Full-day kindergarten and student literacy growth: Does a lengthened school day make a difference?

Keith Zvoch\textsuperscript{a,}*\textsuperscript{,}, Ralph E. Reynolds\textsuperscript{b,}, Robert P. Parker\textsuperscript{c}

\textsuperscript{a} Department of Educational Leadership, University of Oregon, College of Education, 5267 University of Oregon, Eugene, OR 97403-5267, United States
\textsuperscript{b} Department of Educational Psychology, University of Nevada, Las Vegas, United States
\textsuperscript{c} Center for Evaluation and Assessment, University of Nevada, Las Vegas, United States

Received 19 July 2006; received in revised form 9 August 2007; accepted 11 August 2007

Abstract

In the context of a quasi-experimental research design, literacy data obtained on students were examined to assess relationships between kindergarten program model (full- vs. half-day) and student literacy outcomes. Application of multilevel modeling techniques to the time series data collected from kindergarteners in economically disadvantaged school contexts in a large southwestern school district revealed that students exposed to a full day of instruction had greater literacy growth than their peers in half-day classrooms. Further examination of the program model results revealed that the relative efficacy of full-day kindergarten tended to be greater in smaller class size environments. These results, if replicated, suggest that full-day kindergarten initiatives targeted toward students from disadvantaged backgrounds may be more successful when implemented in classrooms with relatively small student enrollments. Implications for instructional policy and practice are discussed.

Keywords: Full-day kindergarten; Class size; Socio-economic context; Multilevel growth models

A trend in early childhood education has been the adoption and implementation of full-day kindergarten programs (Kauerz, 2005; Shin, 2005; Walston & West, 2004). As of 2007, all 50 states had enacted policy allowing school districts to offer a full-day kindergarten program to students. However, a recent review of state support for full-day kindergarten revealed large between-state legislative differences in the funding allocated to school districts (Kauerz, 2005). At present, several southeastern states require and earmark adequate funding for the provisioning of full-day kindergarten while other states fund kindergarten programs at a level that does not fully offset the expense (e.g., staff, classroom space, materials) associated with lengthening the school day. In these states, some districts opt to allocate state funds and/or utilize federal Title I resources to make full-day kindergarten available to certain student populations. In other districts, full-day kindergarten is supported directly by parents who finance a full day of instruction by paying a tuition stipend. The patchwork of approaches used to provide access to full-day kindergarten has resulted in the majority of children in the United States now having the opportunity to attend kindergarten on a full-time basis. A recent estimate indicates that approximately 65% of the nation’s children attend full-day kindergarten (Shin, 2005) although the percentage tends to be higher for students in some population groups. On average, children who are

* Corresponding author. Tel.: +1 541 346 1401.
E-mail address: kzvoch@uoregon.edu (K. Zvoch).
African-American, impoverished, or who attend a private school have greater access to full-day kindergarten when considered relative to their counterparts in other demographic or sector groups (Walston & West, 2004).

The increasing availability of full-day kindergarten has coincided with the rising awareness of K-12 stakeholders (e.g., school administrators, parents, state legislatures) to heightened school accountability pressures, changing household compositions, and mounting evidence suggesting that early intensive instruction can positively benefit students, families, and schools (Brewster & Railsback, 2002; Gullo, 1990; Lee, Burkam, Ready, Honigman, & Meisels, 2006). Proponents of full-day kindergarten argue that a lengthened school day can serve to reduce initial achievement gaps between advantaged and disadvantaged groups (Ackerman, Barnett, & Robin, 2005; Villegas, 2005), better prepare children for the increasing rigor of the elementary school curriculum (Elicker & Mathur, 1997), and ease the child care burden on families (Gullo, 1990; Vecchiotti, 2003). The enthusiasm that K-12 stakeholders express for full-day kindergarten may also stem in part from the comprehensible theory of action that underlies the full-day kindergarten model. The central mechanism of change in the full-day kindergarten environment is the greater instructional window (generally 6 h as opposed to 3 h of daily instruction) that allows teachers more time to use diverse pedagogical approaches (e.g., teacher-directed whole and small group instruction, child-initiated activity) and to present a more comprehensive and challenging curriculum. When effectively utilized, the extended instructional opportunity enables students to thoroughly engage in the varied learning activities (e.g., child-selected projects, teacher-directed group work, exploratory free play) that are thought to facilitate the acquisition of social/behavioral skills as well as a wider range of academic competencies (Elicker, 2000).

Study of processes and outcomes associated with full-day kindergarten tends to support the theory of action that underlies the full-day kindergarten program model. Teachers in full-day classrooms allocate more time to different modes of instruction and provide students with a more diverse curriculum (Elicker & Mathur, 1997; Walston & West, 2004). Students in full-day kindergarten are more likely than students in half-day kindergarten to spend additional time in teacher-directed and child-initiated activities and full-day kindergarteners are also more likely to receive daily instruction in math, science, and social studies. Relative to half-day kindergarten classrooms, more instructional time is also spent on the development of advanced literacy and mathematics skills. For example, in a nationally representative sample of kindergarteners, students in full-day classrooms were more likely to engage in learning activities (e.g., reading aloud, writing letters and numbers, solving math problems) that are generally not offered until students reach the first grade (Walston & West, 2004). Yet, it should be noted that when represented proportionally, the time allocated to some key learning activities (e.g., small group instruction, individual work) in full- and half-day classrooms was quite similar (see Elicker & Mathur, 1997; Walston & West, 2004).

