Published in final edited form as:

Elem Sch J. 2007 March ; 107(4): 327–340. doi:10.1086/516667.

The Short-Term Effect of Grade Retention on Peer Relations and Academic Performance of At-Risk First Graders

Katie A. Gleason, Oi-man Kwok, and Jan N. Hughes

Texas A & M University

Abstract

Using latent variable structural equation modeling, we tested a theoretical model positing that grade retention has a positive effect on children's teacher- and peer-rated academic competencies and on sociometric measures of peer acceptance. We also expected that the positive effect of grade retention on peer acceptance would be mediated by children's ability to meet academic challenges in their classrooms. Participants were 350 (52.6% male) ethnically diverse and academically at-risk first graders attending 1 of 3 school districts in Texas. An individually administered test of academic achievement, teacher-report and peer-report measures of academic competence, and peer-report measures of peer acceptance were collected on children in first grade and 1 year later, at which time 63 children were repeating first grade and 287 were in second grade. The hypothesized model provided a good fit to the data. Children's academic competencies, as perceived by peers and teachers, fully mediated the effect of retention on subsequent peer acceptance.

When a child does not exhibit certain minimum academic competencies, teachers and parents face a dilemma. The child can be promoted to the next grade, with the hope that the child will somehow acquire the necessary competencies. This option is known as "social promotion." An alternative is to retain the child in grade to give the child another opportunity to master academic skills for that grade. This option is known as "grade retention." Passage in 2002 of the No Child Left Behind Act, with its emphasis on mastery of minimum grade-level competencies as a condition for promotion, has renewed discussion of grade retention in the public arena. However, many researchers have examined the effects of grade retention during childhood and adolescence (Alexander, Entwisle, & Dauber, 1994, 2003; Beebe-Frankenberger, Bocian, MacMillan, & Gresham, 2004; Holmes & Matthews, 1984; Jimerson, 1999, 2001; Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997; Mantzicopoulos, 1997; Mantzicopoulos & Morrison, 1992; McCoy & Reynolds, 1999; Meidel & Reynolds, 1999; Meisels & Liaw, 1993; Rumberger, 1995; Shepard & Smith, 1990).

Generally, research has suggested that grade retention has a negative effect on the developmental trajectories of children; however, there are several inconsistencies in the grade-retention literature. For example, Jimerson (2001) conducted a meta-analysis of 20 studies published between 1990 and 1999 that reported the effects of grade retention on academic and socioemotional development in retained students and a comparison group of promoted students. These 20 studies yielded 175 analyses reporting academic outcomes and a mean effect size (Cohen's *d*, based on averaging outcomes for each study and weighing each study equally) of $-.39$. However, effects across studies were varied. Only nine analyses suggested that retained students performed better academically than promoted students. Conversely, 82 analyses indicated that promoted students performed better than retained students, and 84 analyses

yielded no significant differences between retained and promoted students on academic achievement.

Grade Retention and Peer Relations

Although there is a large body of literature on the effects of grade retention on academic achievement and, to a lesser extent, on behavioral adjustment, research examining the relation between retention and peer relations is sparse. The few studies on retention and peer acceptance have relied on teacher or parent reports or child perceptions of peer support rather than on peer assessments of liking for the child. These studies have yielded inconsistent results. For example, when teachers were asked to rate classroom behavior and peer acceptance, they rated retained students as displaying more problem behaviors in the classroom and as being less well liked by classmates than low-achieving promoted students (Jimerson, 1999; Pianta, Teitbohl, & Bennett, 1997). However, when Pierson and Connell (1992) asked third through sixth graders to rate on a 1 (very true) to 4 (not at all true) Likert scale their feelings in situations in which they are with their classmates (e.g., I feel ... relaxed ... ignored ... happy), responses did not differ for retained children ($N = 74$) and matched-ability children who were never retained ($N = 69$). The authors concluded that the results suggested that retention is not harmful to a child's perceived relatedness to peers.

Using an experimental design and hypothetical vignettes, Plummer and Graziano (1987) investigated the effect of retention status (defined as children who are old for their grade) on second and fifth graders' preferences for peers as partners for school-related or non-school-related activities. The manipulation consisted of informing children that the child in the retained vignette was old for his or her grade. The researchers confirmed that this manipulation had the effect of communicating the vignette child's retention status. They found mixed results for the effect of the manipulation on children's peer preferences. Specifically, when children chose a partner for a non-school-related activity, they preferred the promoted (i.e., average age for grade) child. However, when children chose a partner for a school-related activity, children preferred the retained (i.e., old for grade) child. Importantly, when asked the reason for their choice for the school-related activity, children who chose the retained child stated that this child had more experience with classroom activities. Results also indicated that the fact that retained children were taller may have provided them some social advantage. Thus, Plummer and Graziano's findings suggested that retention's effects on a child's peer acceptance might depend on the interaction context.

