higher among non-Hispanic Black children (80% were enrolled in FDK) and children from low-income families (71% from households with incomes between $15,000 and $30,000).

Reasons for the Increase in FDK

Most experts agree that there are four primary reasons for the recent increase in FDK (see Brewster & Railsback, 2002). First, the growth in single-parent families and two-parent families in which both parents work outside the home has heightened the need for full-day care for preschool children of all ages. Second, many young children in the United States do not speak English at home. It is widely held that these children require language instruction as early as possible so that the language barrier does not impede their academic progress. Third, the movements toward (a) increased accountability for schools and (b) rising state standards have heightened pressure on educators to accelerate students' movement through the academic curriculum. And last, it has become a national priority to close the gap in achievement between children coming to school from disadvantaged and/or ethnic minority homes versus middle-class and/or White homes. One solution offered for each of these four societal concerns is to increase the time students spend in school and extend the availability of academic instruction to younger children.

Positive and Negative Effects of FDK

The notion that a full day of kindergarten will positively affect this array of social concerns has both its proponents and skeptics. Table 1 presents a summary of the positive effects of FDK that have been offered by its advocates. Principal among the benefits are the academic advantages. Often pointing to research (Pennsylvania Partnership for Children, 2003; WestEd, 2005), proponents of FDK claim that it contributes to increased school readiness by better preparing students for first grade, leads to higher grades and standardized test scores, and supports language development. They also suggest that FDK lessens the need for grade retention, remedial education, and special education placements. As indirect effects on achievement, FDK proponents suggest that it helps foster more independent learning and greater creativity.

Some proponents of FDK point as well to nonacademic benefits (Brannon, 2005; Brewster & Railsback, 2002). FDK advocates claim that it can improve children's self-esteem and self-confidence and assist with their socialization and peer relations by providing more opportunities to interact with other children. Socialization includes learning more quickly to be more cooperative and properly interact with other children and adults. Finally, if school lunches are available, the nutrition of some children in FDK may improve (Pennsylvania Partnership for Children, 2003).

The mechanisms through which FDK leads to these positive effects are viewed by proponents as the result of salutary effects on teachers and classroom instruction

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 1

Potential positive effects of full-day kindergarten compared to half-day kindergarten

For students
Better academic skill development
Reading readiness
Language development, especially for non-English-speaking students
Higher standardized test scores
Fewer grade retentions, less remediation
Fewer referrals to special education services
More independent learning
Easier transition to first grade
Better socialization and peer relations
More cooperative behavior
More opportunity to interact with other children and adults
Positive influence on self-esteem, self-confidence
Better nutrition
For instruction and teaching
Better student attendance
More individualized instruction
Easier identification of problem areas
Less hurried instruction
More repetition of material
Less transition time between activities
Fewer total students for each teacher to track
For parents
Lower child care costs
Easier scheduling and transportation
More contact with the teacher
For society
Levels the playing field for disadvantaged children
More learning opportunities for low income children
Decreased cost because of reduced need for retention and remediation

(Brewster & Railsback, 2002; Ohio State Legislative Office of Education Oversight, 1997; Pennsylvania Partnership for Children, 2003; WestEd, 2005). They claim that FDK improves school attendance and provides time for more individualized instruction that can lead to earlier identification and remediation of learning problems. They say that instruction in FDK classrooms is less hurried and allows for more repetition. Less time is spent on transitions between activities; those activities that occur at the start and close of the school day consume a smaller portion of the day. Children's attendance may improve. For teachers, the total number of students each must track is halved (assuming that HDK teachers have both a morning and an afternoon session).

With regard to parents and society at large, as previously noted, free public FDK can lower child care costs and ease scheduling and transportation concerns.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 2
Potential negative effects of full-day kindergarten compared to half-day kindergarten

For students
Causes higher expectations for first graders
First grade pushed down to kindergarten
Pressure to achieve things before developmentally prepared
K should be more learning by doing rather than worksheets and teacher-led instruction
Increased fatigue, irritability, aggression
Lengthened adjustment because of separation anxiety
Poor role models in lunchroom, playground
Loss of confidence, enjoyment of learning
Less time for informal learning
Part of day in home is also important
Fewer opportunities to visit informal education settings (e.g., museums)
For teachers and instruction
Less planning time
Greater fatigue because of handling same students all day
For parents
Child care needs of working parents still may not be met
For society
Diminished parent responsibility
Cost (salaries, space)
Takes resources from more effective interventions (e.g., smaller classes)
Access is still unequal for disadvantaged students

Also, because teachers have fewer students, parents may find the teacher is more accessible. From a broader societal perspective, proponents of FDK point to its potential to level the playing field for disadvantaged children and decrease the costs for retention and remediation in later grades (Brewster & Railsback, 2002; Le, Kirby, Barney, Setodji, & Gershwin, 2006; Weast, 2001; WestEd, 2005).

Those who are skeptical about the benefits of FDK not only wonder whether it will deliver on its promises but also raise concerns about some possible negative effects on children (see Table 2). In *The Hurried Child*, David Elkind (2006) wrote,

> Although a developmentally appropriate kindergarten should be a half-day of hands-on learning experiences in the morning and nap and quiet time in the afternoon, this often does not happen. The kindergarten is now seen as a preparation for the first grade and a place where children learn their letters and numbers. (p. 67)

Thus, of most concern among the skeptics is that FDK may raise expectations about what children should know when they enter first grade. This occurs because first grade content is pushed down into kindergarten (Karweit, 1992). But the child may not be ready for the added rigor. Cruikshank (1986) argued that kindergarteners learn by doing: "A 5-year old is still a 5-year old. They need a developmental program that meets their cognitive (academic), social, emotional and motor needs"

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

(p. 12). By bending to societal pressures, skeptical educators argue, kindergarteners may be asked to learn things they are not ready for developmentally, and the effort may have the opposite effects from those intended (Gullo, 1990).

Concerns about FDK go beyond a possible failed attempt to accelerate learning. Some kindergarten teachers worry that the longer school day will lead children to be tired, irritable, and aggressive and will increase the time children need for adjustment to school by prolonging separation anxiety from parents (Good, 1996). Also, they wonder whether exposure to poor role models (e.g., older children misbehaving in the lunchroom) will increase bad behavior on the part of children in FDK. All of this may actually result in a loss of confidence and enjoyment in learning (WestEd, 2005) rather than the boost in self-esteem envisioned by FDK proponents. And, some argue, there are opportunities for informal learning that arise at home and in the community that are lost when young children spend their day in school (e.g., visits to museums with parents; Berson, 1968; Brannon, 2005).

Also in contrast to the arguments of proponents, kindergarten teachers worry that instruction might be adversely affected by FDK because teachers will have less time for planning and be more fatigued because of handling the same students all day (Good, 1996). Working parents may find that even though the school day has been lengthened, it still does not meet their needs for child care, as the beginning or end of the school day does not coincide with the beginning or end of the work day (Olsen & Zigler, 1989). And finally, from a societal viewpoint, FDK may lead to an erosion of parents' sense of responsibility for their children (Gullo, 1990), will increase costs that take resources away from other perhaps more efficacious programs (Karweit, 1992), and still may reveal unequal access that works against the more disadvantaged students (WestEd, 2005).

Factors That Might Influence the Effects of FDK

Program components and instructional techniques. One thing that both the proponents and the skeptics of FDK agree on is that, regardless of its length, the content and instructional strategies used in the kindergarten program are paramount to its success or failure. Thus, the list given earlier of potential effects of FDK on instructional and teacher practices might also be viewed as mediators of FDK effects on students; their presence or absence might determine whether FDK has positive, negative, or no effects on student outcomes. For example, FDK might lead some teachers to spend more time working with students individually whereas other FDK teachers do not. We might then expect the effects of FDK to be more positive for students in need of remediation in the former classrooms than the latter ones.

Miller (2002), citing a document from the ERIC Clearinghouse on Early Childhood Education, listed 11 characteristics of effective kindergarten programs, regardless of their length. Effective programs (a) integrate new learning with past experience in projects; (b) use mixed-ability, mixed-age groups of students; (c) provide an unhurried setting; (d) involve children in firsthand experience; (e) provide for informal interaction with objects, other children, and adults; (f) emphasize language development and preliteracy skills; (g) share information and build understanding with parents; (h) emphasize reading to children at school and at home; (i) balance small-group, large-group, and individual instruction; (j) assess

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

progress through close observation, and (k) develop social skills, including conflict resolution.

In keeping with most of these dicta, Clark (2001) cautioned that kindergarten providers should "resist the pressure to include more didactic academic instruction in all-day kindergarten programs. . . . An all-day kindergarten program can provide children with the opportunity to spend more time engaged in active, child-initiated, small-group activities" (p. 4). Thus, assuming these characteristics are indeed the keys to effective kindergarten programs, the extent to which FDK encourages, discourages, or leaves unaffected the use of these strategies should mediate the appearance of positive, negative, or no effects on children.

Individual differences. It is also recognized by those who study kindergarten that the potential exists for either positive or negative effects of FDK to occur depending on the readiness of the child for a longer day in a school-like setting. However, the issue of individual differences is complicated by the fact that individual differences that might be used to predict children's reactions to FDK can also serve as the outcomes that FDK is meant to inculcate in children. For example, the list of potential positive effects of FDK includes the value of socialization to other children, whereas the list of negative effects includes concerns about separation anxiety. But a lack of separation anxiety and greater prior socialization to other children might also be individual differences that moderate whether or not a child will benefit from FDK in other areas of development.

The issue of individual differences in children's reactions to FDK must consider how characteristics and abilities of children interact to determine who may benefit from FDK. For example, by far the most frequent individual differences mentioned as moderators of FDK effects are the child's economic circumstances and whether or not English is spoken at home. However, children from a disadvantaged home where English is not spoken might gain the most from FDK but only if they possess the needed social skills and experience little separation anxiety.