Examination of student outcomes suggests that full-day kindergarteners tend to outperform their half-day peers on a variety of achievement and social indicators. In recent studies and reviews, the weight of the evidence suggests that children exposed to a complete day of instruction achieve greater academic gains and/or attain higher test scores than comparable children who have received a half day of instruction during the first year of formal schooling (Elicker, 2000; Fusaro, 1997; Guarino, Hamilton, Lockwood, & Rathbun, 2006; Lee et al., 2006; Walston & West, 2004). Moreover, some studies indicate that full-day kindergarteners also are more likely to have better attendance, experience less grade retention, and attain greater social adjustment than half-day kindergarteners (Cryan, Sheehan, Wiechel, & Bandy-Hadden, 1992; Gullo, 2000). However, there is also some indication that students who experience a full day of kindergarten are more likely than their half-day counterparts to later have less self-control, poorer interpersonal skills, and more internal and external behavioral issues (Vi-Nhuan, Kirby, Barney, Setodji, & Gershwin, 2006).

Although the evidence to date suggests that students who participate in full-day kindergarten generally benefit from the lengthened period of instruction, concerns have been raised regarding the methodological procedures used to evaluate the full-day kindergarten model (Ackerman et al., 2005; Cannon, Jacknowitz, & Painter, 2006; Puleo, 1988). The validity of research designs is one such concern as full-day kindergarten “effects” have generally been examined using pre-post or post-test only measurement designs with non-representative student samples and few controls for the non-random assignment of students to kindergarten program conditions (Elicker, 2000; Lee et al., 2006). A closely related issue is the means by which data from kindergarten program model studies have been analyzed. Although based in school settings, multilevel analytic techniques that better account for the hierarchical nature of school data structures (by allowing the separation and modeling of variance occurring at different levels of the data hierarchy) have frequently not been used. Instead, single-level methodologies (e.g., ANCOVA, OLS regression) which require data to be aggregated to the class (or school) level or disaggregated to the student level have been employed (Lee et al., 2006).
The methodological and analytic limitations identified in prior kindergarten model investigations suggest that additional study of full- and half-day kindergarten with diverse samples, better controls, repeated measures, and the use of appropriate analytic techniques is requisite to further elucidate the manner and degree to which a complete day of instruction differentially benefits kindergarten students. Given this backdrop, the current study addresses general efficacy issues related to full-day kindergarten program adoption by examining data obtained on children attending full- and half-day kindergarten programs in a large southwestern school district. The study is both a replication and an extension of previous work in the area. The study is a replication in that a new data set is utilized to investigate the relative efficacy of full-day kindergarten. However, the study also builds on prior investigations by employing a relatively well-controlled time series design and multilevel modeling techniques to assess the literacy status and literacy acquisition rate of students exposed to a full or half day of kindergarten instruction. The following research questions were investigated: (1) Do children who participate in full-day kindergarten acquire literacy skills at a faster rate than comparable children who attend a traditional half-day kindergarten program? (2) Do traditionally disadvantaged student populations (e.g., English language learners) differentially benefit from participation in full-day kindergarten? And (3) Do classroom characteristics such as class size and class calendar moderate the relationship between kindergarten program model and student literacy acquisition rates?

1. Method

1.1. Data source

Student literacy and kindergarten program model data were obtained from a large, rapidly growing school district. The district currently has over 300 schools and serves close to 300,000 students annually. At the elementary school level, the district has 188 schools that provide kindergarten instruction. Beginning in the 2004–2005 school year, full-day kindergarten was offered to children in the most economically disadvantaged elementary schools using Title I funding.

1.2. Analytic sample

As a component of a larger evaluation of the adoption and implementation of full-day kindergarten, the full-/half-day program model evaluation reported here was designed by the district under study to allow for an examination of kindergarten literacy outcomes in relatively similar economically disadvantaged school contexts. Full-day kindergarten students in six schools receiving Title I funding (i.e., >61% of students eligible for free or reduced priced lunch, FRL) constituted the treatment condition. The comparison condition was comprised of half-day kindergarten students in six schools that fell just under the qualifying poverty level for Title I funds (i.e., 55–61% FRL eligible). To ensure that students were exposed to similar instructional content, all students regardless of kindergarten program model received the same core instructional curriculum, Harcourt Trophies (Harcourt School Publishers, 2005).

As the primary intent of the current study was to identify the kindergarten program model “effect” on students who received a year of the instructional treatment, the analytic sample was restricted to those students who remained in the same classroom for the entire school year. The selection of non-transient students ensured that students were exposed to the same teacher and same kindergarten instructional program throughout the school year. However, the selection of non-transient students reduced the percentage of full-day kindergartners and the percentage of students in some special populations. Table 1 presents the characteristics of the analytic sample and the initial cohort of students from which the sample was drawn. In Table 1, it can be seen that the analytic sample was somewhat less disadvantaged than the initial student cohort taken as a whole. Relative to the enrollment characteristics of students at the beginning of the kindergarten year, the percentage of students participating in full-day kindergarten and the percentage of non-Asian ethnic minority and special education students was reduced by approximately 4, 2, and 1%, respectively.

Chi-square tests comparing the analytic sample (N = 443) with the group of students lost from the initial student cohort (N = 375) on student demographic characteristics revealed that students in special populations were not statistically over-represented in the group of excluded students (p > 0.05 in each comparison). However, a greater than expected number of full-day kindergartners were excluded from the analytic sample (χ²(1) = 5.98, p < 0.05). The loss of a higher percentage of the more economically disadvantaged full-day kindergartners indicates that analytic sample was not completely representative of the initial student cohort. In all, out-of-district transfers and the move-
Table 1
Analytic sample and the initial student cohort characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Analytic sample (N=443)</th>
<th>Initial cohort (N=818)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Percent</td>
</tr>
<tr>
<td>Full-day kindergarten</td>
<td>228</td>
<td>51.5</td>
</tr>
<tr>
<td>Female</td>
<td>220</td>
<td>49.7</td>
</tr>
<tr>
<td>Non-Asian minority</td>
<td>284</td>
<td>64.1</td>
</tr>
<tr>
<td>English language learner</td>
<td>143</td>
<td>32.3</td>
</tr>
<tr>
<td>Special education</td>
<td>19</td>
<td>4.3</td>
</tr>
</tbody>
</table>

The management of students to different (non-study) schools within the district resulted in 46% of the initial cohort of students being excluded from the analysis. Limitations associated with the relatively high transiency rate will be addressed in Section 3.