Developmental Significance of Peer Relations

It is important to determine the effects of grade retention on peer relations because children who have problems developing and maintaining positive peer relations are more likely to develop adjustment problems (Parker & Asher, 1987; Rubin, Bukowski, & Parker, 1998). Research has documented associations between peer rejection and aggression (Coie, Dodge, & Kupersmidt, 1990). Children who have difficulties with peer relations are also more likely to withdraw from school, become involved in delinquent activities, and suffer from mental health problems than are children without such difficulties (Furrer & Skinner, 2003; Kupersmidt, Coie, & Dodge, 1990; Parker & Asher, 1987). For example, Buhs and Ladd (2001) found that rejected kindergarten children were more likely to experience negative peer treatment and that this treatment mediated the link between peer rejection and indexes of school adjustment. Specifically, rejected children showed decreases in classroom participation and were more likely to report loneliness, express a desire to avoid school, and perform less well on achievement measures. Furrer and Skinner (2003) found that third through sixth graders' emotional and behavioral engagement mediated the association between their social relatedness (a composite of peer, teacher, and parent support) and achievement. Furthermore,

peer support uniquely predicted children's engagement in the classroom, after controlling for the effect of support from parents and teachers on achievement. According to this view, peer acceptance affects future academic achievement via its direct effects on academic motivation and engagement.

Given the importance of peer acceptance to children's concurrent and future school adjustment, it is important to investigate the effect of grade retention on peer acceptance. To date, no study has examined the effect of grade retention in the early grades on children's peer-assessed liking. Several authors have speculated that negative effects associated with grade retention, such as behavioral problems and early school withdrawal, are the result, in part, of retained children experiencing peer stigmatization and discrimination (Alexander et al., 1994; Jimerson, Anderson, & Whipple, 2002). However, empirical support for these views is lacking. Indeed, Plummer and Graziano's findings (1987) suggested that retained children may gain in peer status due to their relatively greater physical size and knowledge of grade-level routines and tasks.

Grade Retention and Variables Associated with Peer Acceptance

Due to the dearth of studies investigating the relation of retention to peer acceptance, one might turn to studies of the effects of grade retention on variables associated with peer acceptance, such as academic achievement or socioemotional adjustment, to anticipate how retention might affect peer acceptance. In his meta-analysis, Jimerson (2001) also examined the effects of grade retention on socioemotional outcomes. The 16 studies reporting socioemotional outcomes yielded 148 analyses that examined these outcomes of retained students relative to a comparison group of promoted children. The average unweighted study effect size was $-.22$, favoring promoted children. Of the 148 analyses, 127 (86%) yielded no statistically significant difference between retained and comparison students, eight (5%) favored the retained sample, and 13 (9%) favored the comparison students. Thus, most studies failed to show any effect of retention on socioemotional adjustment.

Results of studies of retention's effects on achievement have also been mixed, though predominantly negative (for reviews see Holmes & Matthews, 1984; Jimerson, 2001; Shepard & Smith, 1990.). When interpreting results of such studies, it is important to distinguish studies that used same-age comparisons from those that used same-grade comparisons. Not surprisingly, following retention, repeaters typically perform worse on tests of achievement, relative to other low-achieving, same-age students who were promoted to the next grade (for a review see Alexander et al., 2003). That is, they show slower academic growth, relative to their same-age promoted peers. Conversely, when repeaters are compared to their grade peers who are younger and completing the grade for the first time, they gain in achievement. That is, they are not as far behind their (younger) classmates after retention as they had been before. Their low-achieving promoted peers do not show as much improvement in their standing relative to their (older) classmates (Alexander et al., 2003; Pierson & Connell, 1992). When peer acceptance is assessed in the context of one's classroom, repeaters may experience a boost in acceptance due to their improved academic standing relative to their classmates.

The purpose of the present study was to investigate the 1-year longitudinal effect of retention in first grade on peer acceptance. Based on Plummer and Graziano's (1987) results as well as on studies reporting that retained children, relative to promoted children, demonstrate greater ability to meet classroom academic expectations (Alexander et al., 1994; Pierson & Connell, 1992), we expected that retained children would experience greater peer acceptance in year 2, relative to their promoted peers, after controlling for year 1 peer acceptance. Furthermore, we predicted that retained children would be perceived by peers and teachers as exhibiting higher academic competence in year 2, relative to promoted children. Finally, we tested a model that

posits that retention's positive effect on peer acceptance is mediated by its positive effect on classroom academic competencies.

Consistent with developmental systems theory (Lerner, 1998), the well-established long-term effects of grade retention on school withdrawal (Alexander et al., 2003; Jimerson et al., 2002) are thought to be the result of multiple and interactive factors. However, research on the effects of grade retention has generally failed to investigate mediating processes by which retention alters children's transactions with their environment. This study, although limited in duration for detecting effects, attempted to identify mechanisms by which retention affects children's peer acceptance the year following retention.

Method

Participants

Participants were 350 (52.6% male) first-grade children attending one of three school districts (1 urban, 2 small city) in central and southeast Texas, drawn from a larger ($N = 784$) sample of children participating in a longitudinal study of the effect of grade retention on academic achievement. Participants were recruited across two sequential cohorts in first grade during the fall of 2001 and 2002. Children were eligible to participate in the longitudinal study if they scored below the median score on a state-approved district-administered measure of literacy. On the basis of this criterion, we considered the sample to be academically at risk. Of 1,374 children who were eligible to participate in the study, written parental consent was obtained for 784 (57%). Children with and without consent to participate did not differ on age, gender, ethnic status, family language, bilingual class placement, or literacy test scores. Children with consent were somewhat more likely to be eligible for free or reduced-price lunch (62%) than children without consent (38%), perhaps because parents of participants received monetary compensation for completing questionnaires.

