Published in final edited form as:
J Sch Psychol. 2015 February ; 53(1): 7–24. doi:10.1016/j.jsp.2014.10.001.

Effect of Retention in Elementary Grades on Grade 9 Motivation for Educational Attainment

Heining Cham,
Department of Psychology, Fordham University, 226 Dealy Hall, 441 E. Fordham Road, Bronx, NY 10458. hcham@fordham.edu

Jan N. Hughes,
Department of Educational Psychology, Texas A & M University, Box 4225, College Station, TX 77843-4225. jhughes@tamu.edu

Stephen G. West, and
Department of Psychology, Arizona State University, P.O. Box 871104, Tempe, AZ 85287-1104. sgwest@asu.edu

Myung Hee Im
Department of Educational Psychology, Texas A & M University, Box 4225, College Station, TX 77843-4225. myunghee.im@gmail.com

Abstract

This study investigated the effect of grade retention in elementary school on students' motivation for educational attainment in grade 9. We equated retained and promoted students on 67 covariates assessed in grade 1 through propensity score weighting. Retained students (31.55%, $n_{retained}$ = 177) and continuously promoted students (68.45%, $n_{promoted}$ = 384) were compared on the bifactor model of motivation for educational attainment (Cham, Hughes, West, & Im, 2014). This model consists of a General factor (student's overall motivation for educational attainment), and three specific factors: student perceived Teacher Educational Expectations, Peer Educational Aspirations, and Value of Education. Measurement invariance between retained and promoted groups was established. Retained students scored significantly higher than promoted students on each specific factor but not on the General factor. Results showed that the retained and promoted students did not significantly differ on the General factor. The retained students had significantly higher scores on each specific factor than the promoted students. The results suggested that grade retention may not have the negative effects so widely assumed in the published literature; it is an expensive intervention with minimal evidence of benefits to the retained student.

Correspondence concerning this article should be addressed to Heining Cham, Department of Psychology, Fordham University, Bronx, NY 10458. hcham@fordham.edu or Jan N. Hughes, Department of Educational Psychology, Texas A & M University, College Station, AZ 77843-4222. jhughes@tamu.edu.

Keywords

Retention; elementary grades; motivation; educational attainment

The substantial percentage of students who leave school without a high school diploma is a major concern for educators, policy-makers, and society at large. In 2009, 8.1 percent of 18- through 24-year-olds had not received a high school diploma or alternative credential and were not currently enrolled in high school (Chapman, Laird, & KewalRamani, 2011). Failure to attain a high school degree predicts life-long economic, occupational, social, and health disparities (Alliance for Excellence in Education, 2007; Pleis, Ward, & Lucas, 2010). The decision to drop out of school is the last step of a gradual process of disengagement that begins as early as middle school and increases in high school (Alexander, Entwisle, & Horsey, 1997; Englund, Egeland, & Collins, 2008; Janosz, Archambault, Morizot, & Pagani, 2008; Reschly & Christenson, 2013). Given the serious negative consequences of dropping out of school for the individual and for society, researchers have sought to identify factors that predict drop out in hopes of developing interventions that ameliorate risk processes.

Retention in grade consistently predicts subsequent dropping out of school (Alexander, Entwisle, & Kabbani, 2001; Bowers & Sprott, 2012). However, this association may be a result of a host of inter-related factors at the individual, family, and school levels that predict not only grade retention but also high school completion. The current study applies more rigorous statistical controls for such confounders than have been employed in prior research to isolate an effect of grade retention in the elementary grades on students' grade 9 motivation to complete high school and pursue post-secondary education.

Grade Retention as Risk Factor for Dropping Out of School

A number of studies report that students who repeated one or more grades in school are at least twice as likely to drop out of school as are students who are continuously promoted (for a review, see Jimerson, Anderson, & Whipple, 2002). The association between grade retention and dropping out of school has led researchers to suggest that grade retention has a negative effect on school completion (Grissom & Shepard, 1989; Jimerson, 1999; Roderick, 1994; Rumberger, 1987). However, these studies generally failed to employ research designs that adequately controlled for pre-existing differences between students who were retained and those who were promoted (Allen, Chen, Willson, & Hughes, 2009). Importantly, students are not randomly selected into the "intervention" of grade retention, and retained students differ from promoted students on a number of confounders that predict academic attainment even prior to grade retention, including low achievement, conduct problems, poor relationships with teachers, low parental involvement in school, and poverty (Barrington & Hendricks, 1989; Englund et al., 2008; Janosz et al., 2008; Roderick, 1994; Wang & Fredricks, 2014; Willson & Hughes, 2006). Thus, grade retention may be confounded with other pre-existing vulnerabilities that place students at-risk for poor academic and behavioral functioning that ultimately leads to dropping out of school. These pre-existing vulnerabilities provide an alternative explanation for why students leave school early. Next, we summarize previous research on the effect of grade retention on academic

achievement and on psychosocial variables that place students at-risk for dropping out of school.

Effect of grade retention on academic achievement

Low academic performance is the strongest predictor of leaving school prior to earning a high school diploma (Battin-Pearson et al., 2000; Newcomb et al., 2002; Wang & Fredricks, 2014). Thus, research on the effect of grade retention on academic performance has the potential to clarify the association between grade retention and school drop out. Early reviews of the literature on the effects of grade retention on subsequent academic achievement concluded that grade retention had a negative effect on achievement (for meta-analytic reviews see Holmes, 1989; Jimerson, 2001; for narrative reviews see Jimerson et al., 2002; Sipple, Killeen, & Monk, 2004). However, most of the studies included in these reviews are plagued by significant methodological limitations, the most important being a lack of a comparison group of promoted peers equivalent prior to retention on achievement and other variables predictive of achievement.

The importance of adequate controls for pre-existing differences between retained and promoted students was highlighted recently in a meta-analysis of studies published between 1990 and 2007 on the effects of grade retention on subsequent academic performance (Allen et al., 2009). This meta-analysis found that the studies employing more rigorous controls for student, school, and family characteristics associated with selection into the grade retention intervention were less likely to find that retention has a negative effect on achievement. Recent studies on the effect of grade retention on academic motivation and achievement have utilized modern methods of controlling for selection effects in observational studies, including instrumental variable analysis (Alet, 2010; Dong, 2010) and propensity score analysis (Dong, 2010; Goos, Van Damme, Onghena, Petry, & de Bilde, 2013; Moser, West, & Hughes, 2012; Wu, West, & Hughes, 2008). In the first approach, instrumental variable analysis, the key procedure is to identify and measure an instrumental variable. An instrumental variable is a variable that is related only to the retention status but not directly related to the outcome variable. An unbiased estimate of the effect of grade retention is then obtained through proper statistical adjustment of the retention-outcome variable relationship using the instrumental variable (DeMaris, in press; Morgan & Winship, 2007, chapter 7). In the second approach, a propensity score, the estimated probability the student will be retained is calculated for each student based on a set of measured covariates. Each covariate potentially confounds the estimate of the effect of grade retention if (1) there are pre-existing (baseline) differences between the retained and promoted students on the covariate and (2) the covariate is related to the outcome variable of interest. Procedures such as matching and weighting can then be used to equate the retained and promoted students on their propensity scores (Schafer & Kang, 2008; West, Cham, Thoemmes, Renneberg, Schulze, & Weiler, 2014). When (a) all the true confounders are measured, (b) the distributions of the propensity scores of the retained and promoted students overlap, and (c) the propensity scores are correctly estimated, successful equating of the retained and promoted students' propensity scores theoretically implies successful equating on each of the covariates used in the calculation of the propensity score. Given successful equating is achieved on all confounders, the propensity score analysis produces an unbiased estimate of

the average effect of grade retention on students. Rubin (2001) and West et al. (2014) have emphasized the importance of attempting to identify and reliably measure all important confounding variables for inclusion in the construction of propensity scores. The propensity score analysis may have advantages over the instrumental variable analysis in some cases because it is less sensitive to violations of its assumptions and it permits probing of the likely consequences of violations of assumptions (Morgan & Winship, 2007; West et al., 2014).

In a study using the same longitudinal sample as the current study, Moser et al. (2012) used propensity score matching to equate the retained and promoted students based on a large and comprehensive set of potential confounders measured prior to any student being retained. In the retention year, the retained students performed better academically on nationally normed reading and math achievement measures, compared to their younger, equally at-risk but promoted grade-mate students. However, by grade 5, the propensity score matched retained and promoted students did not differ in reading and math achievement, although the retained students reached grade 5 one year later than the promoted students. Similar results were found in a study of Flemish students retained in grade 1 (Goos et al., 2013) that also employed propensity score matching. Similar results have been obtained in studies in the United States (Dong, 2010) and in France (Alet, 2010) that used instrumental variable analysis to control for selection factors at the child and school levels.