Table 2 presents the characteristics of the analytic sample by kindergarten program model. A series of chi-square and independent t-tests revealed that the full- and half-day treatment conditions differed in terms of student age (t(441) = 2.00, p < 0.05), number of absences (t(441) = 7.18, p < 0.001), and the percentage of English language learners (χ²(1) = 6.36, p < 0.05). On average, students in the full-day kindergarten condition were older and had more absences than their half-day peers. The full-day treatment condition also contained a relatively higher percentage of English language learners. The observed age and language minority differences suggest that the full- and half-day treatment conditions were not completely equivalent at the outset of the study (i.e., half-day students were somewhat more advantaged than their full-day peers). In the following, relatively complex multivariate models are utilized as a means for disentangling and controlling competing sources of student academic performance. The use of statistical modeling techniques in combination with the aforementioned research design controls aid in strengthening the inference regarding the relationship between program model and literacy outcome. However, the quasi-experimental nature of the present study prevents strong conclusions about kindergarten program “effects” from being drawn.

1.3. Individual measures

Student literacy was assessed using the Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002). DIBELS is a standardized, individually administered measure of early literacy development that is appropriate for students (including English language learners and most students with special needs) for whom learning
to read in English is a goal (Kaminski & Cummings, 2007). DIBELS is comprised of four subtests, Initial Sounds Fluency (ISF), Letter Name Fluency (LNF), Phonemic Segmentation Fluency (PSF), and Nonsense Word Fluency (NWF) that together measure the pre-reading and early reading skill development of students (Good & Kaminski, 2002). Student literacy performance is gauged by the number of correct responses given to an item battery (e.g., a list of letters or spoken three to four phoneme words) during a 7-min assessment frame. Students who perform with speed and precision earn higher scores. Scores can theoretically range from zero to positive infinity. The test developer reports a 1-month alternative form reliability estimate of 0.72 on the ISF, 0.88 on the LNF, 0.79 on the PSF, and 0.83 on the NWF (Good, Kaminski, Simmons, & Kame’enui, 2001).

In the present study, DIBELS was administered three times to students attending traditional 9-month schools (October, January, May). For students attending year-round schools, DIBELS was administered four times (October, January, May, July). In both school types and at each assessment frame, DIBELS assessments were administered over a 2-week period. Correlations across the DIBELS subtests ranged from a low of 0.58 between ISF and NWF to a high of 0.78 between LNF and NWF. For purposes of investigating relationships between kindergarten program model and student literacy outcomes, DIBELS subtest scores were summed to derive a composite literacy score.

The literacy composite provides a general measure of children’s fluency with essential pre-reading decoding skills. Examination of the composite score distributions revealed that scores were approximately normally distributed (i.e., skew and kurtosis values <1) on three of four test administrations. On the first DIBELS administration, skew and kurtosis values greater than one indicated the initial score distribution was positively skewed. An examination of relationships between predictor variables and a logarithmically transformed composite score distribution revealed that the pattern of relationship was consistent across the original and transformed scores. As a consequence, scores were left in the original metric to facilitate estimation and interpretation of student growth trajectories. Across the four administrations, composite DIBELS scores ranged from 0 (minimal pre-reading skill development) to 333 (extensive pre-reading skill development) with seventy scores of 0 and one score of 333 registered. Eighty percent ($N = 56$) of the zero scores occurred on the first test administration. Language and non-Asian ethnic minority students were more likely than their counterparts to register a zero ($p < 0.05$) on the first assessment, suggesting that kindergarten literacy acquisition rates may have been slightly underestimated for these student groups. Composite literacy score correlations were 0.82 between administration one and two, 0.84 between administration two and three, and 0.90 between administration three and four.

Outcome-relevant student background characteristics were collected from district records. Indicator codes were used to classify students on several characteristics that differed across program conditions and have been predictive of kindergarten literacy outcomes in other reported research (Lee et al., 2006; Rathbun & West, 2004; Walston & West, 2004). A series of dummy coded variables identified female, special education, and language and non-Asian minority students. Similarities identified in preliminary analyses regarding the initial literacy status and growth of Asian students and non-Hispanic Whites and similarities in the literacy performance of non-Asian minority students (i.e., African-American, Hispanic, and Native Americans) served as a basis for grouping students into two larger analytic categories. All non-Asian and language minority students present during the DIBELS administrations were assessed in English and included in the study. Following DIBELS protocol, standard administration procedures were maintained when assessing non-native English speakers and students with special needs (see Kaminski & Cummings, 2007).

Other individual measures were kindergarten students’ number of absences and students’ age in months at entry to kindergarten. Age at entry was included as a covariate given the treatment condition age differences and literature suggesting that older children have a slight academic advantage both at the beginning of kindergarten and over the first few years of elementary school (Stipek, 2002). Student absences were measured and utilized to account for kindergarten program model differences in treatment receipt, the extent to which a child was exposed to the 180-day instructional intervention (see Shadish, Cook, & Campbell, 2002). Table 2 presents descriptive statistics by program condition for each of the individual predictor variables.