A total of 350 (45%) participants (ethnic composition was 74 African American, 132 Hispanic, 130 caucasian, 12 Asian or Pacific Islander, and 2 other) had complete data on teacher questionnaires and peer sociometric evaluations administered in year 1 (when all students were in first grade) and 1 year later as well as data on retention status following first grade (i.e., promoted to second grade or retained in first grade). To determine whether the 350 children with complete data differed from the 434 children without complete data on any of six demographic variables (age, gender, ethnicity, limited English proficiency, eligibility for free or reduced-price lunch, highest adult educational level in the home) and on all study variables at baseline, we conducted either chi-square contingency analyses for categorical variables or independent sample *t*-tests for continuous variables. No statistically significant differences were found. At entrance to first grade, participants' mean age was 6.56 ($SD = .35$) years. Their mean intelligence, as measured with the Universal Nonverbal Intelligence Test (Bracken & McCallum, 1998), was 93.64 ($SD = 14.52$), and their mean age standard scores on the Woodcock Johnson III (Woodcock, McGrew, & Mather, 2001) Broad Reading and Broad Math Scales were 97.96 ($SD = 18.07$) and 102.12 ($SD = 13.58$), respectively. Based on family income, 59.70% of participants were eligible for free or reduced-price lunch. For 25.4%, the highest educational level in the household was a high school certificate or less. School records showed that 22% of children had limited proficiency in English. The 350 students were located in 116 classrooms and 30 schools.

Design Overview

From November to March of year 1, when study participants were in first grade, research staff individually administered tests of reading and math achievement. These tests were readministered 1 year later. In March of each year, teachers were mailed a questionnaire packet

for each child. This packet included the measures of the teacher's perception of the child's emotional symptoms, conduct problems, and academic competencies. Teachers received compensation for completing and returning the questionnaires. Classmates' perceptions of the child's emotional symptoms, conduct problems, and academic competencies as well as their liking for the child were obtained via individual interviews conducted between February and May of year 1 and again 1 year later. We obtained written parent consent for each child who participated in the sociometric interview. However, all children in a classroom were eligible to be rated or nominated. Terry (1999) reported that reliable and valid sociometric data can be collected using the unlimited-nomination approach when as few as 40% of children in a classroom participate. When participation rates fall below 40%, results may not generalize to those that would have been obtained under conditions of full participation. Thus, we computed sociometric scores only for children in classrooms in which more than 40% of classmates participated in the sociometric assessment. The mean rate of classmate participation in sociometric administrations was .65 (range .40 to .95). Sociometric assessments were conducted in 180 of 198 classrooms in year 1 and in 224 of 274 classrooms in year 2. In year 2, 63 retained children were in first grade and 287 promoted children were in second grade.

Materials

Teacher perception of academic competence—Teachers were asked to describe participants' academic performance on three items using a Likert-type scale ranging from 1 (almost never) to 6 (almost always). The items included "Performing academically at grade level," "Able to read grade-level material and answer questions about what he/she has read," and "Able to solve grade-level math problems." The internal consistency of the scale was .94 for this sample. Teachers were also asked to describe the child's engagement in the classroom. This 10-item scale is comprised of eight items from the Conscientiousness Scale of the Big Five Inventory (BFI, John & Srivastava, 1999) and two items from the Social Competence Scale (Conduct Problems Prevention Research Group, 2004). Example items from the BFI include "Is a reliable worker," "Perseveres until the task is finished," and "Is easily distracted." The two items from the Social Competence Scale include "Sets and works toward goals" and "Turns in homework." The internal consistency for the 10-item scale for our sample was .86.

Peer-report measures—We used a modified version of the class play method (Masten, Morrison, & Pelligrini, 1985) to assess peers' perceptions of children's academic, social, and behavioral competencies. Children were asked to name classmates who best fit each of several behavioral descriptors. Children were told they could list as few or as many classmates as they wanted for each descriptor. We obtained a child's peer nomination score for each item by summing all nominations received. Scores were standardized within classrooms. Three peer-nomination items that assess peers' perceptions of classmates' academic competencies are relevant to this study: best at schoolwork ("These kids are best at schoolwork. They almost always get good grades and teachers often use their work as examples for the rest of the class"); best at math ("These kids are best in math. They almost always get good grades in math and the teacher calls on them to work hard math problems"); and best at reading ("These kids are best in reading. They usually get good grades in reading, and the teacher calls on them to read aloud or read hard words"). We computed a composite score, peer perception of academic performance (peer academic), as the average standardized score on these three items ($\alpha = .84$). During the sociometric interview, children were also asked to name all the children in their classrooms whom they "liked the most."

Children also were asked to indicate their liking for each child in the classroom on a five-point scale. Specifically, the interviewer named each child in the classroom and asked the child to point to one of five faces ranging from sad (1 = don't like at all) to happy (5 = like very much). A child's mean liking score was the average rating received by classmates. Following Coie,

Dodge, and Coppotelli (1982), we computed social preference scores as the standardized liked-most nomination score minus the standardized liked-least scores. To avoid asking children to nominate disliked children, a rating of “1” was considered equivalent to a “liked least” nomination score (Asher & Dodge, 1986). All sociometric scores were standardized within classrooms.