Effect of grade retention on psychosocial adjustment and motivation

Some researchers have suggested that retained students, being older than the majority of their classmates and having experienced a highly salient form of academic failure, may feel less attached (sense of belonging) to school and less academically capable, leading to reduced academic effort and persistence (Jimerson, 2001; Roderick, 1994). However, studies reporting on the effects of retention on students' subsequent psychosocial adjustment and educational motivation (e.g., liking for school, academic self-efficacy, behavioral engagement in school, peer acceptance, and conduct problems) have led to inconsistent results (for meta-analytic review see Jimerson, 2001; for narrative review see Wu, West, & Hughes, 2010). These inconsistencies may be due to study differences, including the adequacy of controls for pre-retention group differences, the timing of retention, or the length of time post-retention that outcomes were assessed. In a prior study of the current longitudinal sample (Wu et al., 2010), students retained in grade 1 had more favorable scores than their propensity score matched, promoted students 4 years later on several measures of psychosocial adjustment, including lower teacher-rated hyperactivity and higher teacher-rated behavioral engagement, fewer peer-rated sad and withdrawn behaviors, and greater self-reported academic self-efficacy. Despite these positive effects of earlier grade retention, the researchers cautioned that a pattern of an increase in peer acceptance and perceived school belonging immediately after retention, followed by declining levels over the next 3 years, could portend "trouble on the horizon" (p. 148).

According to developmental systems theory (Cox, Mills-Koonce, Propper, & Gariépy, 2010; Lerner, 1998) and life course theory (Elder, 1998), grade retention might have "sleeper" effects on students' psychosocial adjustment and educational motivation that emerge at

periods of developmental transitions, such as the transition to middle school or high school. Using the same longitudinal sample as the current study, Im, Hughes, Kwok, Puckett, and Cerda (2013) investigated the effect of retention in grades 1–5 on students' psychosocial and academic adjustment in middle school. Using propensity score matching to equate the retained and promoted students, piecewise growth curve modeling showed that students retained in the elementary grades and their promoted peers did not differ on measures of teacher-rated engagement, student reported school belonging, or academic achievement in the last year of elementary school, nor did they differ in their post-transition growth trajectories on these variables. Similarly, Goos et al. (2013) used propensity score matching to equate retained and promoted students in the Flanders region of Belgium. They found that students retained in 1st grade and equally at-risk but promoted grade-mate students did not differ in their growth trajectories on measures of psychosocial adjustment from first grade until the start of secondary school (at approximately age 12 for continuously promoted children).

Effect of Grade Retention on Motivation to Complete High School

Even if grade retention does not harm students' academic achievement or psychosocial adjustment, relative to same-grade peers, previously retained students may still be more likely to leave school without earning a high school diploma. Indeed, Alexander, Entwisle, and Dauber (2003) found that early-retained students in Baltimore schools, relative to matched promoted students, were more likely to drop out of school in adolescence, despite performing better in their coursework than promoted children. Alexander et al. 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).

To the best of our knowledge, no study has examined the effect of grade retention in the elementary grades on students' motivation to complete high school, measured at the beginning of high school. The transition from middle school to high school often involves larger schools, more extensive academic tracking, increasingly rigid social cliques defined by shared school identities, and frequent monitoring of progress toward meeting graduation requirements (Benner, 2012; Cox et al., 2010). Relative to their same-age peers who were continuously promoted in the elementary grades, students retained in the elementary grades are, on average, one year older in grade 9.[1] Not only are they "old for grade," but they are also one year closer to reaching the legal age for dropping out of school and the legal age for obtaining full time employment. Thus, their motivation to persist another 4 years to graduation may be less than that of their same-grade but younger peers.

[1]Studies of the effect of grade retention use either same-grade or same-age comparisons. In same-grade comparisons, retained and promoted students are compared on the outcome measure during the same academic year. Typically the retained students will be one grade behind and one year older than their promoted classmates at the time of comparison. Several researchers (e.g., Karweit, 1999) have suggested that same-grade comparisons are more appropriate for educational outcomes because retention effectively "recalibrates" students' performance to standards based on their current classmates. In contrast, psychosocial outcomes not directly tied to the educational context typically use same age comparisons.

The Role of Motivation for Education Attainment in High School Drop Out

Motivation for educational attainment and the factors that give rise to it have been the topic of extensive investigation for decades (e.g., Deci & Ryan, 2000; Pintrich, 2003; Wigfield, Cambria, & Eccles, 2012). Evidence accumulated over the past several decades (e.g., Deci & Ryan, 2000; Pintrich, 2003; Wigfield et al., 2012) documented the critical role of students' self-system beliefs (e.g., perceptions of academic ability, value of academic achievement, and academic goals) that underlie students' achievement-related behaviors, such as effort and persistence in a course and choice in coursework. Among the strongest and most consistent predictors of academic effort and engagement are students' academic competence beliefs and beliefs about the value of educational achievement. When students believe that they are capable of engaging in those behaviors necessary to achieve and value academic success, they are likely to invest energy in academic tasks and to persevere in the face of challenges (Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Eccles & Wigfield, 1995). High school students' beliefs in the value of educational attainment and in their academic competence are also strong predictors of their declared intention to drop out of school (Alivernini & Lucidi, 2011; Caprara et al., 2008; Legault, Green-Demers, & Pelletier, 2006).

According to both expectancy value theory (Wigfield & Eccles, 2000) and self-determination theory (Deci & Ryan, 1991), students' academic competence beliefs and valuing of education are shaped in important ways by the level of social support for achievement provided by key social agents, including parents, teachers, and friends. Consistent with these theoretical perspectives, a large body of research supports the importance of the social context on students' academic motivation (Deci & Ryan, 1991; Legault et al., 2006; Murdock, 1999). For example, Murdock (1999) reported that adolescents' perceived levels of teacher educational expectations for them and perceptions of their friends' academic aspirations were stronger predictors of their behavioral engagement at school than were demographic risk variables. Furthermore, these two social contexts partially mediated the effect of demographic risk on school engagement. With respect to the peer social context, grade 6 students who reported that their friends were academically engaged increased more in achievement over the next two years than students who reported lower levels of friend academic engagement (Véronneau & Dishion, 2011). Parental expectations for their children's academic success is also a key factor in students' engagement and achievement (Hill & Tyson, 2009; Hong & Ho, 2005). For example, Tynkkynen, Tolvanen, and Salmela-Aro (2012) found that adolescents' perceptions of their parents' educational expectations predicted their own 5-year trajectories for educational attainment.

Conceptualization and Measurement of Motivation for Educational Attainment

Conceptualization

Consistent with dynamic systems theory (Lerner, 1998), the current study views students' competence beliefs related to educational attainment, their valuing of educational

attainment, and the social context for educational attainment as influencing each other over time, as part of a dynamic, self-reinforcing motivational system (Legault et al., 2006; Wigfield et al., 2006). For example, students who value achievement, make good grades, and aspire to higher levels of educational attainment select friends who support and encourage academic attainment (Altermatt & Pomerantz, 2005; Kindermann, 2007; Kiuru, Aunola, Vuori, & Nurmi, 2007). Similarly, academically motivated and engaged students receive higher levels of teacher support which, in turn, promotes students' academic effort and persistence (Hughes, Luo, Kwok, & Loyd, 2008). Additionally, parent educational expectations for their students are both a cause and an effect of students' level of academic achievement (Hughes, Kwok, & Im, 2013). Based on this dynamic view of beliefs and social context, the present study conceptualizes motivation for educational attainment as a multi-dimensional construct that includes students' competence beliefs for educational attainment, valuing of educational attainment, and the social context for educational attainment provided by key social agents.

Measurement

The conceptualization above is consistent with a strong consensus that motivation for educational attainment is a multi-dimensional construct. Cham, Hughes, West, and Im (2014) developed a 32-item measure of adolescents' motivation for educational attainment, defined as completing high school and pursuing post-secondary education. Among the 32 items, 20 items were drawn from Murdock's (1999) measure of perceived motivational context (perceived teacher educational expectations, teachers' emotional support, peers' educational aspirations, and the economic value of educational success). An additional 12 items were developed by the researchers to assess students' perceived competence (Zimmerman & Cleary, 2006), expectations for graduation from high school and enrollment in post-secondary education, and perceptions of parents' and friends' expectations for their high school and postsecondary attainment. Using the same longitudinal sample as the current study, Cham et al.'s factor analysis supported a bifactor model of the assessment (Figure 1; see Appendix A for the full set of the items). The bifactor model consists of a General factor, on which all 32 items load that captures the commonality of the full set of motivation for educational attainment items. In other words, the General factor reflects a student's basic, overall motivation for educational attainment. The General factor was consistently predictive of several criterion variables (student-reported school belonging and conduct problems, teacher-reported relationship warmth and conflict, disciplinary infractions, behavioral engagement in the classroom, and grades; Cham et al., 2014).

In addition to the General factor, there are three specific factors in the model[2], which are consistent with the above conceptualization: (a) Teacher Educational Expectation (including the item "My teachers believe that I will graduate from high school"), (b) Peer Aspirations (including the item "My friends expect me to graduate from high school"), and (c) Value Of Education (including the item "Graduating from high school is not as important to me as

[2]Cham et al. (2014) also investigated a full bifactor model in which items 1–4 and 6–12 loaded on an additional specific factor termed Competence and Effort Beliefs. However, the loadings of the items on the specific factor were small and in unpredicted directions and the specific factor accounted for only a trivial proportion of the unique variance. Consequently, this specific factor was deleted from the bifactor model as it was almost entirely redundant with the General factor (Chen, West, & Sousa, 2006).

getting a good paying job," which is reversed scored). The items reflect each of the specific factors in addition to the General factor. Each specific factor captures a unique effect over and above the General factor and the other specific factors. The criterion-related validity correlation of each specific factor with an external criterion variable represents the degree to which the specific factor predicts the criterion over and above the General factor and the other specific factors. Investigation of the criterion-related validities of the specific factors showed that the Teacher Educational Expectations factor uniquely predicted students' sense of school belonging, and the Peer Aspirations factor uniquely predicted teacher-reported grades and behavioral engagement (Cham et al., 2014).

