Academic Performance Gap Between Summer-Birthday and Fall-Birthday Children in Grades K–8

T. C. OSHIMA
CHRISTOPHER S. DOMALESKI
Georgia State University

ABSTRACT **Much interest exists among parents and researchers regarding the benefits and drawbacks of delaying kindergarten entrance to acquire academic advantage ("redshirting"). How evident is this assumed advantage at the kindergarten level and beyond? The authors evaluated large-scale test data from Grades K–8 to investigate the difference in performance between younger children (summer birthday) and older children (fall birthday). The performance gap evident in kindergarten decreased rapidly in Grades 1–3 but persisted up to Grade 5, until leveling off at middle school. The performance gap in the early grades that resulted from birth date was much larger than was the gap caused by gender difference.**

Key words: academic performance gap, fall- and summer-birthday students, Grades K–8

Headlines in the Atlanta Journal-Constitution ("Kindergarten," 2003) read, "Kindergarten: Is older wiser?" "Some say 'redshirting' improves readiness." The article featured success stories for children with summer birthdays who delayed entrance to kindergarten. Parents of preschoolers, after reading newspaper articles or talking with other parents, or both, are keenly aware that some parents purposefully delay their child's entrance to kindergarten. A number of parents begin the process even earlier by having their child repeat a 3-year-old preschool program. That phenomenon, especially common in affluent communities and among boys, causes parents with a 5-year-old summer-birthday child to wonder whether they should delay enrollment even when their child seems ready for kindergarten.

According to the newspaper article mentioned in the preceding paragraph, "The National Center of Education Statistics reports that 6% to 9% of kindergarten-aged children in the United States start a year late." However, those figures may be even higher. *Gifted Child Today* ("Delaying Entrance Into Kindergarten," 2004) reported that "According to the Census Bureau, 22% of first graders were 7 or older in Oct. 2002, up from 13% in 1970." "Redshirting" is more common among White parents than it is among parents from other races (Diamond, Reagan, & Bandyk, 2000).

Abundant early childhood research concerns kindergarten readiness (e.g., Andrews & Slate, 2002; Kurdek & Sinclair, 2001), as well as contradicting reports on the benefits of redshirting. Although most researchers seem to agree with the short-term academic and behavioral benefits of redshirting, there is a wide range of conflicting reports on its long-term benefits. Rusch (1998), for example, argued that being an older student in a class can backfire, especially in the upper grades. Conversely, Deutsch (2003) suggested that the youngest children in a classroom suffer more psychiatric disorders than do other students. The benefit (or harm) of redshirting is just a small part of the problem for an individual. At the classroom level, redshirting widens the age range for students and makes teaching practices more difficult. The presence of substantially older children (i.e., more than a year older) can affect other children academically and socially. Therefore, the increasingly popular trend of delaying kindergarten entry must be evaluated carefully.

To investigate redshirting, researchers should first determine whether there is an academic performance difference between children with summer birthdays and children with fall birthdays. We expected a large difference at the kindergarten level because educators and parents agree that 1 year can make a substantial difference in learning at a young age. Several studies confirm the effect of age on academic achievement at the kindergarten level (e.g., Kurdek & Sinclair, 2001; Meisels, 1992). It is not clear, however, when or if the difference in academic performance diminishes as students grow older. *Gifted Child Today* ("Delaying Entrance Into Kindergarten," 2004) reported that the academic advantages of redshirting diminish by the third grade, whereas Crosser (1991) stated that the difference is still distinct for fifth or sixth graders, especially for boys. Kurdek and Sinclair (2001) found no effects of age for either reading or mathematics achievement at fourth grade.

Address correspondence to T. C. Oshima, Department of Educational Policy Studies, College of Education, P. O. Box 3977, Atlanta, GA 30302-3977. (E-mail to oshima@gsu.edu)

Our purposes in this study were as follows:

1. To first delineate the difference at the kindergarten level; and
2. To subsequently investigate the difference through the elementary and middle school grades by using a statewide criterion reference test.