A Summary of Past Synthesis Findings

Table 3 presents brief summaries of the conclusions reached by other scholars about the cumulative research on the effects of FDK. These syntheses suggest several areas of consensus and confusion. First, although the literature on FDK suggests a rich array of potential impacts on children, research has focused on only a small subset of these. Academic effects predominate. Although other measures of effects appear sporadically, they rarely are mentioned repeatedly by those who have examined the literature.

Second, the research methodologies used in FDK studies leave much to be desired. Reviewers bemoan the lack of studies that randomly assigned children to FDK and HDK programs. They also note that small sample sizes reduce the power of studies to detect small effects and that a lack of diverse student populations in most studies reduces researchers' ability to examine the generality (and specificity) of effects.

That said, several past interpreters of FDK research do see in the literature support for immediate positive academic effects of FDK for children in general and

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 3

Summary of conclusions of other reviews of the literature on the effects of full-day kindergarten

Puleo (1988)

Reviewed 19 studies with 30 outcomes related to basic academic skills. Found 22 outcomes favored FDK.

"Much of the research suffers from serious problems in internal and external validity" (p. 428). More recent studies (since 1980) favor FDK. Strong FDK effects for low SES children. "Research on long-term effects supports the value of full-day" (p. 431). "Most investigators report that fatigue [among children] is not a problem" (p. 432).

Olsen and Zigler (1989)

"In general, extended-day programs do seem to bring about short-term increases in standardized test scores, particularly with children who are disadvantaged, bilingual, or 'least-ready' for school. There is little evidence to suggest, however, that positive changes in academic performance is maintained beyond the early elementary school years or that there are changes in motivation or general intellectual ability that would be likely to support significant long-term change" (p. 179).

Karweit (1992)

Mentioned 4 experiments and 12 quasi experiments. "The academic benefits of full-day kindergarten, then, are inconsistent and most clearly demonstrated in the case of at-risk populations. . . . Full-day kindergarten will have to compete with alternatives such as establishing prekindergartens, reducing class size in early grades, or providing tutors in early grades. Unfortunately, little research documents the relative effectiveness of these different approaches" (p. 83).

Clark and Kirk (2000)

"Most of the recent research on all-day kindergarten indicates positive benefits for children in terms of academic achievement and behavior" but "what children *do* in kindergarten may be more important than how long they are in the classroom each day" (p. 231).

Elicker (2000)

FDK students progress further academically compared to HDK or alternate-day students. Tentative evidence that FDK has stronger, longer-lasting academic benefits for children from low-income families. No current, strong evidence that academic achievement gains for FDK persist beyond first grade. No evidence of detrimental effects of FDK.

Jones (2002)

Meta-analysis of 22 studies that compared academic achievement of FDK and HDK students. Overall effect size was $d = 0.56$. FDK had a smaller effect on math than on reading or language.

Education Commission of the States (2004)

"Experts now are in general agreement that there are no detrimental effects to attending full-day kindergarten and, in fact, students in full-day programs show significantly stronger academic gains over the course of the kindergarten than their half-day counterparts. The research also finds that poor and minority students especially can benefit from participation in full-day programs. There is less agreement about the degree to which benefits gained from attending full-day kindergarten carry forward throughout the student's academic career" (p. 1).

(continued)

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 3 (continued)

Plucker (2005)
FDK "is associated with a wide range of positive outcomes, including increased student achievement and social and behavioral development." "Positive effects appear larger for disadvantaged students." FDK is expensive. Many comments regarding quantity versus quality but this is an oversimplification: FDK "fundamentally changes the nature of activities that occur in that program" (p. 6).

WestEd (2005)
"Children benefit from a developmentally appropriate, full-day program, most notably in terms of early academic achievement" (p. 1). Benefits include increased school readiness, higher academic achievement, improved student attendance, literacy and language development, social and emotional, decreased retention and remediation.

Le, Kirby, Barney, Setodji, and Gershwin (2006)
Definitive conclusions "have been stymied by studies with small sample sizes, lack of statistical control, potential self selection bias, and other methodological weaknesses. . . . There is no consistent definition of full-day kindergarten, and little research that explicates the nature, features, or structure of particular programs. As a result, generalizations about full-day versus half-day programs must be made with caution" (p. 6).

Note. FDK = full-day kindergarten; HDK = Half-day kindergarten.

for children from disadvantaged homes in particular. Lesser consensus exists about whether positive effects of FDK persist beyond the end of kindergarten or first grade. Here, some reviewers find more support than others.

The Present Synthesis

With this as background, we undertook the current synthesis with hopes of improving in several ways the current state of knowledge about the effects of FDK. In addition to conducting the most comprehensive meta-analysis of the effects of FDK on achievement, we were the first to use statistical methods to aggregate nonacademic effects of FDK. Also, we attempted to meta-analytically test the general consensus of past interpreters of the research that FDK has more positive effects for children from disadvantaged homes. Then, we looked at whether enough evidence had accumulated to draw conclusions about the moderating influence of other individual differences or program components. Next, we attempted to bring together all of the research on the long-term effects of FDK. This has been an area of contention among interpreters of the FDK research literature; the longitudinal evidence deserves a more thorough and systematic treatment than it has received in the past. Finally, we are the first to systematically integrate research that compared HDK and FDK programs that meet every school day to FDK programs that meet on alternate days. Alternate-day FDK (AD-FDK) is often adopted as a mechanism to save transportation costs without increasing instructional costs. Also, by gauging effects of the two FDK variations, we may be able to parse out

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

which, if any, FDK effects are from the longer school day or the increase in instructional time.

Method of Literature Search and Judging Study Relevance

Literature Search Procedures

First, we searched six different electronic reference databases for reports related to FDK: ERIC, PsycINFO, Sociological Abstracts, Dissertation Abstracts, EconLit, and Google Scholar. The searches were conducted during October 2009 and covered all years available in the database. The terms *all day kindergarten* or *full day half day schedules* were used in these searches. Two researchers then examined each title and abstract in the document file and judged whether they felt the document (a) was irrelevant (e.g., the document mentioned FDK only in passing but was primarily focused on another topic), (b) likely contained background information on FDK (e.g., an opinion piece or description of a program) but not empirical evidence on its effects, or (c) likely contained empirical evidence on FDK associations. If either researcher felt that the document might contain data relevant to FDK, we obtained the full document. In total, 655 document records were examined. Of these, 290 were deemed potentially relevant by at least one document record reader. We then obtained the 290 potentially relevant documents along with numerous documents we used "background" and examined these in their entirety.

Next, we employed two direct-contact strategies to ensure that we tapped sources that might have access to FDK research that would not be included in the reference databases. First, we contacted through e-mail 153 deans, associate deans, or chairs of colleges, schools, or departments of education at institutions of higher education and requested that they ask their faculty to share with us any research they had conducted that related to FDK. Second, we sent a similar e-mail request to the National Association of Test Directors, the Education Commission of the States, and the Regional Educational Laboratories.

Finally, we examined the references in previous syntheses of the FDK literature to determine whether these contained mention of any reports we had not encountered through the reference database and direct-contact searches.

Criteria for Including Studies

For a study to be included in the research synthesis, several criteria had to be met. Most obviously, the study had to have focused on the difference between kindergarten programs that operated on a half-day schedule versus a full-day schedule on a measure of student academic achievement or readiness, some other measure of student development or well-being, or some measure of classroom process.

We employed five additional screens to eliminate studies before coding began. First, we set aside studies that compared HDK to FDK that met on alternating school days. These studies were examined separately. Second, the only sampling restriction placed on studies was that they had to study kindergarten programs based in the United States or Canada ($k = 2$).[2] Third, we eliminated studies that intentionally confounded the FDK variable with another instructional intervention.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

For example, we eliminated a study if the report stated that in addition to going to school all day students in FDK classes were also provided services not available to HDK students (e.g., special reading or literacy programs or medical services for FDK students only) or were instructed using a different curriculum.[3] Fourth, we included only outcomes measured at the end of the kindergarten year or the beginning of first grade. Any data on the lasting effects of FDK were set aside for separate analysis. Finally, the report had to contain enough information to permit the calculation of an estimate of the relationship between the length of the kindergarten day and the outcome measure.

Studies Comparing FDK to HDK at the Conclusion of the Kindergarten Year

Method

Information retrieved from studies. Numerous different characteristics of each study were included in the database. These characteristics encompassed six broad distinctions among studies: (a) the research report, (b) the research design, (c) the kindergarten programs themselves, (d) the setting of the study and the sample of students, (e) the outcome measures, and (f) the estimate of the relationship between attending FDK versus HDK on the outcome variable. As is true in all meta-analyses, many of the study characteristics we coded either were not reported often enough or exhibited too little variability across studies to be examined as moderators of FDK effects. Full descriptions of all the coded characteristics are available on-line at http://rer.sagepub.com/supplemental/. All appendices referred to below are available on this website.

Effect size estimation. We used the standardized mean difference, or *d*-index, to estimate the effect of the length of the kindergarten day on student outcomes (Cohen, 1988). In the meta-analysis, we subtracted the HDK mean from the FDK mean and divided the difference by their weighted average standard deviation. Thus, positive *d*-indexes indicate that the students in FDK programs had better achievement or academic readiness or higher scores on other measures of development or well-being or that the classroom process occurred more frequently in FDK classes. If available, we calculated effect sizes based on the means and standard deviations of the student outcomes. If means and standard deviations were not available, we indirectly retrieved the information needed to calculate *d*-indexes from inferential statistics (see Borenstein, 2009). Effect sizes that adjusted or control for other variables also were retrieved or calculated if the needed information was available.