1.4. Classroom measures

As with outcome-relevant student background differences, the classroom environments that students experience can also differ on characteristics that are predictive of student achievement outcomes. In the present study, two classroom characteristics (class size and class calendar) that can be expected to inhibit or enhance instruction were also collected from district data records (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996; Finn, Gerber, Achilles, & Boyd-
Zaharias, 2001; Grissmer, 1999; Nye, Hedges, & Konstantopoulos, 1999). Class size and class calendar information were derived from district records. Analogous to the coding of the program model variable (0 = half-day kindergarten, 1 = full-day kindergarten), dummy codes were used to distinguish classrooms in schools on 9-month or 12-month schedules (0 = 9 months, 1 = 12 months/year round). The bottom panel of Table 2 presents descriptive statistics on each classroom measure by program model. None of the treatment condition classroom differences were statistically significant (\( p > 0.05 \)). Nonetheless, each of the classroom characteristics was tested with the kindergarten program variable in the models specified below to allow for estimation of complex multivariate relationships.

1.5. Analytic procedures

Three-level longitudinal growth models (observations nested within students, students nested within classrooms) were used to describe the initial status and change in children’s pre-reading literacy skills. Using the Hierarchical Linear Modeling (HLM) program, version 6.0 (Raudenbush, Bryk, Cheong, & Congdon, 2004), a model with no covariates (i.e., the unconditional model) was fit first to estimate variance components. Then conditional models with child and classroom characteristics were fit and the unconditional and conditional variance components were compared to provide an estimate of variance accounted for at each level of the model. The models account for the correlations in the data due to repeated assessments of the same children and the nesting of children within classrooms (Raudenbush & Bryk, 2002). The final conditional model provides an estimate of the relationship between the type of kindergarten program children attended (full- or half-day) and children’s literacy outcomes taking into account children’s gender, ethnicity, age, English language and special education status, and the number of school year absences.

Differences in class schedules, assessment time frames, and the number of DIBELS assessments presented a challenge to model specification. As classrooms housed in year-round schools operated on five different tracks (i.e., calendars), as upward to a 2-week period of time could separate a DIBELS administration for students attending classrooms in any assessment window, and as students received a different number of DIBELS assessments (3 or 4) depending on class schedule (9- vs. 12-month school year), a decision was made to model student literacy acquisition rates as a function of the number of elapsed instructional days experienced by students. The number of elapsed instructional days for each student at each assessment point was calculated by matching assessment dates with the relevant school attendance calendar. Taking this approach, the confounding effect of differential instructional exposure (at particular assessment points) was explicitly included in the model. However, aside from instructional exposure, it should be noted that students taught in year-round schools still had a longer developmental window in which to acquire literacy skills (9 vs. 12 months). To account for the increased potential for literacy development in students attending kindergarten in year-round schools, the dichotomous classroom calendar indicator (i.e., 9- vs. 12-month class schedule) described above was also included as a classroom-level predictor.

A three-level longitudinal growth model was used to fit a linear regression function to the time series data obtained on students. Eq. (1) specifies the level-1 model. In Eq. (1), it can be seen that the outcome, \( \pi_{ij} \) is the literacy outcome at time \( t \) for student \( i \) in classroom \( j \), \( \pi_{0ij} \) is the status of student \( ij \) at the time of the first assessment (i.e., the beginning of kindergarten), \( \pi_{1ij} \) is the linear rate of literacy acquisition across the kindergarten year for student \( ij \), and \( \pi_{ij} \) is a residual term representing unexplained variation from the latent growth trajectory. Residuals were assumed normally distributed with mean of zero and variance \( \sigma^2 \).

\[ Y_{ij} = \pi_{0ij} + \pi_{1ij} (\text{elapsed instructional days}) + \pi_{ij} \]  

(1)

In the unconditional model, within-classroom variation (level-2) in the initial status (\( \pi_{0ij} \)) and literacy acquisition rates (\( \pi_{1ij} \)) of students was modeled in terms of the status and acquisition parameters of the student’s classroom and student-level residuals while between-classroom variation (level-3) in the initial status (\( \beta_{0ij} \)) and acquisition rate (\( \beta_{1ij} \)) of classrooms was modeled as a function of grand mean literacy status or grand mean literacy acquisition and classroom-level residuals (see Eqs. (2a)–(3b)).

\[ \pi_{0ij} = \beta_{00j} + r_{0ij} \]  

(2a)

\[ \pi_{1ij} = \beta_{10j} + r_{1ij} \]  

(2b)

\[ \beta_{00j} = \gamma_{00j} + u_{00j} \]  

(3a)

\[ \beta_{10j} = \gamma_{10j} + u_{10j} \]  

(3b)
In the conditional models, student and classroom characteristics were added to the equation. In Eqs. (4a) and (4b), it can be seen that within-classroom variation (level-2) in the literacy status and literacy acquisition of students was modeled as a function of the status ($\beta_{0ij}$) or acquisition rate ($\beta_{1ij}$) of classroom $j$, the student characteristics that were hypothesized to account for observed variation in the parameters of the within-classroom literacy model, and respective residual terms, $r_{0ij}$ or $r_{1ij}$. Note that in the specification of the model for student growth rates, the number of student absences was also added as a predictor.

$$
\pi_{0ij} = \beta_{00j} + \beta_{01j}(\text{female}) + \beta_{02j}(\text{non-Asian minority}) + \beta_{03j}(\text{ELL}) + \beta_{04j}(\text{SPED}) + \beta_{05j}(\text{age in months}) + r_{0ij}
$$

(4a)

$$
\pi_{1ij} = \beta_{10j} + \beta_{11j}(\text{female}) + \beta_{12j}(\text{non-Asian minority}) + \beta_{13j}(\text{ELL}) + \beta_{14j}(\text{SPED}) + \beta_{15j}(\text{age in months}) + \beta_{16j}(\text{number of absences}) + r_{1ij}
$$