Study Purpose

The primary purpose of the current study was to investigate the hypothesis that retention in elementary school impairs students' overall motivation in grade 9 to complete high school and pursue post-secondary education. We also investigated whether retention affected each of the specific factors associated with students' motivation (i.e., Teacher Educational Expectation, Peer Aspirations, and Value of Education). Grade 9 has been described as the "lynchpin year" for ultimate success in completing high school (Donegan, 2008). Motivation to complete high school is measured at the beginning of high school and can be considered as a proximal indicator of leaving school before graduation. By studying the effect of retention in the elementary grades on grade 9 students' motivation to complete high school and pursue post-secondary education, problems associated with the difficulty of tracking students who leave school prior to earning a high school diploma or graduate equivalency degree (GED) are minimized. Only a small percentage of students who eventually drop out of school do so prior to their first year in grade 9. In Texas, the location of the current study, only 0.04% of students drop out of school prior to grade 9, yet 16.10% of grade 9 students fail to earn a high school diploma or a GED within 4 years of entering grade 9 (Texas Education Agency, 2012). In Texas, barring special circumstances, it is illegal for students to drop out of school prior to age 18 or, if enrolled in a GED program, age 16.

A finding that retention in the elementary grades has a negative effect on grade 9 students' motivation for educational attainment would be consistent with Alexander et al.'s (2003) conclusion that "something about the experience [of grade retention] apparently weakened repeaters' attachment (or sense of belonging) to school." Conversely, a finding of no differences or positive effects of grade retention on grade 9 students' motivation would challenge the prevailing view that grade retention negatively impacts the likelihood of high school completion.

Method

Participants

Participants were drawn from a larger sample of academically at-risk students from three school districts (one urban and two small city districts) in Texas, when the students entered grade 1 in the fall of 2001 and 2002 (Hughes & Kwok, 2006). A total of 1,374 of grade 1 students in the three school districts met the following criteria for participation: scored below the median score on a state approved district-administered measure of literacy at

school entrance, spoke either English or Spanish, were not receiving special education services other than speech and language services in grade 1, and were not previously retained in grade 1. 1,200 returned consent forms were collected, of which 784 (65.33%) parents provided consent. No differences were indicated between the eligible students with and without parental consent across a broad array of archival variables, including performance on the district-administered test of literacy (standardized within district, due to differences in test used), age, gender, ethnicity, eligibility for free or reduced price lunch, bilingual class placement, cohort, and school context variables (i.e., % ethnic/racial minority; % economically disadvantaged). At the end of the 5th year of participation in the study, 569 (72.58%) parents provided written consent for continued participation. Almost all nonconsent was due to parental nonresponse. No material differences were indicated between the students with and without renewed parental consent in the analyses of the 67 measured covariates (potential confounders) in the propensity score analysis (see Measures section that follows). The absolute value of standardized mean difference (Cohen's *d*; Cohen, 1988) of the 45 continuous covariates between the students with and without renewed parent consent ranged from 0.00 to 0.48, with a median of 0.07. The Pearson correlations between the 22 binary and ordered covariates and students' renewed consent status ranged from .01 to .14, with a median of .07. Six students who were retained twice and two students who were retained after the transition to middle school were excluded from the current study.

The resulting sample consisted of 561 students (54.37% boys), including 177 students who were retained once and 384 students who were continuously promoted during grades 1 to 5. At entrance into the study, students' average age (in years) was 6.57 (*SD* = 0.38), with 35.29% of the sample being Caucasian, 24.24% African American, 36.36% Hispanic (38.73% of whom had limited English proficiency), and 4.11% other. 55.26% of the students were economically disadvantaged. Their mean full scale IQ based on the Universal Nonverbal Intelligence Test (Bracken & McCallum, 1998) was 93.10 (*SD* = 14.49), and their mean reading and math achievement age-standard scores from the Woodcock–Johnson III (Woodcock, McGrew, & Mather, 2001) or its Spanish-language equivalent (Woodcock–Muñoz Language Test; Woodcock & Muñoz-Sandoval, 1996) were 96.43 (*SD* = 17.67) and 101.48 (*SD* = 14.04), respectively.

Measures

Covariates for propensity score analysis—A total of 67 covariates (potential confounders), all of which were measured in grade 1 before retention occurred, were used to estimate the propensity scores of the retained and promoted students (Im et al., 2013). These 67 covariates were selected to be as comprehensive as possible, including variables that have been shown in prior research to be associated with grade retention or academic achievement. These variables were assessed with direct child testing and interviews (e.g., measures of IQ, effortful control, liking for school, language proficiency, and academic achievement); teacher questionnaires (e.g., behavioral, academic, and social functioning); parent questionnaires (e.g., family demographics, home–school relationship, and educational aspirations); peer sociometric testing (e.g., peer-rated acceptance and prosocial and aggressive behaviors); and school records (e.g., child ethnicity, age, and gender, classroom

ethnic composition, and bilingual class placement). If students or their parents spoke any Spanish in grade 1, students were administered the Woodcock–Muñoz Language Test (Woodcock & Muñoz-Sandoval, 1993) to determine the student's language proficiency in English and Spanish, and all student measures were administered in the language in which the student demonstrated greater proficiency. In grade 1, 55 students (9.80%) received the Spanish Woodcock–Muñoz Language Test. Spanish-speaking parents were sent questionnaires in both Spanish and English. A complete list of the covariates is presented in Appendix B.

Retention status—Students were considered retained in a given grade if they were in the same grade for two consecutive years. Schools provided information on students' grade placements every year.

Motivation for educational attainment—The 32-item measure developed by Cham et al. (2014) was used to measure adolescents' motivation for educational attainment (i.e., completion of high school and pursuit of post-secondary education). Students were asked to indicate the degree to which they agreed or disagreed with each statement (item) using a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*). Following Cham et al., the current study analyzed this outcome measure using the bifactor model, as shown in Figure 1. The conceptualization, interpretation, and criterion-related validity evidence of the General factor and the three specific factors (Teacher Educational Expectation, Peer Aspirations, and Value of Education) were presented in the Introduction. Cham et al. (2014) also reported that the bifactor model achieved measurement invariance across time (Years 9 and 10 of assessment), gender, and ethnicity groups (Caucasian, African American, and Hispanic). For the propensity score equated sample, reliability coefficients (coefficient ω; 1.00 defines perfect reliability; McDonald, 1999) of the General factor (ω = .92), Teacher Educational Expectation specific factor (ω = .80), Peer Aspirations specific factor (ω = .86), and Value of Education specific factor (ω = .85) in grade 9 were all satisfactory (see Reise, 2012). In the current study, we analyzed the factors rather than sum scores. Factors are theoretically assumed to be measurement error free.

Procedure

Of the eligible students (see Participants section for the eligibility criteria), teachers were asked to distribute consent forms to their parents via children's weekly folders. Teachers and parents were told that the purpose of the study was to learn more about factors that influence children's adjustment and success in school. Regardless of whether consent was granted, small gifts to children (e.g., erasers, fancy pencils) and the opportunity win a larger prize in a random drawing were instrumental obtaining a return of 1,200 consent forms. A total of 784 parents (65.33%) provided consent. At the end of the 5th year of participation in the study, parents provided written consent for continued participation. In addition to parental consent, student's assent to complete the questionnaires was obtained. Students were told the purpose of the questionnaire was to help the researcher learn more about factors that influence students' school performance, and they were assured of the confidentiality of their responses. Students received $25.00 for completing the questionnaires. Demographic information, such as student's age, gender, ethnicity, was obtained from school district

records. Teacher-reported and peer-reported data were collected annually beginning when all participants were in first grade. Parents and teachers each received $25 for completing and returning their set of the questionnaires. The research was approved by the Institutional Review Boards of Texas A&M University, Arizona State University, and each school district's research advisory team.

The items comprising the motivation for educational attainment scale were administered to students as part of a longer questionnaire. In order to examine the effect of retention in the elementary grades on students' motivation for educational attainment in the same grade (i.e., grade 9), the motivation items of the retained students measured at Year 10 and those of the promoted students measured at Year 9 were used. At Year 9, questionnaires were administered in individual sessions conducted at school. At Year 10, students were given the option of completing either an on-line or a paper version of the questionnaire. The purpose of offering the option of assessment procedures during Year 10 was to minimize attrition. We found that as children progressed into adolescence, an increasing proportion of students were unavailable to complete the paper version assessment, whereas these students were available for on-line version assessment. At Year 10, 18.54% of the students completed the on-line version and 81.46% completed the paper version. The means of the paper and on-line versions did not differ materially on any of the motivation items (median absolute value of Cohen's $d = 0.14$).

Results

Missing Data

Among the 67 covariates for the propensity score estimation, the median proportion of missing data was 9.98%. When the covariates contain missing values, the propensity score is defined as the variable that balances the completely observed values of the covariates as well as the missing data pattern of these covariates (Rosenbaum & Rubin, 1984). Among the 32 educational motivation items in grade 9, 75 students (11.30% retained and 14.32% promoted) had no observations on any item, whereas 5 students (0.56% retained and 1.04% promoted) did not have complete observations on all items. We discarded the 75 students without data plus the 5 students with incomplete data in the analysis of motivation for educational attainment (listwise deletion; see the description of the maximum likelihood mean-and-variance adjusted (MLMV) estimation in the section that follows). Additional analysis showed that there was no significant association between students' retention status in grade 1 and their missing data patterns on these items, phi coefficient = 0.03, $\chi^2(1) = 0.56$, $p = .45$.