Method

Data

We used data for kindergarten students from a large dataset, "The Early Childhood Longitudinal Study of the Kindergarten Class of 1998–1999" (National Center for Education Statistics, 2001). From 21,260 kindergarten students, we identified 3,862 students as "younger children with summer birthdays (June, July, and August)" and 2,693 students as "older children with fall birthdays (September, October, and November)." To qualify for the younger summer-birthday group, the child had to be a first-time kindergarten student who was born in June, July, or August and be less than 67 months of age at the time of fall testing. To be in the older fall-birthday group, the child had to be a first-time kindergarten student who was born in September, October, or November and be 67 months of age or older at the time of fall testing. Test data came from two test administration periods (fall and spring). We collected data for elementary and middle grades from a statewide criterion-referenced test administered to approximately 115,000 students per grade in spring 2002. In this southeastern state, children had to have been 5 years old as of September 1 to enter kindergarten. Therefore, children born in the summer were the youngest children in the kindergarten class unless they were held back for a year. For each grade, about 10,000 children qualified as "younger children with summer birthdays"; another 10,000 qualified as "older children with fall birthdays." From each pool of approximately 10,000 students, we randomly selected 3,000 students for the present study. The state reported test scores in Grades 1–8 in scaled scores; a score of 300–349 indicated "meeting expectations," and a score of 350 and over indicated "exceeding expectations."

Analysis

We examined students' reading and mathematics performance in this cross-sectional study. For kindergarten students, we added affective variables (approaches to learning, self-control, and social interaction), as well as physical variables (height and weight). The study also includes demographic variables—race and gender. Race included eight levels for kindergarten students: (a) White, (b) African American, (c) Hispanic—Race Specified, (d) Hispanic—Race Not Specified, (e) Asian, (f) Native Hawaiian, (g) Other—Pacific Islander, (h) American Indian or Alaska Native, and (i) Multiracial, and six levels for Grade 1–8 students: (a) Asian, (b) African American, (c) Hispanic, (d) Native American/Alaskan, (e) White, and (f) Multiracial.

In the first part of the analysis, we investigated the mean difference between the two groups of interest for the age variable (summer- vs. fall-birthday children). To make the comparison meaningful across all kinds of measures and tests, we used effect size (*ES*) as an indicator for the difference throughout this study, as well as the independent *t* test. We calculated *ES* as the mean difference of two groups over pooled standard deviation. We plotted *ES* across grades to observe the general trend over time. We repeated the analysis within each gender group.

In the second part of the analysis, we used multiple regression to compare the relationships between three independent variables (race, gender, and age) and the dependent variables (reading or mathematics) to identify how age affected the test scores in relation to other demographic variables. Race, gender, and age were categorical variables that we transformed into *k* dummy variables where *k* was the number of categories, minus 1. For race variables for kindergarten students, we created seven dummy variables. For Grades 1–8, we transformed the race variable into five dummy variables (see Pedhazur, 1997 for procedures regarding dummy coding).

Results

Figure 1 shows *ES* for the difference between two groups (older fall-birthday children and younger summer-birthday children) for kindergarten students. Positive *ES* indicates that the mean for the older children was higher than that of the younger children. Stevens (1999) posited that *ES* = .20 is small, *ES* = .50 is medium, and *ES* > .80 is large. Furthermore, Stevens explained that medium *ES* is apparent