Coder reliability. Each research report was coded by two coders. If there was a discrepancy in coding, this was first discussed by the coders. If the disagreement could not be resolved, the first author (H.C.) was consulted. Because all studies were independently coded twice and continuing disagreements were resolved by a third independent coder, we did not calculate reliability for this process (which would have entailed training three more coders and having them code at least a subset of studies). Out of all possible codes, the two coders had an initial agreement rate of about 89%.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Methods of data integration. First, we examined the distribution of effect sizes to determine if any were statistical outliers. The Grubbs (1950) test, also called "the maximum normed residual test," was applied (also see Barnett & Lewis, 1994). This test identifies outliers in univariate distributions and does so one observation at a time. If outliers were identified (using $p < .05$, two-tailed, as the significance level), these values were set at the value of their next nearest neighbor.

Publication bias. Even though we used several search techniques, there is still the possibility that we did not obtain all studies that have investigated the effects of FDK. Therefore, we used Duval and Tweedie's (2000a, 2000b) trim-and-fill procedure to test whether the distribution of effect sizes used in the analyses was consistent with variation in effect sizes that would be predicted if the estimates were normally distributed. If the distribution of observed effect sizes was skewed, indicating a possible bias created either by the study retrieval procedures or by data censoring on the part of authors, the trim-and-fill method provides a way to estimate the values from missing studies that need to be present to approximate a normal distribution. Then, it imputes these missing values, permitting an estimate of the impact of data censoring on the observed distribution of effect sizes.

Calculating average effect sizes. We used a weighted procedure to calculate average effect sizes across all comparisons (Borenstein, Hedges, Higgins, & Rothstein, 2005). Also, 95% confidence intervals (CIs) were calculated for average effects. If the CI did not contain zero, then the null hypothesis of no FDK versus HDK difference can be rejected.

Identifying independent hypothesis tests. One problem that arises in calculating effect sizes involves deciding what constitutes an independent estimate of effect. Here, we used a shifting unit of analysis approach (Cooper, 2010). In this procedure, each effect size associated with a study is first coded as if it were an independent estimate of the relationship. For example, if a single sample of students permitted comparisons of FDK effects on math and reading readiness scores, two separate effect sizes were calculated. However, for estimating the overall effect of FDK, these two effect sizes were averaged prior to entry into the analysis, so that the sample contributed only one effect size. To calculate the overall weighted mean and CI, this one effect size would be weighted by the inverse of its variance. However, in an analysis that separately examined the effect of FDK on math and reading readiness scores, this sample would contribute one effect size to each estimate of a category mean effect size.

The shifting unit of analysis approach retains as much data as possible from each study while holding to a minimum any violations of the assumption that data points are independent. Also, because effect sizes are weighted by sample size in the calculation of averages, a study with many independent samples containing just a few participants will not have a larger impact on average effect size values than a study with only a single, or a few, large independent samples.

Tests for moderators of effects. Possible moderators of difference between FDK and HDK were tested using homogeneity analyses (Cooper, Hedges, & Valentine, 2009).

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Fixed and random effect. When an effect size is said to be "fixed," it assumes error is solely from differences among participants sampled in the study. However, it is also possible to view studies as containing other random influences, including differences in teachers, facilities, community economics, and so on. This view assumes kindergarten classrooms, schools, or even school districts in our meta-analysis also constitute a random sample drawn from a (vaguely defined) population of conditions under which kindergarten programs take place. If it is believed that such random variation in programs is a significant component of error, a random effect model should be used that takes into account this study-level variance in effect sizes (for a discussion of fixed and random effects, see Hedges & Vevea, 1998).

Rather than opt for a single model describing the underlying variation in effects, we chose to apply both models to our data. We conducted all our analyses twice, once employing fixed-effect assumptions and once random-effects assumptions. By employing this sensitivity analysis (Greenhouse & Iyengar, 2009), we could examine the effects of different assumptions on the outcomes of the synthesis. Differences in results based on different assumptions could then be part of our interpretation of results. For example, if an analysis reveals that a moderator variable is significant under fixed-effect assumptions but not under random-effects assumptions, this result suggests a limit on the generalizability of inferences about the moderator variable.

Software. All statistical analyses were conducted using the Comprehensive Meta-Analysis statistical software package (Borenstein et al., 2005). Only outcomes measures that were collected in studies described in at least four separate reports with at least four independent samples had meta-analytic procedures applied to them.

Results[4]

The literature search located a total of 40 reports with usable data that compared students who attended FDK to students who attended HDK on some measure of immediate achievement. The 40 reports provided 254 separate effect sizes (152 unadjusted and 102 adjusted) based on 55 separate samples. The 40 reports appeared between 1979 and 2009. The sample sizes ranged from 7 to 12,790. Seven of these reports were published.

Of the 40 reports, 30 contained 43 separate samples with information that allowed FDK students to be compared to HDK students without statistical adjustments to the comparison. Of these, 16 reports and 23 samples also provided a comparison of FDK to HDK that was statistically adjusted post hoc to make the groups more comparable. There was an additional 10 reports containing 12 samples for which only an adjusted comparison could be obtained. Thus, a total of 26 reports containing 35 samples adjusted for various factors to make the groups more comparable.

Six reports that assessed achievement and two additional reports (including nine independent samples total) provided information on student attendance. The eight reports provided 12 separate effect sizes. The reports appeared between 1980 and 2007. The sample sizes ranged from 18 to 13,742. All of these reports were unpublished. Eight of these samples provided only an unadjusted comparison of

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

FDK to HDK student attendance, and one provided both an unadjusted comparison as well as a comparison statistically adjusted for various factors to make the groups more equivalent.

Five reports that assessed achievement and one additional report (six samples total) provided information on parent or teacher reports of improvement in students' self-confidence and improvement in their ability to work and play with others. In all cases, these outcomes were assessed with single items from a questionnaire filled out by parents or teachers. The six samples provided eight separate effect sizes for self-confidence and seven for student ability to work and play with others. The six reports appeared between 1980 and 2005. The sample sizes ranged from 29 to 250. All of these reports were unpublished. Five of these samples provided only an unadjusted comparison of FDK to HDK on confidence and ability to work and play with others. One report provided an unadjusted effect as well as a comparison statistically adjusted for various factors to make the groups more equivalent.

Three reports that assessed achievement and one additional report (four samples total) provided information on improvement in student independence. The four reports provided four separate effect sizes. The four reports appeared between 1983 and 2005. The sample sizes ranged from 29 to 123. All of these reports were unpublished. All studies measured this outcome with a single item from a questionnaire filled out by either parents or teachers. All samples provided only an unadjusted comparison of FDK to HDK child independence.

Finally, five reports that assessed achievement and two additional reports (seven samples total) provided information on parent preference for FDK. In all cases, this outcome was assessed with a single questionnaire item asking parents to indicate their preferred choice for a kindergarten program. The seven reports provided eight separate effect sizes and appeared between 1980 and 2005. The sample sizes ranged from 83 to 250. Six of the reports were unpublished and one was published. Six of the samples provided only an unadjusted comparison of FDK to HDK preference, and one report provided an unadjusted effect as well as a comparison statistically adjusted for various factors to make the groups more equivalent.

Overall Association of FDK With Academic Achievement[5]

Table 4 presents the results of the analyses examining the overall effect size for FDK using achievement and other indicators of adjustment and well-being as the outcome variable, after the test for statistical outliers and Winsorization (see Appendix D for details of the outlier tests and adjustments). The unadjusted effects sizes for FDK versus HDK on achievement ranged between $d = -0.98$ and $+1.83$. Of the 152 unadjusted effect sizes, 128 were in a positive direction, 23 were in a negative direction, and 1 was exactly zero. Using fixed-effect assumptions, the weighted mean d-index was 0.24 and was significantly different from 0 (95% CI = 0.23/0.25). Using a random-effects model the weighted average d-index was 0.31 (95% CI = 0.25/0.38).

The adjusted effect sizes on achievement varied between $d = -0.98$ and $+1.82$. Of the 102 adjusted effect sizes, 89 were in a positive direction and 13 in a negative

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 4
Overall effects of full-day kindergarten on immediate outcomes

	Unadjusted *d*-index						Adjusted *d*-index					
		Fixed		Random				Fixed		Random		
Outcome	*k*	*d*	−/+CI	*d*	−/+CI	*Q*	*k*	*d*	−/+CI	*d*	−/+CI	*Q*
Achievement	43	0.24	0.23/0.25	0.31	0.25/0.38	591.26***	35	0.33	0.30/0.36	0.35	0.23/0.46	362.55***
Attendance	10	0.10	0.08/0.13	0.10	0.06/0.15	9.26	1	0.09	−0.32/0.50	0.09	−0.32/0.50	0.00
Confidence	6	0.56	0.41/0.72	0.55	0.32/0.79	10.56†	1	1.36	0.91/1.82	1.36	0.91/1.82	0.00
Ability to work or play with others	6	0.55	0.39/0.71	0.67	0.29/1.05	24.80***	1	1.06	0.63/1.49	1.06	0.63/1.49	0.00
Independence	4	0.38	0.15/0.61	0.28	−0.25/0.80	14.37**						
Preference for FDK	7	1.75	1.61/1.90	1.77	1.46/2.09	25.79***	1	−0.11	−0.52/0.30	−0.11	−0.52/0.30	0.00

Note. CI = confidence interval; FDK = full-day kindergarten; HDK = half-day kindergarten.
†$p < .10$. **$p < .01$. ***$p < .001$.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

direction. Using fixed-effect assumptions, the weighted mean *d*-index was 0.33 and was significantly different from 0 (95% CI = 0.30/0.36). Using a random-effects model, the weighted average *d*-index was 0.35 (95% CI = 0.23/0.46).

The trim-and-fill analyses were conducted in several different ways (details can be found in Appendix E). Four different analyses involving the unadjusted effect sizes produced new distributions with imputed values that ranged from $d = 0.22$ to 0.27, and all remained statistically significant. For the adjusted effect sizes, no additional effect sizes were imputed. Thus, even when testing for possible data censoring, the effect of FDK on achievement was positive and significantly different from zero under all models.