Propensity Score Estimation

As discussed in the Introduction, the first step in propensity score analysis is to estimate each student's propensity score (conditional probability of being retained given the student's score on the covariates). The random forests method (Breiman, 2001) was used to estimate the propensity scores using the R package party Version 1.0–6 (Hothorn, Bühlmann, Dudoit, Molinaro, & Van Der Laan, 2006; Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008; Strobl, Boulesteix, Zeileis, & Hothorn, 2007). The random forests method automatically

identifies complex and nonlinear relations of covariates with students' grade retention status (retained and promoted). Compared with other methods of estimating propensity scores (e.g., logistic regression), the random forests method may lead to greater bias reduction in the estimate of the effect of grade retention on grade 9 motivation for educational attainment (Drake, 1993; Lee, Lessler, & Stuart, 2010). Following Rosenbaum and Rubin (1984), additional binary missingness indicators (= 1 when datum is observed and, = 0 when datum is missing) corresponding to incompletely observed covariates were included in the propensity score estimation model to adjust for missing data of the covariates. The model specification details for the random forests method are presented in Appendix C.

Propensity Score Equating

We used the weighting by the odds method (Hirano, Imbens, & Ridder, 2003) to equate the estimated propensity scores distribution between the retained and promoted students. In this method, each retained student was given a weight of 1.00, whereas each promoted student was given a weight of $\frac{\hat{\pi}}{1-\hat{\pi}}$, where $\hat{\pi}$ is the promoted student's estimated propensity score. The participants' propensity score weights are then accounted for using survey sampling weighting procedures to estimate the parameters and their standard errors (Asparouhov, 2005; Lohr, 2010). This weighting by the odds method estimates the grade retention effect for the students who were actually retained once in the elementary grades compared to closely equated students who were promoted (Schafer & Kang, 2008). Figure 2 shows the boxplots of propensity scores between the retained and promoted students before (i.e., pre-existing difference) and after equating. Before equating, as expected, the retained students had higher propensities to be retained whereas the promoted students had lower propensities to be retained. After equating, the propensity scores distributions were balanced in the retained and promoted groups.

Covariates' distributions balance—The performance of propensity score equating procedure was also assessed based on the balance of (1) the distributions of the completely observed values of the covariates and (2) the missing data pattern of the covariates between the retained and promoted students (Rosenbaum & Rubin, 1984; Stuart, 2010). We calculated the absolute standardized mean difference (SMD; for which 0 indicates perfect balance; Rubin, 2001; Stuart, 2010) and variance ratio (VR; for which 1.00 indicates perfect balance; Rubin, 2001) of the 67 covariates and each covariate's missing data pattern between the retained and promoted students before (i.e., pre-existing difference) and after propensity score equating. Table 1 shows the proportions of SMDs and VRs that fall within the ranges suggested by Hughes, Chen, Thoemmes, and Kwok (2010) and Rubin (2001), according to the covariate types (binary, ordinal, and continuous) plus the missing data pattern. After propensity score equating, higher proportions of covariates were classified into the categories of SMD close to the point of perfect balance (0). Before propensity score equating (pre-existing difference), a high proportion of the covariates were classified into the categories of VR close to the point of perfect balance (1.00). The results of VR of covariates did not materially change after propensity score equating. Table 2 shows the SMDs and VRs of 20 selected covariates that are highly correlated with the general motivation factor (without propensity score correction). The selected covariates cover a

broad range of dimensions (demographics, performance, behavior, motivation, social, personality, parent involvement, and home–school relationship) and sources (archival, performance, student-report, classmates-report, parent-report, and teacher-report). The results showed that, in general, these covariates had lower SMDs and VRs after propensity score equating. For the balance on the missing data patterns, a high proportion of the missing data patterns had SMDs and VRs that were close to the perfect balance before propensity score equating. After propensity score equating, a lower proportion of the missing data patterns had SMDs and VRs that were close to perfect balance. One potential reason is that the missing data rates of the covariates were not very high (median rate = 9.98%). Given the relatively small proportion of missing data, the propensity score equating procedures did not lead to improved balance on missingness.

We concluded that the weighting by the odds method together with the random forests propensity scores successfully balanced the distributions of propensity scores, distributions of the observed values of the covariates, and the missing data patterns of the covariates between the retained and promoted students. Therefore, we concluded that the weighting procedure successfully equated the retained and promoted students on the set of 67 covariates measured in grade 1.

Retention Effect on Motivation for Educational Attainment

Multi-group structural equation modeling was conducted to compare the bifactor educational motivation model between the retained and promoted students in grade 9 using M*plus* Version 7 (Muthén & Muthén, 1998–2012). In order to account for students' propensity score weights which were used in the weighting by the odds method, we used maximum likelihood MLMV estimation, which computes the χ^2 test statistic, the χ^2 difference test, and standard errors of parameter estimates that are appropriate for analyses involving weighting procedures (Asparouhov, 2005; Asparouhov & Muthén, 2006; 2010). MLMV estimation involves a correction to normal-theory maximum likelihood estimation; the χ^2 test statistic and the estimated standard errors of the parameter estimates are corrected for both the sampling weights (Asparouhov, 2005) and the non-normal distributions of the items (Bentler & Dudgeon, 1996). Compared with the alternative maximum likelihood mean adjusted (MLM) and maximum likelihood robust (MLR) estimations, Asparouhov (2005) showed that, when participants' sampling weights need to be accounted for, the MLMV estimation produces a χ^2 test statistic that has the expected Type I error rate when the model is correctly specified. MLMV is also robust to non-normality (Bentler & Dudgeon, 1996). The skewness (retained: range = −2.29 to 0.17, median = −0.86 and promoted: range = −2.69 to −0.03, median = −0.91) and excess kurtosis (retained: range = −0.93 to 7.90, median = 0.29 and promoted: range = −1.11 to 10.69, median = 0.43) of the motivation items differed from the values of 0 that characterize the normal distribution. Consequently, MLMV estimation was chosen. One limitation is that the MLMV estimator currently requires listwise deletion. Five students (0.56% retained and 1.04% promoted) who did not have complete observations on the 32 motivation items needed were discarded.

Measurement invariance of the bifactor model between the retained and promoted students is a prerequisite for investigating the effect of grade retention on educational motivation

(Millsap, 2011; Widaman & Reise, 1997). Table 3 shows the χ^2 test statistic, the RMSEA fit index with its 90% confidence interval, and the χ^2 difference test of the models that investigate the measurement invariance of factor pattern, factor loadings, latent intercepts, measurement error variances and covariances, and factor variances in the sequential order. The RMSEA of all the models suggested acceptable fit to the data. The null hypothesis of the χ^2 difference test is that the more restricted invariance model (e.g., same factor pattern + identical factor loadings) fits the data equally well as the less restricted invariance model (e.g., same factor pattern only). Although the χ^2 difference test should ideally be nonsignificant, all χ^2 difference tests were significant or marginally significant at $\alpha = .05$. We further investigated the modification indices; the modification indices did not suggest potential improvements could be made to these tested models (maximum value = 5.59). Based on the similar magnitudes of the RMSEA (i.e., no material change across all tested models; Cheung & Rensvold, 2002), the bifactor model of educational motivation appeared to have the same measurement structure in retained and promoted students.

We then compared the mean differences of each factor (General factor and specific factors of Teacher Educational Expectations, Peer Aspirations, and Value of Education) between the retained and promoted students in grade 9. The variance of the factors equals 1.0, so that mean differences can be interpreted as standardized effect sizes (Cohen's *d*). The retained students had lower scores on the General factor than the promoted students (mean difference = −0.25), but the difference was not statistically significant, $z = -0.81$, $p = .42$. The retained students had significantly higher scores on each of the three specific factors than the promoted students: (1) Teacher Educational Expectations, mean difference = 0.68, $z = 4.97$, $p < .001$; (2) Peer Aspirations, mean difference = 0.37, $z = 2.75$, $p = .006$; and (3) Value of Education, mean difference = 1.00, $z = 4.67$, $p < .001$. Table 4 shows the unstandardized factor loadings, factor means, and factor variances of the bifactor model for motivation for education attainment.

Two additional analyses were conducted to examine the robustness of the current findings. First, we conducted the sensitivity analysis proposed by Rosenbaum (1986; see also West et al., 2014) to investigate how the current results might be affected if there were a hidden confounder (or a composite of several hidden confounders). Following Rosenbaum (1986), a "worst case" scenario is assumed in which a hypothetical unmeasured covariate exists that has (a) a mean difference equal to the largest observed baseline difference of any covariate in the data set, and (b) this unobserved covariate correlates with the outcome variable to the same degree as the largest covariate-outcome correlation observed in the data set. Under this worst case scenario, the mean difference (Cohen's *d*) for the General factor, Teacher Educational Expectations, Peer Aspirations, and Value of Education changed to −0.16, 0.59, 0.28, and 0.90, respectively. We concluded that the current results were unlikely to be materially changed if a hidden confounder or a composite of several hidden confounders existed. Second, we also used logistic regression to estimate students' propensity scores. After propensity score equating, the results of the tests of the effect of retention on the general and three specific factors did not change materially, leading to identical conclusions to those obtained using the random forests procedure.