(4b)

To assess whether the slopes relating individual background characteristics to student literacy status and acquisition rate varied across classrooms, a multi-parameter hypothesis test of the variance components associated with the within-classroom model was conducted. The test was run to determine if there was a basis for investigation of research question 2 (i.e., did students in special populations differentially benefit from participation in full-day kindergarten?). Comparison of a model where all predictors were group mean centered and slopes were free to vary (deviance = 13894.96, parameters estimated = 108) with a model where all slopes were held constant across classrooms (deviance = 13953.20, parameters estimated = 20) indicated that the relaxation of parameter constraints did not statistically improve the fit of the model ($\chi^2(88) = 58.24$, $p > 0.05$). This result indicates that estimated literacy status and acquisition differences between student demographic groups were relatively similar (i.e., within-sampling error) across classroom contexts. As there was no evidence of differential student subgroup advantage or disadvantage in certain classrooms, each student predictor was treated as a fixed effect and grand mean centered in order to adjust the classroom literacy status and acquisition estimates for the nonrandom allocation of students to classrooms. The grand mean centering of student predictor variables allowed for a test of the relationship between program and classroom characteristics and student literacy outcomes at the mean or mean ratio (for dichotomous variables) of each student-level predictor.

Eqs. (5a) and (5b) present the final conditional between-classroom model (level-3) where the initial literacy status and acquisition rate of classrooms is conceived as a function of grand mean literacy status ($\gamma_{000}$) or grand mean literacy acquisition rate ($\gamma_{100}$), classroom characteristics, and respective classroom-level residual terms, $u_{00j}$ or $u_{10j}$. Note in the specification of the model for classroom literacy acquisition rates, a term representing the program model (i.e., full-/half-day kindergarten) by class size interaction was also added as a predictor.

$$
\beta_{00j} = \gamma_{000} + \gamma_{001}(\text{full-day kindergarten}) + \gamma_{002}(\text{class size}) + \gamma_{003}(\text{year-round schedule}) + u_{00j}
$$

(5a)

$$
\beta_{10j} = \gamma_{100} + \gamma_{101}(\text{full-day kindergarten}) + \gamma_{102}(\text{class size}) + \gamma_{103}(\text{year-round schedule}) + \gamma_{104}(\text{program model} \times \text{class size}) + u_{10j}
$$

(5b)

The program model-by-class size interaction term was sequentially added to the model after testing for the main effects of the classroom-level predictors (i.e., program model: full-/half-day kindergarten, class size; class calendar: 9-/12-month schedule). Estimating the main effects in one block allowed for a test of research question 1, do children who participate in full-day kindergarten acquire literacy skills at a faster rate than comparable children who attend a traditional half-day kindergarten program? Adding the interaction term in a second block allowed for a test of research question 3, does class size moderate the relationship between kindergarten program model and the rate of literacy acquisition? Detail on the sequential testing of other classroom-level interaction terms is presented in the results.

2. Results

Three separate headings are used to identify and separate the results associated with the unconditional and conditional analytic models described above. Unconditional model results include estimates of the mean literacy status of all students upon entry to kindergarten as well as the mean rate of literacy acquisition of all students across the kindergarten year. Unconditional estimates of the amount of within- and between-classroom variation are also presented to highlight
the degree to which the literacy status and literacy acquisition rates of students vary within classrooms and the degree to which mean literacy status and mean literacy acquisition rates vary between classrooms. Next, the conditional within-classroom model results present the relative difference in initial literacy status and literacy acquisition rate by student background characteristic. These estimates reflect the gap in literacy level between identified student groups upon entry to kindergarten and provide indication of the degree to which literacy acquisition rates between special student populations diverge over the kindergarten year. Finally, the between-classroom results are estimates of the relationship between classroom characteristics (i.e., kindergarten program model: full-/half-day, class size; class calendar: 9 month/year round and their joint relationship) and student literacy at the beginning of kindergarten (literacy status) and over the course of the kindergarten year (literacy acquisition). Variance estimates associated with the conditional models reflect the degree to which within and between-classroom differences in literacy status and acquisition rate are reduced (or explained) by the inclusion of student and classroom predictor variables.

2.1. Unconditional three-level student literacy model

Table 3 presents the results of the three-level student literacy models. The two fixed effects presented in the top panel of column 1 are estimates of the mean literacy score upon entry to kindergarten ($\gamma_{000} = 6.43$) and the mean rate of literacy acquisition per instructional day ($\gamma_{100} = 0.63$) across all students. Variance estimates are presented in the middle of Table 3. Chi-square tests indicated that there was statistically significant variation across all parameters. These results suggest that there were differences from one student and one classroom to another in initial literacy status as well as in the rate of literacy acquisition throughout the kindergarten year. Calculation of the percentage of variation attributable to classrooms indicated that 10% of the variability in entry status (55.82/55.82 + 526.54) and 50% of the variability in literacy acquisition rates (0.05/0.05 + 0.05) was due to classroom-to-classroom differences.

2.2. Within-classroom student literacy model

Results of the within-classroom model are presented in the second column of Table 3. In column 2, it can be seen that upon entry to kindergarten, non-Asian minority students and English language learners had lower literacy scores relative to their counterparts. Estimates of the relationship between student characteristics and literacy acquisition rates are presented next. These estimates demonstrate that males acquired literacy skills at a slower rate than females and non-Asian minority, English language learners, and special education students acquired literacy skills at a slower rate than their counterparts. A negative relationship between student absences and literacy acquisition rates was also observed.