Academic achievement—The WJ-III Tests of Achievement (Woodcock et al., 2001) is an individually administered measure of academic achievement for individuals ages 2 to adulthood. For our purposes we used the WJ-III Broad Reading scores (Letter-Word Identification, Reading Fluency, and Passage Comprehension subtests) and the WJ-III Broad Math scores (Calculations, Math Fluency, and Math Calculation Skills subtests). Both age and grade standard scores were used. Extensive research has documented the reliability and construct validity of the WJ-III and its predecessor (Woodcock & Johnson, 1989; Woodcock et al., 2001). The 1-year stability for this age group ranges from .92 to .94 (Woodcock et al., 2001).

The Bateria Woodcock-Muñoz: Pruebas de Aprovechamiento—Revisada (Woodcock & Munoz, 1996) is the comparable Spanish version of the Woodcock-Johnson Tests of Achievement—Revised (WJ-R; Woodcock & Johnson, 1989), the precursor of the WJ-III. If children or their parents spoke any Spanish, children were administered the Woodcock-Munoz Language Test (Woodcock & Munoz-Sandoval, 1993) to determine the child’s language proficiency in English and Spanish and selection of either the WJ-III or the Bateria-R. The Woodcock Compuscore (Woodcock & Munoz-Sandoval, 2001) program yields scores for the Bateria-R that are comparable to scores on the WJ-R.

Results

The hypothesized model is shown in Figure 1. The boldface arrows indicate the target mediation effects in which teacher-rated academic competence and peer-rated academic competence mediate the short-term effect of retention on peer acceptance measured in the following year. Table 1 presents descriptive statistics (i.e., means and standard deviations) for the manifest variables for retained children ($N = 63$) and promoted children ($N = 287$), respectively. At time 1, when all children were in first grade, teachers perceived children who were subsequently retained as less engaged ($F(1, 265) = 7.10, p = .008$) and as achieving less ($F(1, 316) = 29.03, p < .001$). Retained and promoted children did not differ at time 1 on any of the peer-rated variables. Table 2 presents the correlations between the continuous manifest variables. All correlation coefficients were significant at $p < .01$.

We examined all structural models by using maximum-likelihood estimation with robust standard errors and a mean-adjusted chi-square statistic test (MLR; Muthén & Muthén, 2004). To account for the dependency among the observations (students) within clusters (classrooms), analyses were conducted using the “complex analysis” feature in Mplus (Muthén & Muthén, 2004), which accounts for the nested structure of the data by adjusting the standard errors of the estimated coefficients.¹

The first step in testing for mediation is to establish a statistically significant relation between the predictor and outcome variables (Baron & Kenny, 1986; Kenny, Kashy, & Bolger, 1998). We examined a model with only direct effects from retention status and time 1 peer acceptance

¹Current implementations of Multilevel Structural Equation Models (MSEM) under Mplus are restricted to general two-level models. In fact, Moerbeek (2004) showed that ignoring a level of the data structure (e.g., fitting a two-level model to a dataset with a three-level structure) resulted in losing statistical power. Hence, the statistical tests we obtained from Mplus that assumed a two-level structure (students nested within classrooms) were relatively conservative (i.e., with lower statistical power) compared with the statistical tests in which all three levels of the data structure (i.e., students nested within classrooms nested within schools) could be taken into account.

to time 2 peer acceptance. This model fit the data adequately ($\chi^2(3) = 1.07, p = .78$; CFI = 1.00; RMSEA = .000; SRMR = .004). All estimates were significant at $p < .05$ except for the covariance between retention and time 1 peer acceptance. Retention had a significant positive effect on peer acceptance measured in the following year after controlling for the previous peer-acceptance score ($\gamma = .11, p < .05$). We also analyzed the same model using the missing procedure (i.e., Type = Missing) under Mplus, which is a full information maximum-likelihood (FIML) method of addressing missingness. We found a similar pattern of results.

The second and third steps of testing mediation are to establish a statistically significant relation between the predictor and mediator variables and between the mediator and outcome variables (Baron & Kenny, 1986; Kenny et al., 1998). We examined two mediators simultaneously, teacher- and peer-perceived academic competence; the results are presented in Figure 2. To obtain stronger inferences of the mediation effects, we also controlled for the effects of both time 1 peer acceptance and academic competence on the corresponding time 2 measures.² The mediation model fit the data reasonably ($\chi^2(23) = 69.78, p < .001$; CFI = .97; RMSEA = .076; SRMR = .040). Retention had significant effects on both teacher-perceived academic competence [$\hat{\alpha}$ (retention & academic competence2 standardized) = .24, $\hat{\alpha}$ (retention & academic competence2 nonstandardized) = .60, $SE_{\hat{\alpha}}$ (retention & academic competence2) = .13, $p < .05$] and peer-perceived academic competence [$\hat{\alpha}$ (retention & peer academic2 standardized) = .25, $\hat{\alpha}$ (retention & peer academic2 nonstandardized) = .49, $SE_{\hat{\alpha}}$ (retention & peer academic2) = .12, $p < .05$] measured in the following year. In turn, both teacher- and peer-perceived academic competence also had significant positive effects on peer acceptance measured in the following year [$\hat{\beta}$ (academic competence2 & peer acceptance2 standardized) = .18, $\hat{\beta}$ (academic competence2 & peer acceptance2 nonstandardized) = .18, $SE_{\hat{\beta}}$ (academic competence2 & peer acceptance2) = .07, $p < .05$; $\hat{\beta}$ (peer academic2 & peer acceptance2 standardized) = .35, $\hat{\beta}$ (peer academic2 & peer acceptance2 nonstandardized) = .43, $SE_{\hat{\beta}}$ (peer academic2 & peer acceptance2) = .09, $p < .05$] after controlling for the peer acceptance measured in the previous year.