Discussion

In the current study, we conducted a rigorous propensity score analysis to control for the pre-retention differences between students who were (a) subsequently retained once versus (b) continuously promoted (never retained) in the elementary grades. Our propensity score weighting by the odds method successfully equated the retained and promoted students on a comprehensive set of 67 covariates measured in first grade prior to any retention. After successfully equating these two groups, we examined the effect of grade retention on students' motivation for educational attainment in grade 9. Grade 9 was selected based on the reasoning that it is the "lynchpin year" for ultimate success in completing high school (Donegan, 2008). We conducted same grade comparisons on our outcome variables measuring students' motivation for educational attainment (see note 1).

The current study found no evidence that retention in the elementary grades impairs students' general motivation for educational attainment in grade 9. Indeed, the results suggest that retention bestows benefits on three specific factors of motivation (Teacher Educational Expectations, Peer Aspirations, and Value of Education) relative to a carefully matched group of continuously promoted students who were at equal risk for retention when assessed in grade 1. In interpreting the findings of the study, it is important to understand the bifactor model of motivation for educational attainment. The General factor represents the construct that is measured in common by the entire set of items on the scale; it reflects each student's basic, overall motivation for educational attainment. Each of the three specific factors (Teacher Educational Expectations, Peer Aspirations, and Value of Education) represents the unique covariance in the set of items comprising the specific factor that exist over and above the other specific factors and the General factor. Thus, the effects of grade retention on the specific factors should be interpreted as the residual effects on each of the three specific constructs after they have been adjusted for the student's basic, overall motivation for educational attainment. Brunner, Nagy, and Wilhelm (2012) discussed the advantages of the bifactor model in its ability to represent measures in which there is interest in both a general construct and more specific facets. Over and above students' general motivation for educational attainment, retained students, relative to their propensity score matched, promoted peers, obtained higher scores on each of the three specific factors. As reported in Cham et al. (2014) with the current longitudinal dataset, in Year 9 the teacher educational expectations factor uniquely predicted students' sense of school belonging and the peer aspirations factor uniquely predicted teacher-reported grades and behavioral engagement.

Ten years into this longitudinal study of the effects of early grade retention on students' academic and psychosocial adjustment (Im et al., 2013; Moser et al., 2012; Wu et al., 2010), our results have not supported the popular view within the educational literature that grade retention harms students' educational success. Instead, we have either found advantages for the retained group or have failed to reject the null hypothesis of no difference between the retained and promoted groups. We have made comparisons between promoted and retained students on measures of psychological adjustment during the same measurement year (i.e., same age; see note 1 for the differences between same age and same grade comparisons). At year 4 of measurement (mean age = 10.57), students retained in first grade demonstrated

better psychosocial adjustment than did propensity-matched promoted peers (Wu et al., 2010). We have made comparisons between promoted and retained students in the same grade on education-related outcomes. At grade 5, no differences on a standardized measure of reading and math achievement between early retained and promoted students were detected (Moser et al., 2012), although retained students were one year older. Across the transition from elementary school to middle school, the expected negative effects of retention on measures of academic achievement, teacher-reported engagement, or school belonging were not found (Im et al., 2013). Now, at the critical point of the transition to high school, we found no ill effects of retention in the elementary grades on students' overall motivation to complete high school and pursue post-secondary education. Indeed, retained students relative to promoted matched peer controls who had entered grade 1 at the same time were somewhat more likely to perceive teachers as expecting them to succeed academically and that their peers had high educational expectations.

Comparison with Previous Research on Retention Effects

Our use of more adequate controls for differences between retained and promoted students on an extensive set of measured covariates that may be potential confounders of the grade retention effect may explain why our results diverge from those of other researchers (Alexander et al., 2003; Griffith, Loyd, Lane, & Tankersley, 2010; Jimerson, 1999). Cohort effects may also explain the differences in results. The current study was initiated in the midst of implementation of state-wide policies including high-stakes testing beginning in grade 3 (Texas Education Agency, 2007) that strongly discouraged social promotion; such policies were not in effect at the time previous research was conducted.[3] Thus, retention decisions are made in a context that may differ in important ways from the context of earlier studies. For example, when retention decisions are driven primarily by failure to pass a test of grade-level competencies rather than more subjective considerations documented in earlier research (Jimerson, Carlson, Rotert, Egeland, & Sroufe, 1997; McCoy & Reynolds, 1999), retention in grade may be more beneficial (Chen, Hughes, & Kwok, 2014). Demographic differences may also contribute to differences in results. Many of the previous prospective, longitudinal studies were conducted in the 1980s and 1990s in urban school districts serving predominantly poor African American students (Alexander et al., 2003; Temple, Reynolds, & Meidel, 2000), whereas the present study included a balance of ethnic groups and urban and smaller city schools.

From a statistical point of view, the theory of propensity scores requires that all pre-treatment confounders be identified (Rosenbaum & Rubin, 1983) to eliminate bias. Rubin (2001) and West et al. (2014) suggested that researchers attempt to identify all potential pre-treatment confounders for inclusion in the propensity score analysis. A comprehensive set of potential confounders can minimize the any bias in the estimate of the average treatment effect. This result occurs because the comprehensive covariate set reduces the chances that other unmeasured confounders that provide unique contributions to the confounding effect, over and above of those of the measured potential confounders (Cook, Steiner, & Pohl, 2009; Steiner, Cook, Shadish, & Clark, 2010).

[3]Griffith et al. (2010) used NELS data for participants who were in grade 8 in 1988.

Study Limitations

Study findings need to be interpreted in light of limitations of our measure of motivation for educational attainment. First, although students themselves are the best reporters of their beliefs about their own competence and valuing of education, students' perceptions of the social context (i.e., peers' educational aspirations and teachers' and parents' competence beliefs and educational expectations) may be biased relative to the perspective of other reporters. Nevertheless, according to motivational theorists (Deci & Ryan, 1991; Wigfield & Eccles, 2000), a student's perception of social support from others, whether congruent or not with other sources, is expected to affect students' engagement in learning and motivation for educational attainment, over and above other reports (Hughes, 2011). Second, Cham et al. (2014) noted that the measure had relatively few parent support items and so that this facet of motivation for educational attainment may not have been fully represented. Previous research has suggested that adolescents' perceptions of parents' warmth and closeness may be important to students' commitment to educational attainment, via their impact on students' valuing of educational attainment (Legault et al., 2006).

Despite the finding of no ill effects of retention on students' motivation for educational attainment in grade 9, we cannot definitively rule out, in the present study, the possibility that students retained in grade will be less likely to actually complete high school. Although we carefully equated retained and promoted students on a large set of covariates assessed in grade 1, the intervention of retention itself leads to confounding on a key covariate: age. Retained students are one year older than comparable promoted students when they enter grade 9. When the retained students reach the legal age to leave school or to work, they are one year further away from completing high school than they would have been if they had been continuously promoted. At that time point for these low achieving students, the cost–benefit analysis of continuing an extra year in school may not be as positive as that of the comparable continuously promoted student who does not face this extra year. This personal calculus may increase the probability that retained students will drop out of school. If this is the case, retention may impair students' graduation rates, even if does not harm their commitment for educational attainment at the point at which they embark on their high school career. Additional waves of data are necessary to answer this ultimate question.

Implications of Findings for Research and Practice

These findings have implications for future research on the effect of grade retention on completing high school and pursuing post-secondary education. Foremost is the demonstrated efficacy of propensity score weighting as a method of adjusting for the potential confounders associated with both grade retention and motivation for educational attainment, thereby providing a less biased estimate of the effect of grade retention. After propensity score weighting, retained and promoted students were virtually identical on 67 pre-retention covariates (potential confounders) that are potentially associated with both grade retention and motivation for educational attainment in grade 9. This level of statistical control for potential confounders was possibly due to the assessment at baseline of a large number of factors associated with academic risk in prior research, going beyond demographic and academic achievement to include peer, teacher, and parent ratings of children's behavioral regulation and social competence, teacher–student relationships,

parent school involvement, the home-school relationship, and classroom instructional resources. Based on developmental systems theory (Lerner, 1998), small differences between children on these variables at school entrance may have an out-sized difference on subsequent school performance via developmental cascades (Masten et al., 2005). Studies on the effect of grade retention on subsequent achievement and educational attainment that do not assess such a comprehensive set of potential confounders prior to grade retention are likely to provide biased estimates of the effects of grade retention. Is retention helpful to students? Unfortunately, this study, nor any single study, can answer that ultimate question. It is important that results from this longitudinal program of research be replicated with samples of students representing diverse geographical regions of the country as well as rural and urban communities. Future research on potentially modifiable school factors that may moderate the effect of retention on subsequent academic performance is also recommended. For example, grade retention effects may differ based on the decision-making process that determines who is retained, the level and type of instructional resources provided to students during their repeat year, or the proportion of students within a school who are retained in a given year. Particularly informative would be studies of differences within samples of retained students that moderate the success of grade retention, so that this intervention can be employed with more precision. In the meantime, we echo the recommendation of Reschly and Christenson (2013) that the academic progress of all learners should be carefully monitored, and evidence-based interventions that address children's specific learning need be employed when students begin to struggle academically. Although grade retention may not have the negative effects so widely assumed in the published literature, it is an expensive intervention with minimal evidence of benefits to the retained student. On the other hand, perpetuating negative expectations for retention effects that are not supported by contemporary, rigorous studies does not serve the interests of the children and families we serve.