Overall Association of FDK With Nonacademic Outcomes

As noted above, five nonacademic outcomes were studied often enough to warrant quantitative synthesis of results.

Attendance. The unadjusted effect sizes using attendance as the outcome variable varied between $d = -0.14$ and +0.27. Using fixed-effect assumptions, the weighted mean *d*-index for unadjusted effect sizes was $d = 0.10$ (95% CI = 0.08/0.13). Using a random-effects model, the weighted average *d*-index was also 0.10 (95% CI = 0.06/0.15). One study tested the adjusted effect size on attendance. It revealed a nonsignificant positive effect of FDK ($d = 0.09$, 95% CI = –0.32/0.50).

Self-confidence. The unadjusted effect sizes using self-confidence as the outcome variable varied between $d = +0.15$ and +0.99. Under fixed-effect assumptions, the weighted mean *d*-index was 0.56 (95% CI = 0.41/0.72). Under random-effects assumptions, the weighted mean *d*-index was 0.55 (95% CI = 0.32/0.79). One study tested the adjusted effect size on self-confidence. It revealed a significant positive effect size favoring FDK ($d = 1.36$, 95% CI = 0.91/1.82).

Ability to work or play with others. The unadjusted effect sizes for ability to work and play with others varied between $d = +0.18$ and +1.06. Using fixed-effect assumptions, the weighted mean *d*-index for unadjusted effects was 0.55 (95% CI = 0.39/0.71). Using a random-effects model, the weighted average *d*-index was 0.67 (95% CI = 0.29/1.05). One study tested the adjusted effect size on student ability to work and play with others. It revealed a significant positive effect size for FDK ($d = 1.06$, 95% CI = 0.63/1.49).

Child independence. The unadjusted effect sizes for child independence varied between $d = -0.36$ and +0.79. Using fixed-effect assumptions, the average *d*-index was 0.38 (95% CI = 0.15/0.61). Using a random-effects model, the average *d*-index was not significantly different from 0 ($d = 0.28$, 95% CI = –0.25/0.80). There were no tests of child independence that used adjusted effect sizes.

Parent preference for FDK. The unadjusted effect sizes for parent preference for FDK varied between $d = +1.09$ and +2.38. Using fixed-effect assumptions, the weighted mean *d*-index was 1.75 (95% CI = 1.61/1.90). Using a random-effects

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

model, the weighted average *d*-index was 1.77 (95% CI = 1.46/2.09). One study tested the adjusted effect size for preference for FDK. It revealed a nonsignificant negative effect size for FDK ($d = -0.11$, 95% CI = –0.52/0.30).

In sum then, we found sufficient data to conduct meta-analyses on five nonacademic outcomes that proponents claimed would be positively affected by FDK: attendance, self-confidence, ability to work and play with others, child independence, and parents' preference for FDK. For proponents of FDK, the results were mixed. There was some evidence of a positive association between FDK and the child's self-confidence and ability to work and play with others, but evidence of a positive association between FDK and child independence and attendance was more tentative.

Methodological Moderators of the Association of FDK With Achievement

Next, we conducted analyses exploring five moderators of the FDK–achievement association related to the methodology of the studies: (a) the type of report, (b) the organization in which the study was conducted, (c) the research design, that is, whether or not the study used some procedure to equate students in the FDK and HDK groups, (d) if students were equated, whether this was done on prior achievement only or on prior achievement plus other student differences, and (e) the total study sample size. Appendix F presents the details of these analyses. Below, we briefly summarize the results.

Type of report. The unadjusted effect size from published studies was not significantly different from unpublished studies using either fixed-effect or random-effects assumptions. Among the adjusted effect sizes, the effect sizes varied significantly for published ($d = 0.45$, 95% CI = 0.38/0.52) versus unpublished studies ($d = 0.30$, 95% CI = 0.27/0.33) under fixed-effect assumptions but not random-effects assumptions.

Organization. The average unadjusted effect size significantly varied by organization under a fixed-effect model but not under a random-effects model. Under fixed-effect assumptions, the unadjusted effect size was greater among studies conducted by researchers affiliated with a university ($d = 0.28$, 95% CI = 0.26/0.31) than a government agency ($d = 0.23$, 95% CI = 0.21/0.24). Among the adjusted effect sizes, the average effect size did not vary significantly by organization under either a fixed-effect model or a random-effects model.

Research design. Samples were considered to have employed a matching procedure if they matched individual students in FDK to individual students in HDK. Samples in which matching was conducted at the school level were not considered to have employed matching for the purposes of this analysis. Using both fixed-effect and random-effects models, the unadjusted effect size for matched designs (fixed-effect and random-effects: $d = 0.46$, 95% CI = 0.37/0.55) was significantly greater than that of nonmatched designs (fixed effect: $d = 0.24$, 95% CI = 0.22/0.25; random effects: $d = 0.30$, 95% CI = 0.23/0.37). In all reports for which a statistically adjusted effect size could be

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

obtained, a nonequivalent control group design with post hoc equating was employed. The adjusted effect size was significantly larger for matched designs compared to nonmatched designs under fixed-effect (matched: $d = 0.40$, 95% CI = 0.32/0.47; nonmatched: $d = 0.31$, 95% CI = 0.28/0.34), but not random effects models.

One study (Elicker & Mathur, 1997) not included in the above analyses employed random assignment, making it a true experiment rather than a nonequivalent control group design. In this study, enrollment in four FDK classes was determined by randomly drawing from a pool of all incoming kindergarten students. The remaining students were assigned to eight half-day classes. At the end of the kindergarten year all students were evaluated by their teachers on their readiness to continue to first grade as well as a number of other academic outcomes measured at later time points. The researchers found that FDK students were rated by teachers to be more ready to continue to first grade compared to HDK students, controlling for a prekindergarten screening test and family income.

Equating variables. The association between the magnitude of effect sizes and the variables used to equate groups was examined. Under a fixed-effect model, the adjusted effect size was greater for samples adjusted on an achievement premeasure plus additional characteristics ($d = 0.38$, 95% CI = 0.32/0.44) compared to samples adjusted on preachievement alone ($d = 0.31$, 95% CI = 0.27/0.34). The effect size did not significantly vary by the equating variables under a random-effects model.

Total sample size. The unadjusted association of FDK with achievement varied for different total sample sizes under fixed-effect assumptions but not under random-effects assumptions. Contrasts revealed that the unadjusted effect size from samples with fewer than 200 participants ($d = 0.35$, 95% CI = 0.29/0.41) was significantly greater than the unadjusted effect size from samples with 200 or greater participants ($d = 0.24$, 95% CI = 0.22/0.25). The adjusted effect size significantly varied by total sample size under both fixed-effect assumptions and random-effects assumptions. Under a fixed-effect model, the adjusted effect size from samples with fewer than 100 participants ($d = 0.63$, 95% CI = 0.55/0.70) was significantly greater than the adjusted effect size of samples with between 100 and 200 participants ($d = 0.43$, 95% CI = 0.36/0.49). There was no difference between these groups under random-effects assumptions. The adjusted effect size for samples with 200 participants or more (fixed effect: $d = 0.24$, 95% CI = 0.20/0.27; random effects: $d = 0.18$, 95% CI = 0.06/0.30) was significantly lower than the adjusted effect size for samples with fewer than 200 participants under both fixed-effect and random-effects models.

Taken together, the tests of methodological moderators revealed results not uncommon in meta-analyses. Published studies, studies conducted at universities, and studies with smaller samples tended to reveal larger effects. With regard to research design, studies using matching revealed larger effects than studies not using matching, and matching on more variables revealed larger effects than matching on only one variable.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Treatment and Student Moderators of the Association of FDK With Achievement

We conducted moderator analyses of the effect of FDK on academic achievement using four moderators related to characteristics of the treatment and of the students. Table 5 presents these results.

Year of kindergarten program. We used the year of the report's appearance to serve as a proxy for the year in which the FDK versus HDK comparison took place. Using a fixed-effect model, the unadjusted effect sizes indicated that the advantage for FDK was smaller in studies conducted before 1990 (d = 0.19, 95% CI = 0.18/0.21) than in 1990 or after (d = 0.34, 95% CI = 0.32/0.36). The relationship was in the same direction using a random-effects model, but the difference was not significant (pre-1990: d = 0.28, 95% CI = 0.18/0.38; 1990 or after: d = 0.36, 95% CI = 0.28/0.43). For adjusted effect sizes, both the fixed-effect and random-effects models revealed stronger associations after 1990 (fixed effect pre-1990: d = 0.26, 95% CI = 0.20/0.31; 1990 or after: d = 0.36, 95% CI = 0.32/0.39; random effects pre-1990: d = 0.20, 95% CI = 0.05/0.35; 1990 or after: d = 0.57, 95% CI = 0.38/0.77).

FDK minutes per day. Under a fixed-effect model, the unadjusted effect size when full-day students spent less than 360 minutes in school (d = 0.30, 95% CI = 0.27/0.33) was significantly lower than the effect size when full-day students spent 360 or more minutes at school per day (d = 0.37, 95% CI = 0.32/0.41). Under the random-effects model, the unadjusted effect size when full-day students spent less than 360 minutes in school (d = 0.20, 95% CI = 0.09/0.31) was also significantly lower than the effect size when students spent 360 or more minutes at school per day (d = 0.49, 95% CI = 0.28/0.70).

The adjusted effect size when full-day students spent less than 360 minutes in school (fixed effect: d = 0.20, 95% CI = 0.13/0.26; random effects: d = 0.07, 95% CI = –0.11/0.25) was again significantly lower than when full-day students spent 360 or more minutes at school per day (fixed effect: d = 0.44, 95% CI = 0.37/0.51; random effects: d = 0.43, 95% CI = 0.18/0.67) under both fixed-effect and random-effects assumptions.