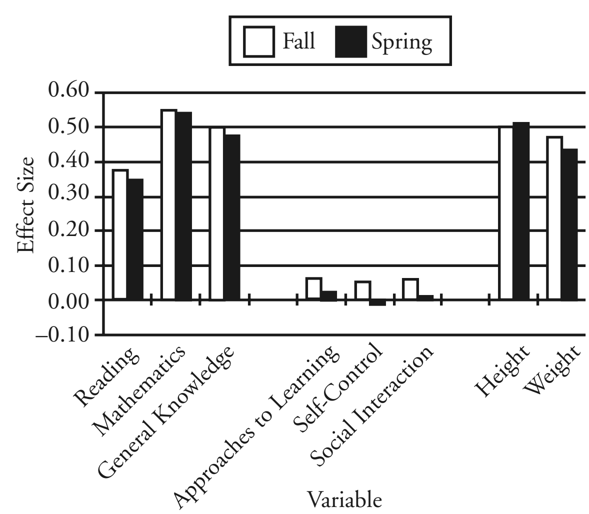

FIGURE 1. Effect sizes for younger versus older kindergarten students by academic, affective, and physical variables.

to a researcher. Stevens provided the IQ difference between semiskilled workers and professionals or managers as an example of medium *ES* (.50). On the basis of Stevens's criteria, the differences of those two groups are considered to be average for academic variables, as well as for physical variables.

Our data show that older kindergarten students are 1.95 in taller than are their younger peers in the fall (*ES* = .50). In other words, older kindergarten students average one half of a standard deviation taller than younger kindergarten students. That physical difference is probably apparent to parents and educators. For academic variables, *ES* also is about .50 (.38 for reading, .55 for mathematics, and .50 for general knowledge). In other words, differences existed between the two groups in academic areas as much as in the physical areas. The differences observed for affective variables (approaches to learning, self-control, and social interaction) were small (see Figure 1). Also shown in Figure 1 is the slight decrease in *ES* from fall to spring, except *ES* for height, possibly indicating a slight narrowing of the gap during the kindergarten year.

Table 1 reports the means and standard deviations for reading and mathematics scores for the two groups across kindergarten through Grade 8. For kindergarten, the scores are expressed in a *T* score. The *T* score is a standardized score with a mean of 50 and a standard deviation of 10. Table 1 also lists *p* values for the test of equal means (independent *t* tests) and equal variances (Levine's test). Because of a large number of significance tests combined with large sample sizes, we used a conservative level of alpha (.001) as an indicator of significance. The differences of means were significant from kindergarten to Grade 5 for reading and mathematics. The differences were not significant for Grades 6–8. Also noteworthy was a consistent

TABLE 1. Means and Standard Deviations for Reading and Mathematics Scores

	Reading			Mathematics		
Grade	Younger	Older	*p*	Younger	Older	*p*
Kindergarten[a]						
Fall						
M	44.82	50.75	<.001	47.09	53.48	< .001
SD	16.37	14.61	<.001	11.77	11.52	.901
Spring						
M	46.77	51.62	<.001	48.13	53.85	< .001
SD	14.57	13.05	<.001	10.67	10.32	.430
1[b]						
M	338.31	349.67	<.001	327.85	339.46	< .001
SD	36.30	38.37	.072	29.68	31.98	< .001
2						
M	339.80	348.38	<.001	328.89	335.51	< .001
SD	41.28	43.24	.141	28.87	31.84	< .001
3						
M	338.26	343.61	<.001	328.81	332.47	< .001
SD	34.93	35.75	.293	29.04	29.56	.241
4						
M	343.09	348.37	<.001	317.45	320.64	< .001
SD	43.91	46.84	.002	30.60	31.97	.041
5						
M	334.24	338.06	<.001	324.50	327.94	< .001
SD	33.07	33.62	.239	28.82	29.52	.210
6						
M	343.24	344.68	.205	324.02	324.83	.361
SD	42.78	45.01	.013	33.74	35.01	.018
7						
M	339.96	341.26	.145	321.38	322.61	.072
SD	33.63	35.17	.034	25.73	27.36	.001
8						
M	350.20	351.36	.339	320.49	320.22	.759
SD	46.57	47.51	.402	32.75	33.76	.132

Note. Younger = children with summer (June, July, August) birthdays; Older = children with fall (September, October, November) birthdays.
[a]For kindergarten, *n* (younger students) = 3,862 and *n* (older students) = 2,693 in the fall; *n* (younger students) = 3,718 and *n* (older students) = 2,586 in the spring. [b]For Grades 1–8, *n* (younger students) = 3,000 and *n* (older students) = 3,000.

trend of standard deviations. For Grades 1–8, although seldom reaching significance, the younger group's variation was less than was that of the older group's variation. That trend reversed for kindergarten students.