According to Baron and Kenny's (1986) procedure, the mediation effects can be tested by multiplying the nonstandardized path coefficients corresponding to the mediation effects. Sobel's (1982) test of mediation effects, that is,

$$Z_{\hat{\alpha}\hat{\beta}} = \frac{\hat{\alpha} * \hat{\beta}}{SE_{\hat{\alpha}\hat{\beta}}}$$

along with the delta method standard error (i.e., $SE_{\hat{\alpha}\hat{\beta}} = \sqrt{\hat{\alpha}^2 * (SE_{\hat{\beta}})^2 + \hat{\beta}^2 * (SE_{\hat{\alpha}})^2}$) was employed (Krull & MacKinnon, 1999, 2001). The result showed that both teacher- and peer-perceived academic competence significantly mediated the short-term relation between retention status at the end of first grade and peer acceptance in the following year [$Z_{\hat{\alpha}}(\text{retention competence2}) * \hat{\beta}(\text{competence2 acceptance2}) = 2.25, p < .05$; $Z_{\hat{\alpha}}(\text{retention smart2}) * \hat{\beta}(\text{smart2 acceptance2}) = 3.10, p < .05$]. The direct effect of retention on peer acceptance as shown in Figure 2 was not significant, indicating a full mediation effect by both teacher- and peer-

²Due to the data-collection procedure, some students had missing responses of peer perceptions of academic performance at time 1. We analyzed the same mediation model shown in Figure 2 using students with available time 1 peer perceptions of academic performance ($N = 202$). A significant effect of retention on peer perceptions of academic performance at time 2 after controlling for peer perceptions of academic performance at time 1 was also found [$\hat{\alpha}$ (retention & peer academic2 standardized) = .22, $p < .05$]. The pattern of results from this subsample was very similar to the one presented in Figure 2 ($N = 350$).

perceived academic competence. We obtained similar results from the data using the FIML procedure to handle missingness (i.e., coefficients presented in the parentheses in Fig. 2).

Because our measure of academic competence was based on teacher and peer perceptions of students' academic performance in the classroom, one cannot reach any conclusions from the above analyses regarding the effect of retention on children's actual academic achievement. Thus, to place our findings in the context of children's actual academic achievement and to permit comparisons of our results with those of published studies on the effect of retention on achievement, we compared the WJ-III Broad Reading and Broad Math age and grade standard scores of retained and promoted children at time 2, controlling for the relevant time 1 scores and taking the dependency into account. Table 1 includes descriptive statistics for these measures. Using the mixed routine under SPSS, we found a significant difference between retained and promoted children on all time 2 test scores [i.e., F (Broad Reading age standard score) (1, 104) = 26.85, $p < .001$; F (Broad Reading grade standard score) (1, 222) = 67.38, $p < .001$; F (Broad Math age standard score) (1, 169) = 12.10, $p < .005$; F (Broad Math grade standard score) (1, 234) = 72.45, $p < .005$]. The denominator degrees of freedom of the F -test are estimated by a general Satterthwaite approximation. Retained students scored significantly lower than the promoted students on both time 2 Broad Reading and Math age standard scores. Conversely, retained students scored significantly higher than the promoted students on both time 2 Broad Reading and Math grade standard scores.

Discussion

As we expected, among a sample of academically at-risk first graders, children who repeated first grade improved more in peer acceptance the following (repeat) year than did children who were promoted to second grade. Furthermore, this direct effect was completely mediated by teacher- and peer-perceived academic competence. It is likely that repeaters gain in peer acceptance because their younger, less experienced classmates view them as meeting the day-to-day academic and behavioral challenges better during their repeat year than their same-age, equally experienced classmates viewed them the prior year. Because promoted children face new academic and behavioral challenges in the year following first grade, their same-age, equally experienced classmates do not view them as more academically competent, and therefore promoted children do not gain in peer acceptance. The descriptive statistics on observed variables provided in Table 1 show that in year 1, when all children were enrolled in first grade for the first time, children who subsequently repeated first grade and children who were subsequently promoted did not differ on peer-rated liking, social preference, and academic competence. However, first-grade teachers rated children subsequently promoted as more engaged and as achieving more than retained children. The following year, however, repeaters made substantial gains on each of these measures. In the second year, repeaters' standard mean liking score and social preference scores were close to the mean classroom scores. Their standard score for peer-rated academic competence surpassed the scores of 61.4% of their classmates. With respect to teacher evaluations of achievement, in year 2 retained children surpassed both the year 1 (first grade) ($F(1, 275) = 11.11, p < .001$) and year 2 (second grade) ($F(1, 163) = 12.13, p < .001$) scores of their promoted peers. Retained children's year 2 scores on teacher evaluations of engagement, however, did not differ significantly from the scores of their promoted peers at year 1 or year 2.