The within-classroom results indicate that the initial gap in literacy performance between language groups (11.23 points, English proficient advantage) increased by an additional 18 points ($0.10 \times 180 = 18.00$, a final status difference of 0.55 of a standard deviation) and the initial gap in literacy performance between ethnic groups (11.83 points, non-Hispanic White/Asian advantage) increased by approximately 13 points ($0.07 \times 180 = 12.60$, a final status difference of 0.46 of a standard deviation) over the kindergarten year. In addition, a final status literacy gap between education groups and between gender groups also emerged. By acquiring literacy skills at a rate of 0.18 points/instructional day less than their general education peers, special education students scored approximately two-thirds of a standard deviation lower than their counterparts by the end of the kindergarten year. A 0.07 point per instructional day literacy acquisition deficit also placed male students two-tenths of a standard deviation behind their female counterparts at years end. Final status literacy gaps were calculated by multiplying the relevant growth estimate by 180, the number of days in the school year and adding the computed value to the estimate that captured group differences in initial literacy performance. For example, the final literacy gap between language groups was $(0.10 \times 180) + 11.23 = 29.23$. Final literacy gap estimates were then divided by the pooled standard deviation (S.D.) of estimated literacy scores at the end of the school year (S.D. = 52.73) to derive an effect size estimate of the final status difference between student groups (e.g., language group standardized difference, 29.23/52.73 = 0.55).

In the middle panel of column 2, conditional variance estimates of students’ literacy status and rate of literacy acquisition are presented. Chi-square tests demonstrated that student characteristics did not completely explain the differences among students in literacy outcomes. A comparison of unconditional and conditional variance estimates revealed that student characteristics accounted for a relatively small amount of the variation in literacy levels and literacy acquisition. Student characteristics accounted for 19% of the variation in students’ initial literacy status and 14% of the
Table 3
Three-level student literacy models

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unconditional (1)</td>
</tr>
<tr>
<td>Mean literacy status ($\gamma_{000}$)</td>
<td>6.43 (2.22)*</td>
</tr>
<tr>
<td>Full-day kindergarten ($\gamma_{001}$)</td>
<td></td>
</tr>
<tr>
<td>Class size ($\gamma_{002}$)</td>
<td>0.68 (0.40)</td>
</tr>
<tr>
<td>Year-round schedule ($\gamma_{003}$)</td>
<td></td>
</tr>
<tr>
<td>Female ($\gamma_{010}$)</td>
<td>−2.02 (2.76)</td>
</tr>
<tr>
<td>Non-Asian minority ($\gamma_{020}$)</td>
<td>−11.87 (3.26)**</td>
</tr>
<tr>
<td>ELL ($\gamma_{030}$)</td>
<td>−11.23 (3.44)**</td>
</tr>
<tr>
<td>Age at entry ($\gamma_{040}$)</td>
<td>0.41 (0.30)</td>
</tr>
<tr>
<td>Special education ($\gamma_{050}$)</td>
<td>−0.61 (6.89)</td>
</tr>
<tr>
<td>Mean literacy growth ($\gamma_{100}$)</td>
<td>0.63 (0.05)**</td>
</tr>
<tr>
<td>Full-day kindergarten ($\gamma_{101}$)</td>
<td></td>
</tr>
<tr>
<td>Class size ($\gamma_{102}$)</td>
<td>0.01 (0.01)</td>
</tr>
<tr>
<td>Year-round schedule ($\gamma_{103}$)</td>
<td></td>
</tr>
<tr>
<td>Program model × class size ($\gamma_{104}$)</td>
<td>−0.03 (0.01)***</td>
</tr>
<tr>
<td>Female ($\gamma_{110}$)</td>
<td>0.07 (0.03)*</td>
</tr>
<tr>
<td>Non-Asian minority ($\gamma_{120}$)</td>
<td>−0.07 (0.03)*</td>
</tr>
<tr>
<td>ELL ($\gamma_{130}$)</td>
<td>−0.10 (0.04)**</td>
</tr>
<tr>
<td>Age at entry ($\gamma_{140}$)</td>
<td>−0.001 (0.004)</td>
</tr>
<tr>
<td>Special education ($\gamma_{150}$)</td>
<td>−0.18 (0.07)*</td>
</tr>
<tr>
<td>Absences ($\gamma_{160}$)</td>
<td>−0.006 (0.002)**</td>
</tr>
</tbody>
</table>

Random effects                  Variance estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual literacy status ($r_{ij}$)</td>
<td>526.54***</td>
<td>429.29***</td>
<td>428.86***</td>
</tr>
<tr>
<td>Individual literacy growth ($r_{ij}$)</td>
<td>0.05***</td>
<td>0.04***</td>
<td>0.04***</td>
</tr>
<tr>
<td>Level-1 error ($e_{ij}$)</td>
<td>330.34</td>
<td>330.57</td>
<td>331.00</td>
</tr>
<tr>
<td>Classroom literacy status ($u_{0j}$)</td>
<td>55.82**</td>
<td>57.03***</td>
<td>24.25**</td>
</tr>
<tr>
<td>Classroom literacy growth ($u_{1j}$)</td>
<td>0.05**</td>
<td>0.05***</td>
<td>0.01***</td>
</tr>
<tr>
<td>Model deviance</td>
<td>14045.80</td>
<td>13956.66</td>
<td>13920.32</td>
</tr>
<tr>
<td>Parameters estimated</td>
<td>9</td>
<td>20</td>
<td>27</td>
</tr>
</tbody>
</table>

Note: The program model-by-class size interaction term was sequentially added to the model after testing for the main effects of the classroom-level predictors. Coefficients from estimation of the complete model are presented; ELL, English language learner.