To track mean differences over time, we plotted *ES* against Grades 1–8 (see Figure 2; *ES* likely declined rapidly during Grades 1–3). However, a small *ES* (.10) remained during Grades 3–5. *ES* reached near zero for Grades 6–8. Figure 2 also illustrates graphs for girls and boys. For girls, a fairly large *ES* exhibited for mathematics in Grade 1 quickly declined by Grade 2. For boys in Grades 2–8, reading tended to exhibit larger *ES* values than did mathematics, indicating that being a younger aged student may have more effect on reading than on mathematics.

Table 2 shows the results from regression analyses. The R^2 values are reported under race, gender, and age. For example, the R^2 value of .182 under race for kindergarten (fall) for reading indicates that 18.2% of variation in reading can be explained by determining race alone. It is similar that 0.03% and 3.3%, respectively, of variation in reading can be explained by gender and age alone. The R^2_{change} refers to the difference in R^2 between the full model (race, gender, and age as independent variables) and the reduced model (race and gender as independent variables), indicating the unique contribution of age in addition to race and gender. In other words, a significant R^2_{change} indicates that adding age in the model would significantly improve the prediction of the test score when the model already contains race and gender.

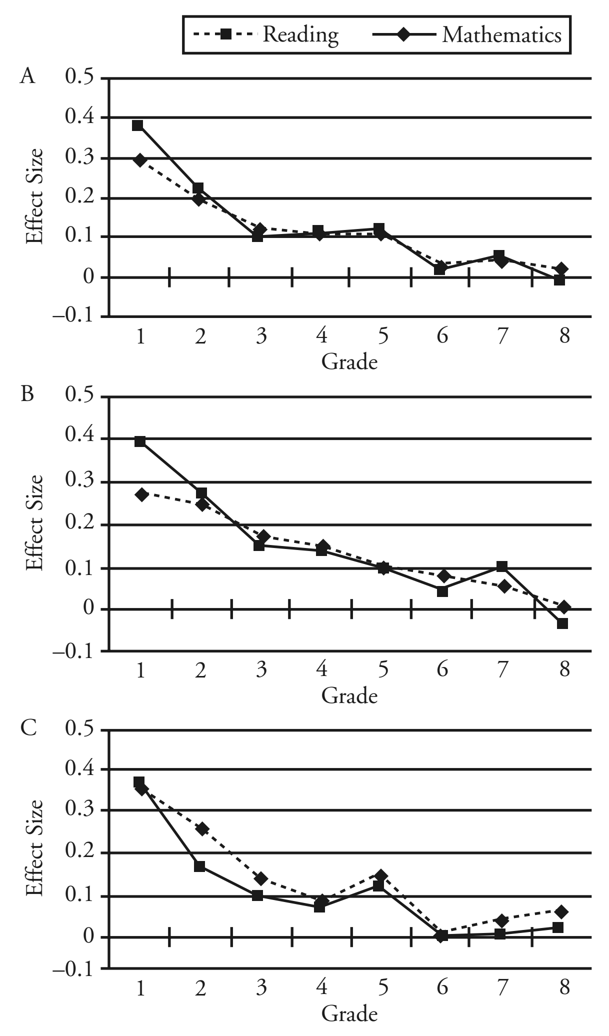

FIGURE 2. Effect sizes for younger versus older children over time (Grades 1–8) for combined sample, female sample, and male sample.