Acknowledgments

This research was supported by a grant from the National Institute of Child Health and Human Development (grant number HD039367) to Jan N. Hughes.

References

Alet E. Is grade repetition a second chance? 2010 from http://www.seg.guanajuato.gob.mx/Ceducativa/CartillaB/6antologia/antecedentes/pdf/14.-%20IS%20GRADE%20REPETITION%20A%20SECOND%20CHANCE.pdf.

Alexander, KL.; Entwisle, DR.; Dauber, SL. On the success of failure: A reassessment of the effects of retention in the primary grades. 2nd ed.. Cambridge, UK: Cambridge University Press; 2003.

Alexander KL, Entwisle DR, Horsey CS. From first grade forward: Early foundations of high school dropout. Sociology of Education. 1997; 70:87–107.

Alexander KL, Entwisle DR, Kabbani N. The dropout process in life course perspective: Early risk factors at home and school. The Teachers College Record. 2001; 103:760–822.

Alivernini F, Lucidi F. Relationship between social context, self-efficacy, motivation, academic achievement, and intention to drop out of high school: A longitudinal study. The Journal of Educational Research. 2011; 104:241–252.

Allen CS, Chen Q, Willson VL, Hughes JN. Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. Educational Evaluation and Policy Analysis. 2009; 31:480–499. [PubMed: 20717492]

Alliance for Excellence in Education. The high cost of high school dropouts. Washington DC: Author; 2007. from http://www.all4ed.org/files/archive/publications/HighCost.pdf [Retrieved June 18, 2013]

Altermatt E, Pomerantz EM. The implications of having high-achieving versus low-achieving friends: A longitudinal analysis. Social Development. 2005; 14:61–81.

Asparouhov T. Sampling weights in latent variable modeling. Structural Equation Modeling. 2005; 12:411–434.

Asparouhov T, Muthen B. Robust chi square difference testing with mean and variance adjusted test statistics. 2006 (Mplus Web Notes: No. 10) from http://statmodel.com/download/webnotes/webnote10.pdf.

Asparouhov T, Muthen B. Simple second order chi-square correction. 2010 May 3. from http://statmodel.com/download/WLSMV_new_chi21.pdf.

Bandura A, Barbaranelli C, Caprara GV, Pastorelli C. Multifaceted impact of self-efficacy beliefs on academic functioning. Child Development. 1996; 67:1206–1222. [PubMed: 8706518]

Barrington BL, Hendricks B. Differentiating characteristics of high school graduates, dropouts, and nongraduates. The Journal of Educational Research. 1989; 82:309–319.

Battin-Pearson S, Newcomb MD, Abbott RD, Hill KG, Catalano RF, Hawkins JD. Predictors of early high school dropout: A test of five theories. Journal of Educational Psychology. 2000; 92:568–582.

Benner AD. Exit examinations, peer academic climate, and adolescents' developmental outcomes. Journal of School Psychology. 2012; 51:67–80. [PubMed: 23375173]

Bentler PM, Dudgeon P. Covariance structure analysis: Statistical practice, theory, and directions. Annual Review of Psychology. 1996; 47:563–592.

Bowers AJ, Sprott R. Examining the multiple trajectories associated with dropping out of high school: A growth mixture model analysis. The Journal of Educational Research. 2012; 105:176–195.

Bracken, BA.; McCallum, RS. Universal nonverbal intelligence test (UNIT). Itasca, IL: Riverside Publishing; 1998.

Breiman L. Random forests. Machine Learning. 2001; 45:5–32.

Brunner M, Nagy G, Wilhelm O. A tutorial on hierarchically structured constructs. Journal of Personality. 2012; 80:796–846. [PubMed: 22091867]

Caprara GV, Fida R, Vecchione M, Del Bove G, Vecchio GM, Barbaranelli C, Bandura A. Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. Journal of Educational Psychology. 2008; 100:525–534.

Cham H, Hughes JN, West SG, Im MH. Assessment of adolescent's motivation for educational attainment. Psychological Assessment. 2014

Chapman, C.; Laird, J.; KewalRamani, A. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education; 2011. Trends in high school dropout and completion rates in the United States: 1972–2009. (NCES 2012-006) from http://nces.ed.gov/pubsearch. [Retrieved November 18, 2011]

Chen Q, Hughes JN, Kwok O. Differential growth trajectories for achievement among children retained in first grade: A growth mixture model. Elementary School Journal. 2014; 114:327–353. [PubMed: 24771882]

Cheung GW, Rensvold RB. Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling. 2002; 9:233–255.

Cohen, J. Statistical power for the behavioral sciences. 2nd ed.. Hillsdale, NJ: Erlbaum; 1988.

Cook TD, Steiner PM, Pohl S. Assessing how bias reduction is influenced by covariate choice, unreliability, and data analytic model: An analysis of different kinds of within-study comparisons in different substantive domains. Multivariate Behavioral Research. 2009; 44:828–847.

Cox MJ, Mills-Koonce R, Propper C, Gariépy JL. Systems theory and cascades in developmental psychopathology. Development and Psychopathology. 2010; 22:497–506. [PubMed: 20576174]

DeMaris A. Combatting unmeasured heterogeneity in cross-sectional studies: Evaluating instrumental-variable and Heckman-selection models. Psychological Methods. (in press).

Deci, EL.; Ryan, RM. A motivational approach to self: Integration in personality. In: Dienstbier, RA., editor. Nebraska symposium on motivation, 1990. Lincoln, NE: University of Nebraska Press; 1991. p. 237-288.

Deci EL, Ryan RM. The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry. 2000; 11:227–268.

Dong Y. Kept back to get ahead? Kindergarten retention and academic performance. European Economic Review. 2010; 54:219–236.

Donegan B. The Linchpin year. Educational Leadership. 2008; 65(8):54–56.

Drake C. Effects of misspecification of the propensity score on estimators of treatment effect. Biometrics. 1993; 49:1231–1236.

Eccles JS, Wigfield A. In the mind of the actor: The structure of adolescents' academic achievement related-beliefs and self-perceptions. Personality and Social Psychology Bulletin. 1995; 21:215–225.

Elder GH. The life course as developmental theory. Child Development. 1998; 69:1–12. [PubMed: 9499552]

Englund MM, Egeland B, Collins WA. Exceptions to high school dropout predictions in a low-income sample: Do adults make a difference? Journal of Social Issues. 2008; 64:77–94. [PubMed: 19779583]

Goos M, Van Damme J, Onghena P, Petry K, de Bilde J. First-grade retention in the Flemish educational context: Effects on children's academic growth, psychosocial growth, and school career throughout primary education. Journal of School Psychology. 2013; 51:323–347. [PubMed: 23816228]

Griffith CA, Loyd JW, Lane KL, Tankersley M. Grade retention of students during grades K-8 predicts reading achievement and progress during secondary schooling. Reading and Writing Quarterly. 2010; 26:51–66.

Grissom, JB.; Shepard, LA. Repeating and dropping out of school. In: Sheppard, LA.; Smith, ML., editors. Flunking grades: Research and policies on retention. New York, NY: The Falmer Press; 1989. p. 34-63.

Hill NE, Tyson DF. Parental involvement in middle school: a meta-analytic assessment of the strategies that promote achievement. Developmental Psychology. 2009; 45:740–763. [PubMed: 19413429]

Hirano K, Imbens GW, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. Econometrica. 2003; 71:1161–1189.

Holmes, CT. Grade-level retention effects: A meta-analysis of research studies. In: Shepard, LA.; Smith, ML., editors. Flunking grades: Research and policies on retention. London, England: The Falmer Press; 1989. p. 16-33.

Hong S, Ho HZ. Direct and indirect longitudinal effects of parental involvement on student achievement: Second-order latent growth modeling across ethnic groups. Journal of Educational Psychology. 2005; 97:32–42.

Hothorn T, Bühlmann P, Dudoit S, Molinaro A, Van Der Laan MJ. Survival ensembles. Biostatistics. 2006; 7:355–373. [PubMed: 16344280]

Hughes JN. Longitudinal effects of teacher and student perceptions of teacher-student relationship qualities on academic adjustment. Elementary School Journal. 2011; 112:38–60. [PubMed: 21984843]

Hughes JN, Chen Q, Thoemmes F, Kwok OM. An investigation of the relationship between retention in first grade and performance on high stakes tests in third grade. Educational Evaluation and Policy Analysis. 2010; 32:166–182. [PubMed: 20628547]

Hughes JN, Kwok O. Classroom engagement mediates the effect of teacher-student support on elementary students' peer acceptance: A prospective analysis. Journal of School Psychology. 2006; 43:465–480. [PubMed: 20431706]

Hughes JN, Kwok O, Im MH. Effect of retention in first grade on parents' educational expectations and children's academic outcomes. American Educational Research Journal. 2013; 50:1336–1359.