FDK selection. Samples were grouped according to whether parents nominated students to be in FDK or if some characteristic of the students, such as low prior achievement or at-risk status, allowed them to be selected for FDK. In some studies, several procedures were employed sequentially. For example, first parents may have nominated their child to attend FDK, and then to further limit the number of students attending FDK a characteristic of the student was used to select from among parent-nominated children. In these cases, samples were classified by the primary mode of selection. So in the former example, this sample would have been classified in the parent-nomination category.

The unadjusted effect sizes did not significantly vary by FDK selection procedure under either fixed-effect or random-effects models. Adjusted effects sizes significantly varied by selection procedure under fixed-effect assumptions but not under random-effects assumptions. Under fixed effect, the adjusted effect size for samples in which FDK students were selected by parent nomination (d = 0.55, 95%

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 5

Treatment and student moderators of the effect of full-day kindergarten on immediate achievement

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	–/+CI	*Qb*	*d*	–/+CI	*Qb*	*k*	*d*	–/+CI	*Qb*	*d*	–/+CI	*Qb*
Report year	43			145.46***			1.38	35			9.92**			9.01**
Pre-1990	26	0.19	0.18/0.21		0.28	0.18/0.38		23	0.26	0.20/0.31		0.20	0.05/0.35	
1990 or after	17	0.34	0.32/0.36		0.36	0.28/0.43		12	0.36	0.32/0.39		0.57	0.38/0.77	
FDK min per day	24			6.56**			5.75*	24			26.89***			5.30*
< 360	12	0.30	0.27/0.33		0.20	0.09/0.31		13	0.20	0.13/0.26		0.07	–0.11/0.25	
≥ 360	12	0.37	0.32/0.41		0.49	0.28/0.70		11	0.44	0.37/0.51		0.43	0.18/0.67	
FDK selection[a]	15			0.90			0.08	13			14.95***			0.00
Parent nomination	8	0.30	0.11/0.49		0.29	–0.16/0.73		7	0.55	0.40/0.70		0.32	–0.04/0.68	
Student characteristic	7	0.40	0.33/0.47		0.36	0.06/0.67		6	0.20	0.12/0.29		0.32	0.06/0.59	
Community	26			4.20*			0.35	17			14.12***			4.32*
Nonurban	8	0.34	0.21/0.47		0.43	0.19/0.67		7	0.23	0.15/0.31		0.18	0.02/0.35	
Urban	18	0.21	0.19/0.22		0.35	0.25/0.45		10	0.43	0.36/0.50		0.49	0.25/0.72	

Note. FDK = full-day kindergarten.

[a]For the FDK selection moderator analysis, samples were grouped according to their primary mode of selecting students into the FDK classrooms. There may have been additional selection methods used to the primary method of selection.

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

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

CI = 0.40/0.70) was greater than the effect for samples in which FDK students were selected by a student characteristic ($d = 0.20$, 95% CI = 0.12/0.29).

Community. The unadjusted effect sizes varied significantly by type of community under fixed-effect assumptions but not random-effects assumptions. Under fixed effect, the unadjusted effect size for nonurban samples ($d = 0.34$, 95% CI = 0.21/0.47) was greater than the effect size for urban samples ($d = 0.21$, 95% CI = 0.19/0.22).

The adjusted effect sizes when participants came from an urban community (fixed effect: $d = 0.43$, 95% CI = 0.36/0.50; random effects: $d = 0.49$, 95% CI = 0.25/0.72) were significantly greater than the effect sizes when participants came from a nonurban community (fixed effect: $d = 0.23$, 95% CI = 0.15/0.31; random effects: $d = 0.18$, 95% CI = 0.02/0.35), under both fixed-effect and random-effects models.

Outcome Moderators of the Association Between FDK and Achievement

Finally, we conducted moderator analyses of the association between FDK and achievement using two moderators related to outcome variable. Table 6 presents these results.

Achievement measure. Outcomes were grouped according to whether a nationally standardized measure of achievement was used versus all other achievement outcomes, including researcher-created scales, class grades, teacher reports, and parent reports. The unadjusted effect size significantly varied by type of outcome measure under fixed-effect assumptions but not random-effects assumptions. Under fixed-effect assumptions, the unadjusted effect size for standardized achievement measures ($d = 0.28$, 95% CI = 0.26/0.30) was greater than the effect for other achievement measures ($d = 0.22$, 95% CI = 0.21/0.24).

The adjusted effect sizes significantly varied by type of achievement measure under a fixed-effect but not random-effects model. Using the fixed-effect model, the adjusted effect size for standardized achievement measures ($d = 0.40$, 95% CI = 0.36/0.45) was greater than the effect for other achievement measures ($d = 0.28$, 95% CI = 0.25/0.32).

Subject matter. Outcomes were grouped according to whether a measure of verbal achievement was used—including reading or reading readiness, writing, and language development—compared to a mathematics measure of achievement. Outcomes that could not be categorized as verbal or mathematics, including those outcomes that were general measures of achievement or measures assessing multiple subjects, were excluded from this analysis.

Under the fixed-effect model, the unadjusted effect size on mathematics achievement ($d = 0.32$, 95% CI = 0.30/0.34) was significantly greater than the effect size for verbal achievement ($d = 0.17$, 95% CI = 0.15/0.18). The unadjusted effect size did not significantly vary by subject matter under the random-effects model. However, for adjusted effect sizes, this pattern was reversed. The adjusted effect sizes for FDK significantly varied by type of subject matter under fixed-effect assumptions but not under random-effects assumptions. Under the

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

TABLE 6

Outcome moderators of the effect of full-day kindergarten on immediate achievement

	Unadjusted *d*-index							Adjusted *d*-index						
		Fixed			Random				Fixed			Random		
Moderator	*k*	*d*	−/+CI	Q_b	*d*	−/+CI	Q_b	*k*	*d*	−/+CI	Q_b	*d*	−/+CI	Q_b
Achievement measure	47			18.30***			0.09	37			15.02***			0.07
Standardized	35	0.28	0.26/0.30		0.30	0.22/0.38		28	0.40	0.35/0.45		0.32	0.16/0.55	
Other	12	0.22	0.21/0.24		0.28	0.16/0.40		9	0.28	0.25/0.32		0.35	0.17/0.48	
Subject	44			162.58***			0.93	30			44.90***			3.47†
Verbal	26	0.17	0.15/0.18		0.28	0.19/0.37		21	0.40	0.37/0.44		0.46	0.32/0.61	
Mathematics	18	0.32	0.30/0.34		0.34	0.26/0.41		9	0.15	0.09/0.22		0.24	0.06/0.43	

†$p < .10$. ***$p < .001$.

fixed-effect model, the effect size for FDK on verbal achievement ($d = 0.40$, 95% CI = 0.37/0.44) was significantly greater than the effect size for mathematics achievement ($d = 0.15$, 95% CI = 0.09/0.22).

Moderators of the Association Between FDK and Nonachievement Measures

No moderator analyses were conducted for the nonacademic outcomes because the number of effect sizes (and effects sizes with the needed information) was too few.

Studies Using Data From the Early Childhood Longitudinal Studies–Kindergarten

In the fall of 1998 the National Center for Education Statistics sponsored a study that began collecting longitudinal data on the incoming class of kindergarten students. The Early Childhood Longitudinal Study–Kindergarten (ECLS-K) involved a nationally representative sample of more than 21,000 children (from an estimated 3,864,000 kindergarteners) and 3,300 of their teachers in 866 public and private schools.

The ECLS-K data represent the largest and broadest sample of kindergartener data assembled and the most in-depth data collection. Being a public database, it is not surprising to discover that it has been used in numerous studies that employed different subsamples of students and different analytic techniques. Many of these studies have included whether the children attended HDK or FDK programs as a variable in their analyses. Because they all draw data from the same sample of children, it would have been inappropriate to include all but one of them in the meta-analysis reported above. However, because studies using ECLS-K data test so many models and report rich data on numerous differences in FDK and HDK classrooms that go beyond achievement outcomes, a summary of these studies fills out our picture of difference between FDK and HDK in important ways.

Methods used in ECLS-K data collection. The ECLS-K kindergarten classes were in schools that also included other grades (some variation of 1st through 12th) or had preschool programs along with kindergarten. Stratified sampling was used in which counties or groups of counties were used to identify 1,277 public and private schools offering kindergarten. About two dozen children were then randomly chosen from each participating school, regardless of its size. Children were oversampled from private schools and from Asian or Pacific Islander ethnic groups. Weights were applied to adjust for the oversampling and to adjust for missing responses. These procedures allow for generalization to the national kindergarten population. A full description of the sample, sampling procedures, and measures can be found in the *Early Childhood Longitudinal Study: Kindergarten Class of 1998–99: Base Year Public-Use Data Files User's Manual* (National Center for Education Statistics, 2001) and *ECLS-K K–5 Longitudinal Electronic Files and Codebook* (National Center for Education Statistics, 2006).

Researchers returned to schools and again collected data on the students in the original sample at the end of their kindergarten year (spring 1999) as well as the fall (testing a 27% subsample for purposes of estimating summer learning loss) and spring of first grade and the spring of third and fifth grades. However, the

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

testing interval in kindergarten and first grade was closer to 6 months than 9 months, suggesting any effects of HDK versus FDK may be underestimated.

Data were collected about the children and schools on a wide variety of variables from the children themselves, from school administrators, from the children's teachers, and through computerized telephone interviews with parents. These data included assessments of children's cognitive ability, behavior problems, and self-perceptions as well as demographics. School-level variables included geographic location, type of community, and class size. Classroom variables included questions about the teacher's background as well as instructional techniques. Family variables included parents' perceptions of their child's progress along with other family variables, for example, maternal employment. Assessment of children's cognitive abilities was collected in reading, mathematics, and general knowledge in kindergarten and first grade and reading, math, and science in third and fifth grade. A two-stage, untimed assessment technique was used in early-grade reading and math in which all children took the same initial test along with a second test based on their initial performance, which could then be used to place all children on a common scale. The reading test included subtests on basic reading skill, vocabulary, and reading comprehension. The math test measured seven early skill domains.