One can make several observations from Table 2: First, age (whether a child was younger or older in the class) was a significant predictor for reading and mathematics through Grade 5. Significant R^2_{change} for Grades K–5 also indicated that age improved the prediction after controlling for race and gender. Second, at any given grade, race always had the highest R^2 with the test scores. The effect of race was fairly stable across Grades K–8. Third, gender was a significant predictor for reading but not for mathematics. Fourth, the order of strength for reading in terms of relationship with test scores was race, age, then gender through second grade. For Grades 3–5, the order was race, gender, then age; for Grades 6–8, race then gender. In other words, age was a better predictor of reading than was gender through Grade 2; gender became a better predictor than age for Grades 3–5. Fifth, the order of strength for mathematics was race, then age through Grade 5; race was the only significant predictor for Grades 6–8. For example, at the beginning of the kindergarten year, 10.6% of variation in mathematics scores could be explained by race, and 6.8% could be explained by age. Gender explained little (< 0.1%) variation.

Discussion

For kindergarten students, the difference observed for academic areas (reading, mathematics, and general knowledge) was as large as was the difference for the height of children from the two groups. By spring, the academic difference narrowed somewhat. For elementary school children, *ES* of the difference of the two groups across five grades declined. There was a rapid decrease up to the third grade and a gradual or no decrease between the third and fifth grades; the difference still existed at fifth grade. Conversely, the difference was near zero in the middle school years. Although redshirting is more common for boys (Bent, May, & Kundert, 1996), the gap between the two groups was not always more pronounced for boys.

How large was the impact of age, whether the child was younger or older in the class, on test scores when compared with other demographic variables, such as race and gender? Whereas race explained 10–15% of variation in test scores in general, age explained up to about 7% of variation, depending on the grade level. That percentage was much larger than the percentage that we obtained from gender

TABLE 2. R^2 **Values for Race, Gender, and Age, With Reading or Mathematics as the Dependent Variable**

	Reading				Mathematics			
Grade	Race	Gender	Age[a]	R^2	Race	Gender	Age[a]	R^2
Kindergarten								
Fall	.182*	.003*	.033*	.022*	.106*	< .001	.068*	.053*
Spring	.128*	.005*	.028*	.020*	.095*	< .001	.067*	.053*
1	.069*	.016*	.022*	.020*	.084*	< .001	.034*	.031*
2	.084*	.006*	.010*	.011*	.118*	.002	.012*	.011*
3	.104*	.014*	.006*	.006*	.131*	.001	.004*	.004*
4	.120*	.008*	.003*	.003*	.118*	< .001	.003*	.002*
5	.102*	.004*	.003*	.003*	.159*	.001	.003*	.002*
6	.128*	.007*	< .001	< .001	.112*	< .001	< .001	< .001
7	.140*	.009*	< .001	< .001	.159*	< .001	< .001	< .001
8	.106*	.007*	< .001	< .001	.128*	< .001	< .001	< .001

[a]Refers to either younger or older students. [b]R^2_{change} refers to the difference in R^2 between the full model (race, gender, and age) and the reduced model (race and gender), indicating the unique contribution of age above and beyond race and gender.
*p = .001.

(1% for reading and <.01% for mathematics). In other words, age difference was as important, or more important, than was gender difference in the early grades.

A plethora of research exists on gender differences. For example, many researchers have shown that reading scores for boys are far behind those of girls (e.g., Gambell & Hunter, 1999; Phillips, Norris, Osmond, & Maynard, 2002; Sommers, 2000), a phenomenon that we confirmed. All through Grades K–8, girls outperformed boys. Researchers also suggest that boys outperform girls in mathematics, perhaps more so during adolescence (e.g., Hedges & Nowell, 1995; Royer & Wing, 2002). In this study, however, we did not observe gender difference for mathematics. Researchers should acknowledge gender difference in academic performance, as well as identify any difference that may exist by birth date. We showed that differences of birth date can be a much larger factor regarding academic performance than gender differences in the early grades and that the differences continue throughout the elementary school years.