Because this study focused on the effect of grade retention on children's peer relations in the classroom, we investigated classroom-based measures of academic competence rather than norm-referenced measures, such as standardized achievement tests. The sociometric measure of academic competence reflects a child's standing relative to classmates. It is reasonable to expect that the teacher ratings of children's engagement and achievement are similarly grounded in the day-to-day realities of a particular classroom and the child's performance

relative to other students in the class. Study measures are subjective evaluations made by individuals who share a child's day-to-day school environment. The finding that these measures predict changes in children's peer acceptance supports the social validity of these measures of competence. The retained child is viewed by classmates and the teacher as more academically capable during the repeat year, a difference that translates into enhanced peer acceptance.

Importantly, the classroom-based measures of academic competence we used do not address children's actual growth in academic skills. Generally, when repeaters are compared with ability-matched, promoted peers on measures of academic achievement, results differ based on whether retained children are compared with their same-age or same-grade peers. Same-age comparisons favor promoted students more often than do same-grade comparisons (for a review, see Alexander et al., 2003). Consistent with past research, our supplementary analyses indicated that retained children fare better than promoted children on grade standard scores and worse than promoted children on age standard scores. These findings underscore the importance of considering different ways of measuring academic competence in evaluating retention effects, because results can differ depending on the measure employed.

Limitations and Future Research

This study stands out from recent studies on the effects of grade retention in that we document positive short-term effects of early grade retention. Our results suggest that repeaters benefit from the extra year in grade in terms of classroom competencies and peer acceptance. In interpreting these results, it is important to realize that students were not assigned randomly to retention and promotion groups, and that unmeasured variables may account for study effects. However, the longitudinal design that permits controlling for year 1 peer acceptance and academic competence offers a strong but not convincing basis for ruling out baseline differences as an explanation for study findings.

A limitation of our study is the lack of additional years of data collection. A critical issue that can only be addressed as more waves of data are collected with the sample is whether these short-term social and academic gains are sustained as repeaters meet new academic challenges. It is well established that retention often has different short- and long-term effects that vary by developmental period (Alexander et al., 2003; Pagani, Tremblay, Vitaro, Boulerice, & McDuff, 2001; Reynolds, 1992). In the most detailed and longest study of the effects of retention in first grade on academic and behavioral outcomes through high school, Alexander et al. (2003) found that early-retained children in Baltimore schools were more likely to drop out of school in adolescence, relative to matched promoted children, despite performing better in their classwork than promoted children. They concluded: "Retention, so far as we can determine, does not impede ... children academically or assault their self-esteem in the early years, yet something about the experience apparently weakened repeaters' attachment to school" (p. ix). Similarly, Pagani et al. (2001) employed autoregressive modeling and documented differences in short- and long-term effects of being retained in the primary grades. They concluded that some negative effects did not appear until several years after retention, and that most negative effects were amplified over time.

A second limitation of our study is the high rate of missing data (45% of participants had complete data). The largest single factor contributing to the missing data is the lack of sociometric data for both years 1 and 2. A total of 331 children were missing sociometric data for at least 1 year. There are many obstacles to obtaining sociometric data, including teachers who do not facilitate the consent return, and students moving to districts outside the three participating ones. Despite the difficulty of collecting sociometric data, such data are considered the gold standard in peer-relations research (see Cillessen & Bukowski, 2000, for

a review). The level of missing data suggests caution in generalizing our results to the original sample. However, we found no statistically significant differences between children with and without complete data at baseline on any demographic or study variable. Also, analyses conducted with data using the full sample and the FIML procedure for missing data and with the sample with complete data yielded similar results. These findings reduce but do not eliminate the concern regarding generalization of study findings.

The differences in short- and long-term effects emphasize the importance of studying developmental processes over time that may account for these effects. Our results suggest that, in the repeat year, repeaters succeed academically and socially. This successful year follows a year of academic and social struggle that culminates in what children refer to as “flunking” and having to repeat a grade while one’s classmates go on to the next grade. Many repeaters will likely struggle again when they are promoted to second grade and are faced with novel academic challenges. If this is the case, our results suggest that repeaters may lose their recently acquired social status, because they will decline in peer- and teacher-perceived ability to meet classroom challenges. This “struggle-succeed-struggle” sequence may have long-term negative consequences for children’s academic motivation and achievement that explain why retained children subsequently become disaffected with school. A more optimistic scenario, one consistent with Furrer and Skinner (2003) and Buhs and Ladd (2001), is that the peer support retained students experience in the repeat year increases their sense of school relatedness and academic motivation, resulting in more positive academic trajectories.

Acknowledgments

This research was supported in part by a grant to Jan N. Hughes from the National Institute of Child Health and Human Development (5 R01 HD39367-02). Authorship order is alphabetical due to equal contribution from each of us.