Variation in students’ rate of literacy acquisition. Model deviance estimates are the final statistics presented in Table 3. A comparison of deviance statistics revealed that the conditional within-classroom model (deviance = 13956.66) fit the data significantly better than the unconditional model (deviance = 14045.80, \(\chi^2\Delta(11) = 89.14, p < 0.001\)).

2.3. Between-classroom student literacy model

Estimates associated with the final conditional between-classroom model are presented in column 3. The coefficients associated with kindergarten program model indicate that students in full-day kindergarten classrooms had lower adjusted initial literacy status but greater rates of literacy acquisition relative to their counterparts in half-day classrooms. These results speak to research question 1 and indicate that the literacy gap favoring students in the relatively less economically disadvantaged half-day classrooms (55–61% FRL eligible) at entry to kindergarten was reversed in favor of students in the relatively more economically disadvantaged full-day classrooms (>61% FRL eligible) by the end of the school year. However, it is important to note that the greater literacy growth of students in full-day classrooms was not constant across class size conditions. Sequentially adding the program model-by-class
Fig. 1. Mean literacy acquisition per instructional day as a function of kindergarten program model and class size.

The coefficient associated with the program model-by-class size interaction speaks to research question 3 and indicates that full-day kindergarten was relatively more effective in below average and average size classrooms. However, in classrooms with larger student enrollments, the full-day kindergarten advantage diminished due to the lower rate of literacy acquisition in full-day classrooms and the unexpected increase in literacy acquisition in relatively large sized half-day classrooms. The joint relationship between kindergarten program type and class size on literacy acquisition is displayed in Fig. 1. Class size \((M = 22.55, \text{S.D.} = 4.75)\) is represented at enrollments of <20, 20–24, and >24 students.

To examine whether student literacy outcomes were also moderated by the joint relationship of other classroom characteristics, terms representing the interaction between program model and class calendar and the interaction between class size and class calendar on classroom literacy status and literacy acquisition were tested as well. Sequentially adding the program model-by-class size, program model-by-class calendar, and class calendar-by-class size interaction terms to the main effects model for literacy status and the program model-by-class calendar and class calendar-by-class size interaction terms for literacy acquisition did not result in a statistically significant improvement in model fit \((\chi^2/\Delta(1) < 2.00 \text{ for each interaction term, } p > 0.05)\). As a result, these terms were not retained in the final model.

At the bottom of Table 3, literacy status and literacy acquisition variance estimates conditioned on classroom (and student) characteristics are presented. Variance estimates associated with the final between-classroom model’s random effects were smaller in magnitude than the random effects associated with the within-classroom model. A comparison of unconditional and conditional variance estimates revealed that classroom characteristics accounted for a moderate amount of the adjusted variation in literacy levels and a large amount of the adjusted variation in literacy acquisition. Classroom characteristics accounted for 57% of the residual variation in the literacy status means and 88% of the residual variation in mean literacy acquisition. Model deviance estimates are the final statistics presented in column 3. Deviance estimates associated with the final between-classroom model were smaller than those associated with the within-classroom model. A chi-square test revealed that the change in deviance was statistically significant \((\chi^2/\Delta(7) = 36.34, p < 0.001)\), indicating that inclusion of classroom characteristics resulted in a significant improvement in model fit.
3. Discussion

The present study was designed to assess relationships between kindergarten program model (full- vs. half-day) and student literacy outcomes. The study was facilitated by the application of multilevel modeling techniques to data collected on students and classrooms in a large school district in the southwestern United States. Results indicated that kindergarten students exposed to a full day of instruction had a faster rate of literacy acquisition relative to their peers in half-day classrooms. However, the relative efficacy of full-day kindergarten varied with respect to classroom enrollment size. In relatively small-sized classrooms (<20 students), full-day kindergarteners rate of literacy acquisition was twice that of their half-day peers. The full-day advantage was maintained to a similar degree in moderately sized classrooms (20–24 students), but in relatively large classes (>24 students), acquisition rates were more similar as full-day students had relatively lower and half-day students had relatively higher rates of literacy acquisition.

The strong negative relationship between class size and literacy acquisition in full-day classrooms \( r = -0.67 \) and the slight positive relationship between class size and literacy acquisition in half-day classrooms \( r = 0.12 \) combined to produce the statistically significant joint relationship between program model, class size, and literacy acquisition rates. Considered by program model condition, these results indicate that the relative rate of literacy acquisition systematically diminished in relation to class size in full-day kindergarten classrooms while students in half-day classrooms acquired literacy skills at rate that was relatively constant across most classroom enrollments. One exception to the within-condition pattern of results was the relatively faster rate of literacy acquisition in large enrollment half-day classrooms. Considered together, these results indicate that full-day kindergarten was relatively less effective (in terms of literacy acquisition rates) with respect to class size both within- and between-program model conditions. In particular, full-day kindergarteners in large enrollment classrooms acquired literacy skills at a slower rate relative to their counterparts in smaller-sized full-day classrooms and at a rate more similar to their peers in large-sized half-day classrooms. Assuming that the observed within- and between-condition differences in the relative efficacy of full-day kindergarten are attributable to something unique about the instructional challenge associated with large student enrollments in full-day kindergarten environments, these results suggest that K-12 stakeholders may need to more closely consider whether lengthening the school day in isolation from class size adjustments will be sufficient to achieve the outcome desired (for students from disadvantaged backgrounds) from the investment in full-day kindergarten.