Hughes JN, Luo W, Kwok O, Loyd L. Teacher–student support, effortful engagement, and achievement: A three-year longitudinal study. Journal of Educational Psychology. 2008; 100:1–14. [PubMed: 19578558]

Im MH, Hughes JN, Kwok OM, Puckett S, Cerda CA. Effect of retention in elementary grades on transition to middle school. Journal of School Psychology. 2013; 51:349–365. [PubMed: 23816229]

Janosz M, Archambault I, Morizot J, Pagani LS. School engagement trajectories and their differential predictive relations to dropout. Journal of Social Issues. 2008; 64:21–40.

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.

Karweit, NL. Grade retention: Prevalence, timing, and effects. Baltimore, MD: The Johns Hopkins University, Center for Social Organization of Schools; 1999. (CRESPAR Report 33).

Kindermann TA. Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. Child Development. 2007; 78:1186–1203. [PubMed: 17650133]

Kiuru N, Aunola K, Vuori J, Nurmi JE. The role of peer groups in adolescents' educational expectations and adjustment. Journal of Youth and Adolescence. 2007; 36:995–1009.

Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. Statistics in Medicine. 2010; 29:337–346. [PubMed: 19960510]

Legault L, Green-Demers I, Pelletier LG. Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. Journal of Educational Psychology. 2006; 98:567–582.

Lerner, RM. Theories of human development: Contemporary perspectives. In: Damon, W.; Lerner, RM., editors. Handbook of child psychology. 5th ed.. Vol. 1. New York, NY: Wiley; 1998. p. 1-24.(Series Ed.) (Vol. Ed.)

Lohr, S. Sampling: Design and analysis. 2nd ed.. Boston, MA: Brooks/Cole, Cengage Learning; 2010.

Masten AS, Roisman GI, Long JD, Burt KB, Obradovic J, Riley JR, Boelcke-Stennes K, Tellegan A. Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. Developmental Psychology. 2005; 41:733–746. [PubMed: 16173871]

McCoy AR, Reynolds AJ. Grade retention and school performance: An extended investigation. Journal of School Psychology. 1999; 37:273–298.

McDonald, RP. Test theory: A unified treatment. Mahwah, NJ: Erlbaum; 1999.

Millsap, RE. Statistical approaches to measurement invariance. New York, NY: Routledge; 2011.

Morgan, SL.; Winship, C. Counterfactuals and causal inference: Methods and principles for social research. New York, NY: Cambridge University Press; 2007.

Moser SE, West SG, Hughes JN. Trajectories of math and reading achievement in low-achieving children in elementary school: Effects of early and later retention in grade. Journal of Educational Psychology. 2012; 104:603–621. [PubMed: 23335818]

Murdock TB. The social context of risk: Status and motivational predictors of alienation in middle school. Journal of Educational Psychology. 1999; 91:62–75.

Muthén, LK.; Muthén, BO. Mplus user's guide. 7th ed.. Los Angeles, CA: Muthén & Muthén; 1998–2012.

Newcomb MD, Abbott RD, Catalano RF, Hawkins DJ, Battin-Pearson S, Hill K. Mediational and deviance theories of late high school failure: process roles of structural strains, academic

competence, and general versus specific problem behavior. Journal of Counseling Psychology. 2002; 49:172–186.

Pintrich PR. A motivational science perspective on the role of student motivation in learning and teaching contexts. Journal of Educational Psychology. 2003; 95:667–686.

Pleis, JR.; Ward, BW.; Lucas, JW. Summary health statistics for U.S. adults: National Health Interview Survey, 2009. National Center for Health Statistics; 2010. Vital Health Stat 10(249).

Reise SP. The rediscovery of bifactor measurement models. Multivariate Behavioral Research. 2012; 47:667–696. [PubMed: 24049214]

Reschly AL, Christenson SL. Grade retention: Historical perspectives and new research. Journal of School Psychology. 2013; 51:319–322. [PubMed: 23816227]

Roderick M. Grade retention and school dropout: Investigating the association. American Educational Research Journal. 1994; 31:729–759.

Rosenbaum PR. Dropping out of high school in the United States: An observational study. Journal of Educational Statistics. 1986; 11:207–224.

Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.

Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. Journal of the American Statistical Association. 1984; 79:516–524.

Rubin DB. Using propensity scores to help design observational studies: application to the tobacco litigation. Health Services and Outcomes Research Methodology. 2001; 2:169–188.

Rumberger RW. High school dropouts: A review of issues and evidence. Review of Educational Research. 1987; 57:101–121.

Schafer JL, Kang J. Average causal effects from nonrandomized studies: A practical guide and simulated example. Psychological Methods. 2008; 13:279–313. [PubMed: 19071996]

Sipple JW, Killeen K, Monk DH. Adoption and adaptation: School district responses to state imposed learning and graduation requirements. Educational Evaluation and Policy Analysis. 2004; 26:143–168.

Steiner PM, Cook TD, Shadish WR, Clark MH. The importance of covariate selection in controlling for selection bias in observational studies. Psychological Methods. 2010; 15:250–267. [PubMed: 20822251]

Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. BMC Bioinformatics. 2008; 9

Strobl C, Boulesteix AL, Zeileis A, Hothorn T. Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics. 2007; 8

Stuart EA. Matching methods for causal inference: A review and a look forward. Statistical Science. 2010; 25:1–21. [PubMed: 20871802]

Temple JA, Reynolds AJ, Meidel WT. Can early intervention prevent high school dropout? Evidence from the Chicago Child-Parent Centers. Urban Education. 2000; 35:31–56.

Texas Education Agency. Grade-level retention in Texas public schools, 2005–06. 2007 (Document No. GE08 601 01). from http://www.tea.state.tx.us/acctres/Retention_2005-06.pdf.

Texas Education Agency. Secondary School Completion and Dropouts in Texas Public Schools, 2010–11. 2012 (Document No. GE12 601 06). from http://www.tea.state.tx.us/acctres/DropComp_2010-11.pdf.

Tynkkynen L, Tolvanen A, Salmela-Aro K. Trajectories of educational expectations from adolescence to young adulthood in Finland. Developmental Psychology. 2012; 48:1674–1685. [PubMed: 22329386]

Véronneau MH, Dishion TJ. Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. The Journal of Early Adolescence. 2011; 31:99–124. [PubMed: 21552353]

Wang M, Fredricks JA. The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. Child Development. 2014; 85:722–737. [PubMed: 23895361]

West SG, Cham H, Thoemmes F, Renneberg B, Schulze J, Weiler M. Propensity scores as a basis for equating groups: Basic principles and application in clinical treatment outcome research. Journal of Consulting and Clinical Psychology. 2014

Widaman, KF.; Reise, SP. Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In: Bryant, KJ.; Windle, M.; West, SG., editors. The science of prevention: Methodological advances from alcohol and substance abuse research. Washington, D. C.: American Psychological Association; 1997. p. 281-324.

Wigfield, A.; Cambria, J.; Eccles, JS. Motivation in education. In: Ryan, RM., editor. The Oxford handbook of human motivation. New York, NY: Oxford University Press; 2012. p. 463-478.

Wigfield A, Eccles JS. Expectancy–value theory of achievement motivation. Contemporary Educational Psychology. 2000; 25:68–81. [PubMed: 10620382]

Willson VL, Hughes JN. Retention of Hispanic/Latino students in first grade: Child, parent, teacher, school, and peer predictors. Journal of School Psychology. 2006; 44:31–49. [PubMed: 20419036]

Woodcock, RW.; McGrew, KS.; Mather, N. Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing; 2001.

Woodcock, RW.; Muñoz-Sandoval, AF. Woodcock–Muñoz Language Survey. Riverside, CA: Riverside Publishing; 1993.

Woodcock, RW.; Muñoz-Sandoval, AF. Batería Woodcock–Muñoz: Pruebas de aprovechamiento—Revisada. Itasca, IL: Riverside Publishing; 1996.

Wu W, West SG, Hughes JN. Effect of retention in first grade on children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. Journal of Educational Psychology. 2008; 100:727–740. [PubMed: 19337582]

Wu W, West SG, Hughes JN. Effect of grade retention in first grade on psychosocial outcomes. Journal of Educational Psychology. 2010; 102:135–152. [PubMed: 20448829]

Zimmerman, BJ.; Cleary, TJ. Adolescents' development of personal agency: The role of self-efficacy beliefs and self-regulatory skill. In: Pajares, F.; Urdan, T., editors. Self-efficacy beliefs of adolescents. Greenwich, CT: Information Age; 2006. p. 45-69.

Appendix A

Adolescent Motivation for Educational Attainment Questionnaire Items

Items		
1		I am confident that I will graduate from high school.
2		I know what courses I need to take to graduate from high school.
3		I am on track to graduate from high school.
4		Nothing will get in the way of my graduating from high school.
5	(R)	Graduating from high school is not as important to me as getting a good paying job.
6		My parents expect me to graduate from high school.
7		I am confident that I will go to college.
8		My parents expect me to go to college.
9		Nothing will get in the way of my going to college.
10		I have started gathering information about vocational schools or colleges.
11		I know what courses and grades it takes to get into the vocational school or college I want to enter.
12		I know how much it costs to go to vocational school or college.
13		My teachers expect that I will do well in the future.
14		I am one of the students teachers believe will be successful.
15		My teachers believe that I will graduate from high school.
16		My teachers consider me to be pretty smart.