References

- Alexander, KL.; Entwisle, DR.; Dauber, SL. On the success of failure: A reassessment of the effects of retention in the primary grades. New York: Cambridge University Press; 1994.
- Alexander, KL.; Entwisle, DR.; Dauber, SL. On the success of failure: A reassessment of the effects of retention in the primary grades. 2. Cambridge: Cambridge University Press; 2003.
- Asher SR, Dodge KA. Identifying children who are rejected by their peers. *Developmental Psychology* 1986;22:444–449.
- Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 1986;5:1173–1182. [PubMed: 3806354]
- Beebe-Frankenberger M, Bocian KM, Mac-Millan DL, Gresham FM. Sorting second-grade students: Differentiating those retained from those promoted. *Journal of Educational Psychology* 2004;96:204–215.
- Bracken, BA.; McCallum, RS. Universal Nonverbal Intelligence Test: Examiner’s manual. Itasca, IL: Riverside; 1998.
- Buhs ES, Ladd GW. Peer rejection as an antecedent of young children’s social adjustment: An examination of mediating processes. *Developmental Psychology* 2001;37:550–560. [PubMed: 11444490]
- Cillessen, AHN.; Bukowski, WM., editors. Conceptualizing and measuring peer acceptance and rejection. San Francisco: Jossey-Bass; 2000.
- Coie JD, Dodge KA, Coppotelli HA. Dimensions and types of social status: A cross-age perspective. *Developmental Psychology* 1982;18:557–569.
- Coie, JD.; Dodge, KA.; Kupersmidt, JB. Peer group behavior and social status. In: Asher, SA.; Coie, JD., editors. *Peer rejection in childhood*. New York: Cambridge University Press; 1990. p. 17-59.
- Conduct Problems Prevention Research Group. Teacher social competence. 2004. [Online]. Available: <http://www.fasttrackproject.org/techrept/t/tsc/>

- Furrer C, Skinner E. Sense of relatedness as a factor in children's academic engagement and performance. *Journal of Educational Psychology* 2003;95:148–162.
- Holmes CT, Matthews KM. The effects of nonpromotion on elementary and junior high school pupils: A meta-analysis. *Review of Educational Research* 1984;54:225–236.
- Jimerson SR. On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology* 1999;37:243–272.
- Jimerson SR. Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review* 2001;30:420–437.
- Jimerson SR, Anderson GE, Whipple AD. Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools* 2002;39:441–457.
- Jimerson S, Carlson E, Rotert M, Egeland B, Sroufe LA. A prospective longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology* 1997;35:3–25.
- John, OP.; Srivastava, S. The big five taxonomy: History, measurement, and theoretical perspectives. In: Pervin, LA.; John, OP., editors. *Handbook of personality: Theory and research*. New York: Guilford; 1999. p. 102-138.
- Kenny, DA.; Kashy, DA.; Bolger, N. Data analysis in social psychology. In: Gilbert, DT.; Fiske, ST.; Lindzey, G., editors. *Handbook of social psychology*. 4. New York: McGraw-Hill; 1998. p. 233-265.
- Krull JL, MacKinnon DP. Multilevel mediation modeling in group-based intervention studies. *Evaluation Review* 1999;23:418–444. [PubMed: 10558394]
- Krull JL, MacKinnon DP. Multilevel modeling of individual and group level mediated effects. *Multivariate Behavioral Research* 2001;36:249–277.
- Kupersmidt, JB.; Coie, JD.; Dodge, KA. The role of poor peer relationships in the development of disorder. In: Asher, SR.; Coie, JD., editors. *Peer rejection in childhood*. New York: Cambridge University Press; 1990. p. 274-305.
- Lerner, RM. Theories of human development: Contemporary perspectives. In: Damon, W.; Lerner, RM., editors. *Handbook of child psychology* (5th ed.). Vol. 1: Theoretical models of human development. New York: Wiley; 1998. p. 685-759.
- Mantzicopoulos P. Do certain groups of children profit from early retention? A follow-up study of kindergartners with attention problems. *Psychology in the Schools* 1997;34:115–127.
- Mantzicopoulos P, Morrison D. Kindergarten retention: Academic and behavioral outcomes through the end of second grade. *American Educational Research Journal* 1992;29:182–198.
- Masten AS, Morrison P, Pelligrini DS. A revised class play method of peer assessment. *Developmental Psychology* 1985;21:523–533.
- McCoy AR, Reynolds AJ. Grade retention and school performance: An extended investigation. *Journal of School Psychology* 1999;37:273–298.
- Meidel WT, Reynolds AJ. Parent involvement in early intervention for disadvantaged children: Does it matter? *Journal of School Psychology* 1999;37:379–402.
- Meisels SJ, Liaw FR. Failure in grade: Do retained students catch up? *Journal of Educational Research* 1993;87:69–77.
- Moerbeek M. The consequence of ignoring a level of nesting in multilevel analysis. *Multivariate Behavioral Research* 2004;39(1):129–149.
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. Los Angeles: Muthén & Muthén; 2004.
- Pagani L, Tremblay RE, Vitaro F, Boulerice B, McDuff P. Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology* 2001;13:297–315. [PubMed: 11393648]
- Parker JG, Asher SR. Peer relations and later personal adjustment: Are low accepted children “at risk”? *Psychological Bulletin* 1987;102:357–389. [PubMed: 3317467]
- Pianta RC, Teitbohl PJ, Bennett EM. Differences in social adjustment and classroom behavior between children retained in kindergarten and groups of age and grade matched peers. *Early Education and Development* 1997;8:137–152.