We provide data on the issues related to the controversial practice of academic redshirting. It was not our intention to argue for or against the practice. We did not examine the population that delayed kindergarten entrance. Redshirting is based partly on the assumption that older children have an advantage over younger children in a classroom. We delineated how much difference there actually is in a given grade.

Numerous issues need to be considered before a decision is made regarding whether a parent should hold a summer-birthday child back for a year: (a) Does the child tend to learn better with older peers? (b) Does the child enjoy interacting with older children? (c) Is she or he easily hurt by not being able to perform as well academically (and physically) as her or his peers? and (d) Is the child a boy, and does he have a reading problem? Although any decision concerning academic redshirting at the kindergarten level (or at a younger age) should be individually based, research in these areas (such as type of children who thrive in an environment in which they are the youngest) should help parents make an intelligent decision concerning redshirting, which is a decision that cannot be reversed once it is made.

We showed the magnitude and duration of the academic gap exhibited by age differences of 7 to 11 months (summer-birthday children vs. fall-birthday children). With those objective data, researchers can investigate the practice of redshirting, which has a profound impact on individuals, as well as on teachers and schools.

REFERENCES

Andrews, S. P., & Slate, J. R. (2002). Public and private prekindergarten programs: A comparison of student readiness. *Educational Research Quarterly, 25*, 59–73.

Bent, D., May, D. C., & Kundert, D. K. (1996). The incidence of delayed school entry: A twelve-year review. *Early Education and Development*, *7*, 121–135.

Crosser, S. L. (1991). Summer birth date children: Kindergarten entrance age and academic achievement. *The Journal of Educational Research*, *84*, 140–146.

Delaying Entrance Into Kindergarten. (2004, Summer). *Gifted Child Today, 27*, 6.

Deutsch, N. (2003). Youngest kids in class suffer more psychiatric disorders. *Medical Post*, *39*, 47.

Diamond, K. E., Reagan, A. J., & Bandyk, J. E. (2000). Parents' conception of kindergarten readiness: Relationships with race, ethnicity, and development. *The Journal of Educational Research*, *94*, 93–100.

Gambell, T. J., & Hunter, D. M. (1999). Rethinking gender differences in literacy. *Canadian Journal of Education, 24*, 1–16.

Hedges, L., & Nowell, A. (1995). Sex differences in mental test scores, variability and numbers of high-scoring individuals. *Science*, *269*, 41–45.

Kindergarten: Is older wiser? (2003, September 20). *The Atlanta Journal-Constitution*, p. G1.

Kurdek, L. A., & Sinclair, R. J. (2001). Predicting reading and mathematics achievement in fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology*, *93*, 451–455.

Meisels, S. (1992). Doing harm by doing good: Iatrogenic effects of early childhood enrollment and promotion policies. *Early Childhood Research Quarterly*, *7*, 155–174.
National Center for Education Statistics. (2002). *ECLS-K base year public-use data files and eletronic codebook*. Washington, DC: Author.
Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York: Holt, Rinehart and Winston.
Phillips, L. M., Norris, S. P., Osmond, W. C., & Maynard, A. M. (2002). Relative reading achievement: A longitudinal study of 187 children from first through sixth grades. *Journal of Educational Psychology*, *94*, 3–13.
Royer, J. M., & Wing, R. E. (2002). Making sense of sex differences in reading and math assessment: The practice and engagement hypothesis. *Issues in Education*, *8*, 77–85.
Rusch, L. (1998). Delaying kindergarten. *Parents*, *73*, 129.
Sommers, C. H. (2000). *The war against boys: How misguided feminism is harming our young men*. New York: Simon & Schuster.
Stevens, J. (1999). *Intermediate statistics: A modern approach* (2nd ed.). Hillsdale, NJ: Erlbaum.