Characteristics of FDK and HDK classrooms. In the ECLS-K sample, about 56% of children attended FDK programs and 44% attended HDK programs. Denton, West, and Walston (2003) reported that FDK was more likely to be provided in the southern region of the United States and in urban and rural communities than in suburban ones. One of the more extensive analyses for the base year was conducted by Walston and West (2004). These researchers looked at several school, family, child, and instruction differences between HDK and FDK. They found that Catholic schools were more likely to provide FDK and both public and private schools had higher concentrations of minority (especially Black) and low-income students in FDK than Catholic schools. Although FDK teachers engaged in more teacher-directed whole-class, small-group, and individual instruction, the proportion of the day spent in each type of instructional grouping was roughly the same for HDK and FDK. Lee, Burkam, Ready, Honigman, and Meisels (2006), also using ECLS-K data, found that even though FDK classes generally last about twice as long as HDK classes, the amount of instructional time spent in FDK classes was not twice that in HDK classes. Across all subject areas, HDK classes spent about 12.4 hours in instruction whereas FDK spent about 17 hours per week (also see Rathburn, Walston, & Germino Hausken, 2000). FDK students received about 30% more language and reading instruction and 46% more time on math. FDK was more likely to use instructional groupings based on the achievement level of the students. Walston and West (2004) also reported that FDK classes were more likely to provide math, social studies, and science instruction every day and were more likely to focus on several reading and math subskills at least weekly. Some of the math subskills covered weekly in FDK classes were more likely to be part of the first grade curriculum.

Outcomes differences. Appendix G provides a summary of studies that examined the effects of the length of the day on student outcomes at the end of kindergarten and includes information on the subsample of students and measures used and the

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

findings. These studies used several different methods to analyze the ECLS-K data, many of which are regression based. They have also used a wide variety of other variables in various combinations in the models they have built. Although this is certainly another positive aspect of these studies, it always remains possible that the employed models omitted critical variables, the absence of which led to spurious results.

Similar to the meta-analysis results, the studies using ECLS-K data generally reveal a greater positive gain in reading and math for children in FDK programs than HDK programs at the end of kindergarten or beginning of first grade. This is true even though FDK students entering kindergarten are more likely to come from minority and poor families and begin the year with lower test scores. However, the data also reveal that FDK children may not have as positive an attitude toward school and may experience more behavior problems in school.

Longitudinal Effects of FDK Versus HDK on Academic Achievement

Studies Using ECLS-K Data

Appendix H presents results for studies using ECLS-K data that measured FDK versus HDK effects in first grade and beyond. Although we must recognize that these studies are all drawing samples from the same database, they do employ different subsamples. They also use a variety of different control variables and analytic techniques. Taken as a whole, these studies consistently find that the positive effect of FDK on achievement dissipates over time and appears to vanish for all subgroups of students by the end of third grade.

Other Studies With Follow-Up Achievement Measures

In addition to the ECLS-K, we found 20 other studies that compared FDK to HDK and included measures of achievement taken at a time other than the end of the kindergarten year. These studies are contained in Appendix I.

Performing a direct combination of the results of these studies at different time points is problematic; different studies would contribute to estimates at different times, meaning each time point estimate would be based on different sample compositions. Thus, there would be no indication of the trend over time in the individual samples. However, a subanalysis could be performed on only those studies that included measurements at two or more time points. Examination of the nine studies in Appendix I that truly are longitudinal reveals no consistent temporal pattern. Some studies show increases in a positive effect of FDK over time, others show decreases in the positive effect, and others with measures at more than two time points reveal both increases and decreases in no consistent pattern. Thus, unlike the ECLS-K data, there is no clear indication that the positive effect of FDK either increases or diminishes over time.

Academic Achievement Outcomes of Alternate-Day FDK Versus (a) FDK Everyday or (b) HDK at the Conclusion of the Kindergarten Year

In these analyses, we examined studies that compared AD-FDK programs to half-day programs or every-day, full-day programs. In AD-FDK programs students attend school for approximately the same amount of time as students in

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

HDK, but they do so by attending on alternating days rather than for half days every day.

In most instances, AD-FDK programs are instituted to save on transportation costs. By switching to such a schedule, school districts eliminate the need for the midday busing of kindergarten students. Because students are still in school only half-time, all other costs remain the same as under an HDK schedule. On the negative side, parents of students in AD-FDK programs express concern about (a) finding child care for kindergarteners when they are not in school and (b) the impact of the irregular schedule on their children's sense of routine.

By comparing AD-FDK to HDK we could assess whether this effort to save money has implications for academic achievement. Also, by comparing AD-FDK to HDK and FDK we might get another indication of whether it is the arrangement of time in school or the amount of time in school that influences students' academic performance. If the former comparison reveals no differences and the latter suggests a positive effect of FDK, this would argue that it is the time students spend in school that provides the benefit for academics. If the reverse findings emerge, it would suggest that it is the arrangement of time that is beneficial.

We found 17 studies that made comparisons between HDK and AD-FDK. Descriptions of the individual studies can be found in Appendix J. With regard to the direction of findings and effect sizes, the results do not clearly favor either program type. The directions of effect in five studies favored AD-FDK on all measures of achievement, whereas two studies uniformly favored HDK. Of six studies that produced a mixed array of effect sizes, two had a preponderance of findings favoring HDK, two favored AD-FDK, and two revealed no favored direction. Four studies simply reported no differences. The average effect size across all studies favored AD-FDK over HDK ($d = 0.27$, 95% CI = 0.18/0.35, using a fixed-error model and $d = 0.20$, 95% CI = 0.00/0.39, using a random-error model). However, it should be kept in mind that four studies finding no difference could not be used in the calculation of the average effect size.

Comparisons of AD-FDK with FDK (Appendix K) are fewer in number ($k = 7$) but did appear to favor FDK over AD-FDK. Five studies revealed effect sizes uniformly favoring FDK on achievement measures (immediately following kindergarten), whereas two studies showed a preponderance of mixed results favoring HDK. The average effect size across all studies favored FDK over AD-FDK ($d = 0.31$, 95% CI = 0.16/0.46, using a fixed-error model and $d = 0.43$, 95% CI = 0.07/0.79, using a random-error model).

In sum then, there appears to be tentative evidence that FDK has a more positive impact on achievement than AD-FDK, which seems to differ little from HDK.

Discussion

There are several conclusions about the effects of FDK that seem warranted by the results of this research synthesis. First, *at the end of the kindergarten year children who attend full-day programs perform better on tests of academic achievement than children who attend only half-day programs.* Regardless of whether this overall association is statistically adjusted to control for other influences on achievement potentially confounded with the length of the kindergarten day and regardless of whether a fixed-effect or random-effects model is used, *at the end of the kindergarten year children in FDK programs appear to score about*

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

one fifth to one third of a standard deviation higher on academic tests than children in HDK programs.[6] And although there was evidence that the length of the kindergarten day might have different strengths of association with different subject matters, the association was positive and statistically significant for all tests, using adjusted and unadjusted effect sizes under fixed-effect and random-effects models on measures of verbal and mathematics achievement. All eight average estimates of association fell between $d = 0.15$ and 0.46. Also, the association of FDK and achievement was positive and significant using both standardized and unstandardized measures of achievement, with a stronger association for standardized tests (using fixed-effect assumptions). Finally, the meta-analysis indicated that over the past three decades as the movement to improve learning in schools has accelerated the difference between achievement in FDK and HDK has increased. Taken as a whole then, we find it reasonable to suggest that a first approximation for the positive association of FDK versus HDK on verbal or math achievement immediately after the kindergarten year would be about one quarter standard deviation. Expressing this association in a different metric, *the average (50th percentile) child in FDK performed better on academic tests than 60% of children in HDK at the end of the kindergarten year.*[7]

There were several sources of evidence to suggest that *the positive association of FDK with academic achievement is a result of the amount of time children spend in school rather than the arrangement of time within the school day.* The meta-analysis revealed that children in FDK programs that lasted more than 360 minutes (6 hours) a day showed a larger positive advantage on achievement tests relative to their HDK counterparts than did FDK students attending school for less than 360 minutes. Also, although the comparisons were varied and difficult to combine, it appeared that results of evaluations involving AD-FDK programs were consistent with this finding. That is, comparisons of AD-FDK with HDK revealed a less consistent direction of findings than AD-FDK versus FDK comparisons, which tended to favor FDK.

We should be quick to point out however that simply adding time to the kindergarten day will not inevitably improve children's achievement. As we stated in the introduction, both the proponents and the skeptics of FDK agree that the content and instructional strategies used in the kindergarten program are paramount to its success or failure. Although added time may create the *opportunity* for increased learning, it is how that time is used that will determine the FDK intervention's ultimate effectiveness.

Regrettably, we were not able to use meta-analytic techniques to compare the strength of the association between FDK and HDK separately for children from disadvantaged and middle-class homes. However, we were able to meta-analytically compare evaluations conducted in urban versus nonurban settings. Interestingly, a significant difference was found for the unadjusted measures of achievement using a fixed-effect model (but not a random-effects model) favoring FDK for nonurban children. *For adjusted measures FDK had a significantly stronger association with higher academic achievement for children attending programs in urban than in nonurban communities* using both fixed and random effect models. And the size of this difference was quite large, with the association for urban children being about twice the size for nonurban children. If we make the assumption that children in urban settings were more likely to come from poorer

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

homes, then this can be taken as indirect evidence for a potentially greater impact of FDK for poorer children, given that the relation is causal.