- Pierson LH, Connell JP. Effect of grade retention on self-system processes, school engagement, and academic performance. *Journal of Educational Psychology* 1992;84:300–307.
- Plummer DL, Graziano WG. Impact of grade retention on the social development of elementary school children. *Developmental Psychology* 1987;23:267–275.
- Reynolds AJ. Grade retention and school adjustment: An explanatory analysis. *Educational Evaluation and Policy Analysis* 1992;14:101–121.
- Rubin, KA.; Bukowski, W.; Parker, JG. Peer interactions, relationships, and groups. In: Damon, W.; Eisenberg, N., editors. *Handbook of child psychology*. Vol. 3. New York: Wiley; 1998. p. 619-700.
- Rumberger R. Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal* 1995;32:583–625.
- Shepard LS, Smith ML. Synthesis of research on grade retention. *Educational Leadership* 1990;47:84–88.
- Sobel, ME. Asymptotic confidence intervals for indirect effects in structural equation models. In: Leinhardt, S., editor. *Sociological methodology*. San Francisco: Jossey-Bass; 1982. p. 290-312.
- Terry, R. Measurement and scaling issues in sociometry: A latent trait approach. Paper presented at the biennial meeting of the Society for Research in Child Development; Albuquerque, NM. 1999 Apr.
- Woodcock, RW.; Johnson, MB. *Woodcock-Johnson Psycho-educational Battery-Revised*. Allen, TX: DLM Teaching Resources; 1989.
- Woodcock, RW.; McGrew, KS.; Mather, N. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside; 2001.
- Woodcock, RW.; Munoz, AF. *Batería Woodcock-Munoz Pruebas de Habilidad Cognoscitiva-Revisada*. Chicago: Riverside; 1996.
- Woodcock, RW.; Munoz-Sandoval, AF. *Woodcock-Munoz Language Survey*. Chicago: Riverside; 1993.
- Woodcock, RW.; Munoz-Sandoval, AF. *Woodcock-Munoz Language Survey Normative Update*. Itasca, IL: Riverside; 2001. *WMLS Normative Update Scoring and Reporting Program* [Computer software].

Fig. 1.
Hypothesized model

Fig. 2.

Mediation effects of academic competence and peer academic on the relation between retention status and peer acceptance. $\chi^2(23) = 69.78, p < .001$; CFI = .97; RMSEA = .076; SRMR = .040. The target hypothesized paths are bold. Peer academic2 has been adjusted by using the corresponding reliability coefficient (i.e., coefficient alpha of peer academic2 = .92). To provide a clear presentation of the model, correlations between the exogenous variables (i.e., academic competence1, peer acceptance1, retention status after year 1) are not presented in the figure. All correlations were significant (at $p < .05$) except the one between peer acceptance1 and retention status ($\chi^2 = -.07, p > .10$). All estimates were significant at $p < .05$ (two-tailed) except the italic ones (e.g., the path coefficient from retention status to peer acceptance2 was not significant). All coefficients were standardized ($N = 350$); coefficients in parentheses were based on the results from the Mplus missing data procedure (i.e., Type = Missing).

Descriptive Statistics for Analysis Variables for Retained and Promoted Children at Baseline and 1 Year Later

Table 1

Variable	Retained (N = 63)				Promoted (N = 287)			
	Year 1		Year 2		Year 1		Year 2	
	M	SD	M	SD	M	SD	M	SD
Peer evaluation:								
Liking rating	-.23	1.00	.02	.98	-.08	.93	-.19	1.02
Social preference	-.18	.94	.01	1.08	-.03	.92	-.14	.96
Peer academic ^a	-.25	.79	.29	1.06	-.29	.68	-.19	.81
Teacher evaluation:								
Achievement	3.46	1.58	4.81	1.04	4.13	1.32	4.15	1.35
Engagement	2.98	.97	3.49	1.08	3.29	1.04	3.39	1.06
Woodcock Johnson III:								
Broad Reading:								
Age standard score	86.97	13.05	85.38	13.00	100.32	8.14	100.85	5.07
Grade standard score	85.41	12.04	106.41	14.71	102.51	16.11	103.60	14.53
Broad Math:								
Age standard score	96.70	11.33	94.83	10.00	103.26	13.76	103.25	12.46
Grade standard score	95.83	11.32	111.97	12.79	104.19	13.10	104.66	12.23

^aThe N for peer academic variable is 48 (retained) and 154 (promoted).

Table 2

Zero-Order Correlations for All Continuous Variables

Scale	1	2	3	4	5	6	7	8	9
1. Achievement1	...								
2. Engagement1	.65	...							
3. Peer liking1	.37	.42	...						
4. Social preference1	.36	.37	.82	...					
5. Achievement2	.54	.43	.29	.31	...				
6. Engagement2	.41	.63	.44	.39	.63	...			
7. Peer academic2	.33	.45	.35	.30	.47	.53	...		
8. Peer liking2	.24	.36	.44	.43	.34	.47	.49	...	
9. Social preference2	.24	.38	.43	.41	.32	.45	.50	.85	...

Note. $N = 350$. All correlations are significant at $p < .01$ (two-tailed).