Aside from the moderating influence of class size, the relationship between kindergarten program model and student literacy acquisition observed in the current study was largely consistent with other recent findings on the relative achievement performance of full- and half-day kindergarten students (Guarino et al., 2006; Lee et al., 2006; Walston & West, 2004; Yan & Lin, 2005). As with other recent studies, students in full-day classrooms outperformed comparable students in half-day classrooms on a key achievement indicator. The current study was also somewhat unique, however, in demonstrating that differential literacy acquisition by kindergarten program model was achieved during the first year of program implementation. Analysis of data generated in conjunction with the full-day program adoption indicated that the policies, procedures, and routines requisite for successful student outcomes to be observed can be quickly developed and implemented. A second unique component was the manner in which program model performance was evaluated. In the current study, students were measured on multiple occasions during the kindergarten year. The multiple measurements on students permitted the change in individual literacy status to be represented as a growth trajectory instead of a pre-post gain score. The estimation of growth trajectories thus facilitated a relatively more accurate representation of inter-individual differences in literacy status and literacy acquisition and permitted examination of relationships between the hypothesized predictors of these distinct literacy outcomes.

The examination of student growth trajectories revealed relatively large status differences between students in special population groups and their counterparts upon entry to kindergarten. In particular, non-Asian and language minority students had literacy scores approximately one-half of a standard deviation behind their respective non-Hispanic White/Asian and English proficient peers. Moreover, differential rates of literacy acquisition over the kindergarten year widened the initial gap in pre-reading literacy between these student groups and also led to the rise of end status literacy gaps between males and females and general and special education students (female and general education student advantage). At present, it is difficult to pinpoint why the growth rates of students in special populations diverged from those of their counterparts. Aside from possible measurement artifacts, it may be that a host of unmeasured factors that can be expected to impact the literacy growth rates of individual students (e.g., teacher/parent expectations, peer...
effects, familial educational resources) may be somewhat less positive or prevalent for traditionally disadvantaged students in disadvantaged contexts. The potential lack of access to certain undefined familial/school resources may be particularly problematic for students in at-risk groups in economically challenged environments as compensating community resources (e.g., positive role models, supplemental educational programs) may also be less available to these students (Coley, 2002; Lee & Burkam, 2002).

As results from the current study contribute to the growing body of literature on early childhood literacy and the efficacy of full-day kindergarten, consideration of relevant sample and data limitations is necessary to contextualize the current findings. Specifically, it should be noted that the study was based on student performance on one measure of early childhood literacy. Use of an alternative literacy measure more or less sensitive to the instructional practice of teachers or use of an indicator representing another content domain (e.g., mathematics) may have produced different estimates of the relative performance and growth of students in full- and half-day kindergarten programs. The current study may have also been limited to some degree as literacy data from students attending schools in a highly transient urban environment were analyzed. The transient scholastic environments studied are likely not representative of those students in more stable contexts. In addition, the disproportionate exclusion of students from the more economically disadvantaged full-day kindergarten classrooms may have produced an upward bias on the literacy estimates associated with this treatment condition (see Zvoch & Stevens, 2005). Results may also have been limited to some degree as the number of English language learners served by the district under study tends to be larger than is found in school districts in other geographic regions.

Concerns surrounding the validity of inferences regarding the relative efficacy of full-day kindergarten also follow from the lack of data on program processes and the lack of control over student assignment to kindergarten program conditions. As with many other investigations of full-day kindergarten, program implementation data were not collected and students were allocated to treatment conditions based on the results of selection factors outside the control of the researchers (see Lee et al., 2006). Not having data on instructional practice and classroom process makes it difficult to distinguish the specific ways in which full- and half-day classrooms differed. Further, the non-random assignment of students to conditions raises the potential for the background characteristics and experiences of students to have systematically biased the investigation of kindergarten program “effects.” In an attempt to minimize internal validity threats, research design and statistical control measures were employed. Holding the kindergarten curriculum (Harcourt Trophies) constant in all study classrooms and employing a matched comparison condition were important design controls. Other student (sex, age, attendance, ethnic, language, and education status) and classroom (class size, class calendar) differences were addressed by the estimation of multivariate statistical models. The use of these control measures aid in improving the strength of the inferences that can be drawn from the current study, but as additional uncontrolled student, family, and classroom factors (e.g., parental involvement, teacher quality) could have been associated with the relative rates of literacy acquisition for students in different kindergarten program conditions, causal conclusions about kindergarten program “effects” are not warranted.

Although the aforementioned limitations suggest a need for additional research on the performance of students in full-day kindergarten programs, both with more comprehensive design controls and across different outcome measures and sampling conditions, the current study does provide some evidence on the potential efficacy of adopting and implementing a full-day kindergarten program. In the present study, disadvantaged students in full-day classrooms had a rate of literacy acquisition that outpaced that of demographically similar peers in half-day classrooms during the first year of a program implementation. However, the efficacy of the full-day kindergarten intervention was contingent upon class size as students in smaller full-day classes had faster literacy acquisition rates than students in larger full-day class size environments. Findings such as these are useful for state legislatures that must consider the manner and degree to which limited resources should be allocated to early childhood education initiatives. In the case of kindergarten programs, the evidence is beginning to suggest that increased access to full-day kindergarten can result in an initial and relatively immediate payoff in greater student literacy, the foundation for future academic and social success. However, factors (e.g., class size, classroom aides) other than the increased instructional time found in full-day kindergarten classrooms also appear to have relevance for children in certain demographic groups or classroom contexts (Lee et al., 2006; Walston & West, 2004). Moreover, questions about the long-term academic and social benefits of full-day kindergarten have recently been raised (Cannon et al., 2006; Rathbun & West, 2004; Saam & Nowak, 2005; Vi-Nhuan et al., 2006). As a result, additional study of the contexts and practices that surround the implementation of full-day kindergarten programs is urged to further elucidate the conditions under which and the students for whom full-day kindergarten is most beneficial.
References