Our synthesis of the data analyses using the ECLS-K revealed that *although FDK children experienced greater growth during the kindergarten year, when subsequent years were added to the analyses, the HDK students showed a stronger growth trajectory and seemed to have made up any early FDK advantage by the end of third grade.* This finding likely will gain considerable attention in policy debates. Therefore, we return to it below for an extended discussion of its potential causes.

Finally, as noted above, the use of outcome measures other than achievement in evaluations of FDK is relatively infrequent. Just a few nonacademic outcomes have generated enough data to warrant mention. The results of the meta-analysis suggested that *the association of FDK versus HDK using nonacademic outcomes suggests no clear potential positive benefit.* Evidence about child independence was inconclusive. There appeared to be a positive association between FDK and child independence (about two fifths of a standard deviation), but only when using a fixed-effect model did this association reach statistical significance. Furthermore, there appeared to be just a small positive association between FDK and attendance (about one tenth of a standard deviation), although this small association was significant under both fixed-effect and random-effects assumptions. Evidence was more suggestive of a positive association of FDK with the child's self-confidence (slightly more than half a standard deviation) and ability to work and play with others (about one half to two thirds standard deviation). In contrast, the analyses of data from the ECLS-K suggested that FDK children may not have as positive an attitude toward school as HDK children and may experience more behavior problems in school. Again however, the nonexperimental nature of all these comparisons complicates their interpretation. To wit, it could be argued that because poor and minority students are more likely to be placed in FDK programs, we would expect children in these programs to perform worse on all of these measures, including attendance, self-confidence, and cooperative behaviors; the fact that they do not could be taken as evidence of a positive effect of FDK.

Two cautions need to be explicitly stated to properly contextualize the interpretation of the evidence we provide above. First, it needs to be stated that *the research designs used to compare the effects of FDK with HDK generally do not permit strong causal inferences.* Only one study employed random assignment of children to FDK and HDK programs. That said, the results of one randomized field trial did favor FDK. Also, the moderator analyses indicated that studies using matching revealed larger positive associations with FDK than studies not using matching, and matching on more variables revealed larger associations than matching on only one variable. Thus, although the studies do point to some tentative conclusions about the relative effects of FDK and HDK, more confident conclusions must await the conducting of multiple evaluations using stronger research designs, especially randomized field trails and regression discontinuity designs (Shadish, Cook, & Campbell, 2002).

Second, not surprisingly, *the effects of FDK and HDK on measures of academic achievement dominate the outcome measures used in research.* Many of the positive and negative effects of FDK versus HDK listed in Tables 1 and 2 are yet to be the focus of study, and those that have been studied have rarely appeared in more

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

than one or two evaluations. This is an important point to make because the lack of evidence on many outcomes should not be used to narrow the debate about the relative effects of the two instructional strategies. Such a narrowing can happen in two ways. First, because certain outcomes have not been studied they may incorrectly be ignored in policy discussions. Second, *the absence of evidence about an effect should not be taken as evidence of the absence of effect.* Because certain effects of FDK versus HDK have not been studied does not mean these potential outcomes are unimportant or are proven to be uninfluenced by the intervention; to the extent that they are valued as outcomes for children they continue to deserve a place in considerations of which kindergarten program is preferable for children. Regrettably, however, the research literature is silent on these issues.

Explanations for the "Fade-Out" of the FDK Effect

Assuming the relation is causal, there are at least three possible explanations for the finding that the academic advantage gained by children in FDK programs disappears by the end of the third grade. First, it may simply be that the FDK "fade-out" is just that: The effect of FDK becomes a smaller and smaller influence as children accumulate more and more experiences in an academic setting. This explanation would suggest that because of its transitory nature, the argument that FDK leads to improved academics is specious in the long run.

Second, it is possible that because of the positive effect of FDK on academics at the end of kindergarten and persisting into first and second grades, children who benefitted from FDK were less likely than HDK children to be eligible for other support services in later grades. Over 3 years, this difference in eligibility may have allowed the HDK students to "catch up" to the FDK students (for a similar argument made regarding the "fade-out" of summer school effects, see Cooper, Charlton, Valentine, & Muhlenbruck, 2000). In this instance then, the achievement trajectories of students in FDK and HDK might or might not have been similar without the added support, but eventually the FDK advantage bestowed on participants was negated by provision of other compensatory services to HDK children.

Third, Votruba-Drzal, Grining, and Maldonado-Carreño (2008) pointed out that between one half (for reading) and one quarter (for math) of the postkindergarten growth advantage for HDK was associated with child and family variables. This argument brings us back to the fact that for all its positive aspects, the ECLS-K was not an experimental study. Therefore, because FDK classes were more likely to serve minority and poor children, the postkindergarten trajectories may be expressing school obstacles these children more frequently face in subsequent years. So although the direct comparison of FDK and HDK students by third grade shows no difference, the trajectory of poor and minority children would suggest that had the FDK intervention been truly inert the FDK group would have been behind the HDK group. If this is the case, FDK had a positive effect on children by boosting their end-of-kindergarten achievement, but it did not change their achievement trajectory.

Determining which, if any, of these explanations for the FDK "fade-out" is supported by the evidence will require new research. The best approach to assessing the third—different trajectories—explanation will require the use of experimental designs involving random assignment of children to FDK and HDK programs. Only in this way can researchers generate estimates of trajectories for children

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

from different backgrounds that are independent from the impact of the kindergarten program. Such studies not only can estimate the immediate impact of FDK on children but also can determine whether FDK has an impact on the subsequent learning trajectory of children.

One impediment to conducting randomized field trials is the ethical concern of using chance to determine who does and does not receive a desirable service. The decision about which children will receive which kindergarten program typically is based on parent nomination or an administrative decision considering student characteristics. We suspect, however, that "natural experiments" (Shadish et al., 2002) occur in this area. That is, whenever an intervention leads more people to volunteer than can be accommodated and a lottery is used to choose participants, a natural experiment has occurred. Data collected from students who were assigned to kindergarten programs through lotteries can be used to obtain a stronger inference about the causal effect of the intervention than the designs that currently dominate FDK research.

Of course, natural experiments are not without problems. For example, in retrospective designs it may be difficult to track down children who received FDK and HDK many years ago. For both retrospective and prospective natural experiments, attrition and treatment cross-over can be problems. Furthermore, because these designs first require an overabundance of volunteers, the results may be generalizable only to children whose parents are interested in having them in FDK at all. Even with these limitations, the use of longitudinal data from natural experiments on FDK would be a welcomed addition to the evidence about this intervention and could give us a much needed assessment of the viability of not only FDK's immediate effects but also whether the apparent fade-out of the effect is real or a function of different educational trajectories associated with differential provision of the program.

The second explanation for the FDK "fade-out"—differential provision of post-kindergarten remedial services—could easily be studied as part of the natural experiments described above. Rather than collecting data on achievement measures only, these longitudinal studies would also collect data on the frequency with which FDK and HDK children receive support services in subsequent years. Also, it is possible to use the ECLS-K study to get a first indication of the viability of this hypothesis. This could be done by examining whether children who were in FDK received fewer remedial academic services after the kindergarten year. Of course because FDK students were more likely to come from poor and minority homes, this student characteristic would need to be controlled in the data analysis. We found no previous use of the ECLS-K data to address this issue.

Finally, given the data, the first explanation for the FDK "fade-out"—that the positive effect of FDK diminishes as the child accumulates more academic-related experiences through the first three school grades—is the alternative hypothesis to the first two; it is simply an unmediated description of the research finding. In other words, we have the evidence that the leveling effect occurs.

Implications for Policy and Practice

We would speculate that, in fact, all three of the explanations for the FDK "fade-out" are operative; it is not necessary that only one prevail. Perhaps the best way to envision the issue is to not think about an "average" FDK effect that is

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

similar for all children but to think of the children in FDK and HDK as a collection of individuals who will be differently influenced by the intervention and its implications for latter instruction. For some FDK students who are from poor or minority families, the years subsequent to kindergarten will be filled with additional challenges that erode the academic advantage they obtained in FDK relative to their more fortunate and majority counterparts in HDK. For some HDK children who are struggling with academic material, the early "boost" provided by FDK they did not receive may be provided by services that come later in their schooling. And for some FDK and HDK students their experiences in later grades (e.g., the superior or subpar first grade teacher, exceptional or disruptive classmates) may serve to level out their achievement, diminishing the impact of their differential kindergarten experience. There is no reason why all three of these processes might not operate simultaneously but in different magnitudes depending on the child and the experiences their schools provide.

The implications for policymakers of these findings and our interpretation of them seem clear. First, FDK may not be the "magic bullet" that alters permanently poor and minority students' academic trajectories. Rather, with regard to academics, *FDK is probably best viewed as one in a continuing series of interventions needed to alter the academic success of students who enter and continue through school with disadvantages.* Furthermore, *other support services provided in later grades may be capable of providing similar levels of academic compensation when FDK is not an option for struggling students.*

The results of research involving outcomes of FDK other than academics are sparse and inconsistent. However, it seems to us that the concerns voiced by many educators and scholars regarding the "hurried child" (Elkind, 2006) remain legitimate; there certainly are some children who are not ready for 6 hours of schooling at the age of 5. These children may not possess the emotional regulation or social maturity needed to benefit from the longer day of instruction. Their attitudes toward school may suffer, and they may be more prone to behavior problems when their attention wanders from the lesson.

That said, it is clear that parents with children in FDK have far more positive attitudes toward FDK than parents of children in HDK. And the availability of FDK serves an important need for single-parent families and families in which both parents work outside the home. What this suggests is that FDK ought to be an option available to families who desire it. Perhaps optimally, decisions about whether a child will benefit from FDK might best be made by parents in consultation with educators, who might assess the social and emotional maturity of children as well as their readiness to learn.

Conclusion

Similar to research on so many other educational programs and interventions, the evidence on the impact of FDK leaves much to be desired. The research designs are weak for making strong causal inferences, and many of the proposed outcomes for FDK are yet to be the focus of study. One positive note is the existence of the ECLS-K study, which provides a longer view of the impact of FDK than is available for many other interventions. Taken as a whole, however, we find that the evidence on FDK—again similar to so many other educational interventions—suggests that its overall effect on academic outcomes will be evident at first but

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

fade over time. For some students, its impact might even be negative on some nonacademic outcomes. Thus, FDK is not a magic bullet that will render meaningless all the impediments to school success that struggling students will face.

Instead, it would be best to think of FDK as one component in the array of interventions parents and educators can use to help all children grow to their full potential. It should be available to all but not necessarily universally prescribed for all. For some struggling students, it will prove an initial boost that, when coupled with other interventions initiated in later years, can help them overcome deficits that otherwise would be predictive of academic failure. For others it may be too much, too soon. Now the task for education researchers is to develop evidence-based algorithms that help educators and parents predict when, and for whom, the FDK experience will be most beneficial.

Notes

The authors wish to thank Jill Kahane and Megan Kuhfeld for their assistance with data retrieval. Correspondence should be sent to Harris Cooper at cooperh@duke.edu.

[1]Half-day kindergarten (HDK) programs typically last about 3 hours and full-day kindergarten (FDK) about 6 hours. Although there is variation around the length of HDK and FDK programs (and thus differences from study to study), we doubt that any program labeled HDK will be as long as an FDK program.

[2]Although it would have been ideal to include studies conducted in other countries, these were simply too few to permit us to include them for purposes of comparison to U.S. and Canadian studies. To include them and claim that findings thus generalized across cultures would have been unwarranted.

[3]For purposes of isolating the effects of an extended day, the provision of additional services and different curricula along with FDK prevents the estimation of the unique effect of extending the kindergarten day. By excluding these studies we can assume that HDK students received the same services as FDK students and were taught using the same curriculum. Thus, any effect is from FDK alone. This was not true in the excluded studies.

[4]Readers can find tables containing descriptions of the individual studies that were included in the meta-analyses on immediate academic outcomes (Appendix B) and immediate nonacademic outcomes (Appendix C).

[5]In the text, only weighted average *d*-indexes and their confidence intervals are reported. The number of samples involved in the analyses and the results of the tests for homogeneity (and their significance levels) can be found in Tables 4 to 6.

[6]The one study that employed random assignment of children to FDK and HDK programs (Elicker & Mathur, 1997) found a positive effect of FDK on reading achievement of approximately three quarters of a standard deviation. This estimate was based on an adjusted effect size, controlling for a prekindergarten screening test and family income.

[7]Significant fixed-effect analyses for study sample size suggested that smaller studies produced larger associations, often an indication that some data censoring (of smaller studies with smaller effect sizes) has occurred. However, the trim-and-fill analysis revealed an estimated fixed-effect association of $d = 0.23$ with imputed values added. And the effect size estimate for studies with samples greater than 200 children ranged from $d = 0.11$ to $d = 0.22$, leaving our rough one quarter standard deviation estimate unchanged.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

References

Barnett, V., & Lewis, T. (1994). *Outliers in statistical data analysis* (3rd ed.). New York: John Wiley.

Beatty, B. (1995). *Preschool education in America.* New Haven, CT: Yale University Press.

Berson, M. P. (1968). All-day kindergarten: Opinions differ. *Today's Education, 57,* 27–29.

Borenstein, M. (2009). Effect sizes for studies with continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221–252). New York: Russell Sage.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). Comprehensive Meta-Analysis (Version 2.1) [Computer software]. Englewood, NJ: BioStat.

Brannon, D. (2005). Kindergarten: What parents pick and why. *Education Digest, 70,* 57–62.

Brewster, C., & Railsback, J. (2002). *Full-day kindergarten: Exploring options for extended learning.* Portland, OR: Northwest Regional Educational Laboratory.

Brosterman, N. (1997). *Inventing kindergarten.* New York: Harry N. Abrams.

Child Trends Data Bank. (2003). *Full-day kindergarten.* Retrieved November 20, 2008, from http://www.childtrendsdatabank.org/indicators/102Fulldaykindergarten.cfm

Clark, P. (2001). *Recent research on all day kindergarten.* Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education.

Clark, P., & Kirk, E. (2000). All-day kindergarten. Review of research. *Childhood Education, 76*(4), 228.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.

Cooper, H., Charlton, K., Valentine, J., & Muhlenbruck, L. (2000). *Making the most of summer school.* Monographs of the Society for Research in Child Development. Malden, MA: Blackwell.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *Handbook of research synthesis and meta-analysis.* New York: Russell Sage.

Cruikshank, S. B. (1986). All-day kindergarten: What every parent should know. *PTA Today, 11,* 12–13.

Denton, K., West, J., & Walston, J. (2003). *Reading—Young children's achievement and classroom experiences: Findings from the Condition of Education 2003* (NCES 2003–070). Washington, DC: U.S. Department of Education.

Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95,* 89–98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 276–284.

Education Commission of the States. (2004). *The progress of education reform 2004: Kindergarten: Full-day kindergarten programs improve chances of academic success.* Denver, CO: Author.

Elicker, J. (2000). Full-day kindergarten: Exploring the research. From inquiry to practice. Bloomington, IN: Phi Delta Kappa International.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Elicker, J., & Mathur, S. (1997). What do they do all day? Comprehensive evaluation of a full-day kindergarten. *Early Childhood Research Quarterly, 12*, 459–480.

Elkind, D. B. (2006). *The hurried child.* Cambridge, MA: De Capo Press.

Good, L. (1996). *Teacher perceptions of the all-day, alternating day kindergarten schedule.* Brookings: South Dakota State University.

Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *Handbook of research synthesis and meta-analysis* (2nd ed., pp. 417–434). New York: Russell Sage.

Grubbs, F. E. (1950). Sample criteria for testing outlying observations. *Journal of the American Statistical Association, 21*, 27–58.

Gullo, D. F. (1990). The changing family context: Implications for the development of all day kindergarten. *Young Children, 45*, 35–39.

Hedges, L. V., & Vevea, J. L. (1998). Fixed and random effects models in meta-analysis. *Psychological Methods, 3*, 486–504.

Jones, S. S. (2002). The effect of all-day kindergarten on student cognitive growth: A meta analysis. (Unpublished doctoral dissertation). University of Kansas.

Karweit, N. (1992). The kindergarten experience. *Educational Leadership, 49*, 82–86.

Le, V. N., 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.

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*, 163–208.

Miller, A. (2002). *Frequently requested information: Full-day kindergarten.* Champaign, IL: ERIC Clearinghouse on Elementary and Early Childhood Education.

Mink, G. W. (1937). An all day kindergarten in a rural consolidated school. *Childhood Education, 14*, 262–265.

National Center for Education Statistics. (2001). *Early Childhood Longitudinal Study: Kindergarten Class of 1998–99: Base year public-use data files user's manual* (NCES 2001–029). Washington, DC: Government Printing Office.

National Center for Education Statistics. (2006). *ECLS-K K–5 longitudinal electronic files and codebook* (NCES 2006032-2). Washington, DC: Government Printing Office.

Ohio State Legislative Office of Education Oversight. (1997). *An overview of full-day kindergarten.* Columbus: Author.

Olsen, D., & Zigler, E. (1989). An assessment of the all-day kindergarten movement. *Early Childhood Research Quarterly, 4*, 167–186.

Pennsylvania Partnership for Children. (2003). *Every child, every school, success from the start. Special report.* Harrisburg, PA: Author.

Plucker, J. A., & Zapf, J. S. (2005). *Short-lived gains or enduring benefits? The long-term impact of full-day kindergarten.* Bloomington, IN: Center for Evaluation and Education Policy, Indiana University.

Puleo, V. T. (1988). A review and critique of research on full-day kindergarten. *The Elementary School Journal, 88*, 427–439.

Rathburn, A. H., Walston, J. T., & Germino Hausken, E. (2000, April). *Kindergarten teachers' use of developmentally appropriate practices: Results from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–1999.* Paper presented at the annual conference of the American Educational Research Association, New Orleans, LA.

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Shapiro, M. S. (1983). *The child's garden: The kindergarten movement from Froebel to Dewey.* University Park: Pennsylvania State University Press.

Votruba-Drzal, E., Grining, L., & Maldonado-Carreño, C. (2008). A developmental perspective on full- versus part-day kindergarten and children's academic trajectories through fifth grade. *Child Development, 79,* 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: National Center for Education Statistics.

Weast, J. D. (2001). Why we need rigorous full-day kindergarten. *Principal, 80,* 6–9.

Weber, E. (1984). *Ideas influencing early childhood education.* New York: Teachers College Press.

WestEd. (2005). *Full-day kindergarten: Expanding learning opportunities* (Policy brief). San Francisco: Author.

Authors

HARRIS COOPER is professor and chair of the Department of Psychology & Neuroscience at Duke University, Box 90086, Durham, NC 27708-0086; e-mail: *cooperh@duke.edu.*

ASHLEY BATTS ALLEN is a graduate student in social psychology at Duke University, Box 90086, Durham, NC 27708-0086; e-mail: *ashleybatts@gmail.com.*

ERIKA A. PATALL is assistant professor of educational psychology at the University of Texas at Austin, 1 University Station, D5800, Austin, TX 78712-0383; e-mail: *erika.patall@mail.utexas.edu.*

AMY L. DENT is a graduate student in developmental psychology at Duke University, Box 90086, Durham, NC 27708-0086; e-mail: *ald.duke@gmail.com.*

This content downloaded from 155.33.16.124 on Mon, 3 Mar 2014 14:48:47 PM
All use subject to JSTOR Terms and Conditions