Items		
17	(R)	My teachers don't think I'll go to college.
18	(R)	Most of my good friends will quit high school when they are old enough.
19		Most of my good friends plan to go to college.
20		Most of my good friends won't drop out.
21		Most of my good friends will get a high school diploma.
22	(R)	I don't think many of my friends will graduate from high school.
23	(R)	Lots of my good friends won't be able to go to college.
24	(R)	I don't think education will guarantee that I get paid well.
25	(R)	I can make good money without an education.
26	(R)	Many of the things we do in school seem useless to me.
27	(R)	If I get bad grades, I can still get a good job.
28	(R)	I could be successful in life without an education.
29	(R)	I know many people who have done well in life with little education.
30	(R)	School is not that important for future success.
31		I will make more money someday if I do well in school.
32		If I work hard in school, I will get a better job than the kids who don't try hard.

Note. (R) indicates reverse-scored item. Students were asked to indicate the degree to which they agreed or disagreed with each of the statements using a 5-point scale (1 = *strongly disagree*; 5 = *strongly agree*).

Appendix B

List of Covariates for Propensity Score Analysis

Covariates	Source
1. Student's ethnicity (African-American vs. other ethnic groups)	Archival
2. Student's gender	Archival
3. Student's age at eligibility determination	Archival
4. Student's limited English proficiency status	Archival
5. Student's bilingual class status	Archival
6. Student's English as a second language status	Archival
7. Student's tested language (English vs. Spanish)	Archival
8. Student's enrollment in pre-first-grade	Archival
9. Percentage of Caucasians in the same class	Archival
10. Student's composite district literacy test score	Performance
11. Student's IQ score (UNIT full scale)	Performance
12. Student's Woodcock-Johnson III broad reading standard score	Performance
13. Student's Woodcock-Johnson III broad math standard score	Performance
14. Student's effortful (inhibitory) control	Performance
15. Student's Dweck puzzles task choice	Performance
16. Student's academic self-efficacy	Child
17. Student's sense of school belonging	Child
18. Student's trouble (peer-report)	Classmates

Covariates	Source
19. Student's aggression(peer-report)	Classmates
20. Student's prosocial behaviors (peer-report)	Classmates
21. Student's ADHD behaviors (peer-report)	Classmates
22. Student's sad/withdrawn (peer-report)	Classmates
23. Student's teacher support (peer-report)	Classmates
24. Student's liking by peers (peer-report)	Classmates
25. Student's social preference by peers (peer-report)	Classmates
26. Number of children (under age 18) living in household	Parent
27. Household employment status	Parent
28. Household highest level of education	Parent
29. Family adversity	Parent
30. Family social economic status	Parent
31. Student's enrollment in kindergarten	Parent
32. Parent's sense of responsibility for child's education	Parent
33. Parent's perceived teachers responsibility for child's learning and behavior	Parent
34. Parent's self-efficacy for helping child in school	Parent
35. Parent's positive perceptions about school	Parent
36. Parent's satisfaction with home-school communication	Parent
37. Parent's perceived parent-teacher shared responsibilities	Parent
38. Parent's perceived school-based involvement	Parent
39. Student's ADHD behaviors (parent-report)	Parent
40. Student's prosocial behaviors (parent-report)	Parent
41. Student's conduct problems (parent-report)	Parent
42. Student's internalizing behaviors (parent-report)	Parent
43. Student's status receiving teacher instruction in reduced class size	Teacher
44. Student's status receiving 1-1 tutoring by an adult	Teacher
45. Student's status receiving 1-1 tutoring by a peer or older student	Teacher
46. Student's status receiving remedial instruction outside classroom	Teacher
47. Student's status receiving instruction with an aider	Teacher
48. Student's status receiving remedial instruction before / after school	Teacher
49. Student's status receiving 1-1 tutoring by an adult before / after school	Teacher
50. Student's status receiving individual counseling	Teacher
51. Student's status receiving speech therapy	Teacher
52. Student's status receiving small group intensive tutoring	Teacher
53. Student's ego control (teacher-report)	Teacher
54. Student's agreeableness personality (teacher-report)	Teacher
55. Student's conscientiousness personality (teacher-report)	Teacher
56. Student's ADHD behaviors (teacher-report)	Teacher
57. Student's prosocial behaviors (teacher-report)	Teacher
58. Student's conduct problems (teacher-report)	Teacher
59. Student's internalizing behaviors (teacher-report)	Teacher
60. Teacher's educational aspirations for student	Teacher

Covariates	Source
61. Student's achievement (teacher-report)	Teacher
62. Student's school engagement (teacher-report)	Teacher
63. Home-school relationship (alliance)	Teacher
64. Home-school relationship (parent's involvement)	Teacher
65. Home-school relationship (teacher's initiation)	Teacher
66. Teacher-student conflict (teacher-report)	Teacher
67. Teacher-student support (teacher-report)	Teacher

Appendix C

Model Specification for the Random Forests Method

Number of Random Samples	1000
Percentage of Students of Random Sample to that of Original Data	63.20%
Sampling with Replacement to Draw Random Samples	No
Test Statistic Selecting the Covariate and the Split Value	Quadratic Form
Adjustment to Multiple Testing of Covariates	No Adjustment
Number of Covariates to Randomly Sample in Each Selection of Covariate	All Covariates
Minimum Number of Students in Each Intermediate Node	100
Minimum Number of Students in Each Terminal Node	1
Maximum Depth of Classification Tree Model	No control
Data Used to Calculate Propensity Scores Using Each Classification Tree	Out-of-bag Data

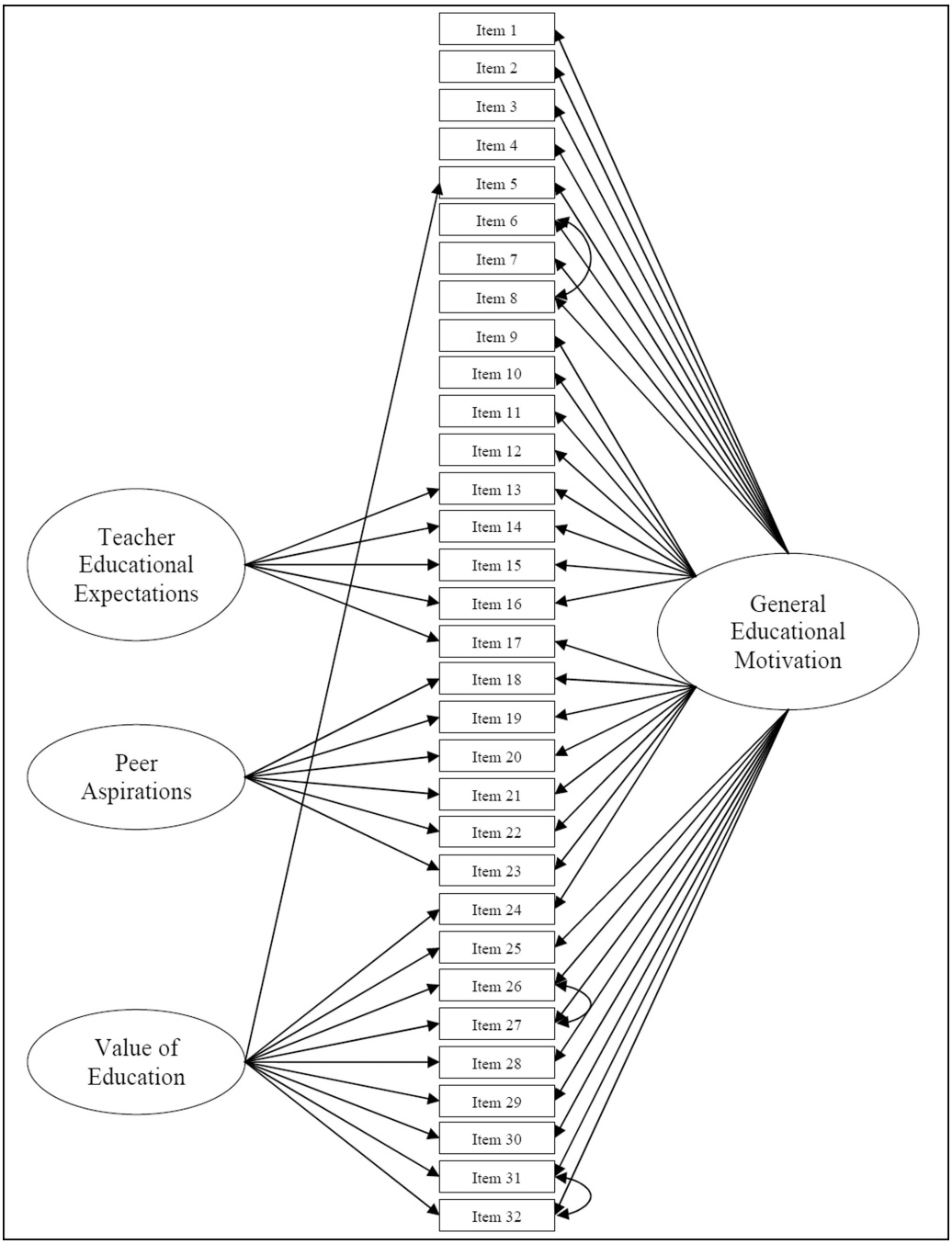

Figure 1.

Path diagram of bifactor model for Adolescents' Motivation for Educational Attainment. The ellipses are the factors and the rectangles are the measured items. The double-headed arrows between pairs of items are the correlated uniquenesses. See Appendix A for specific items.

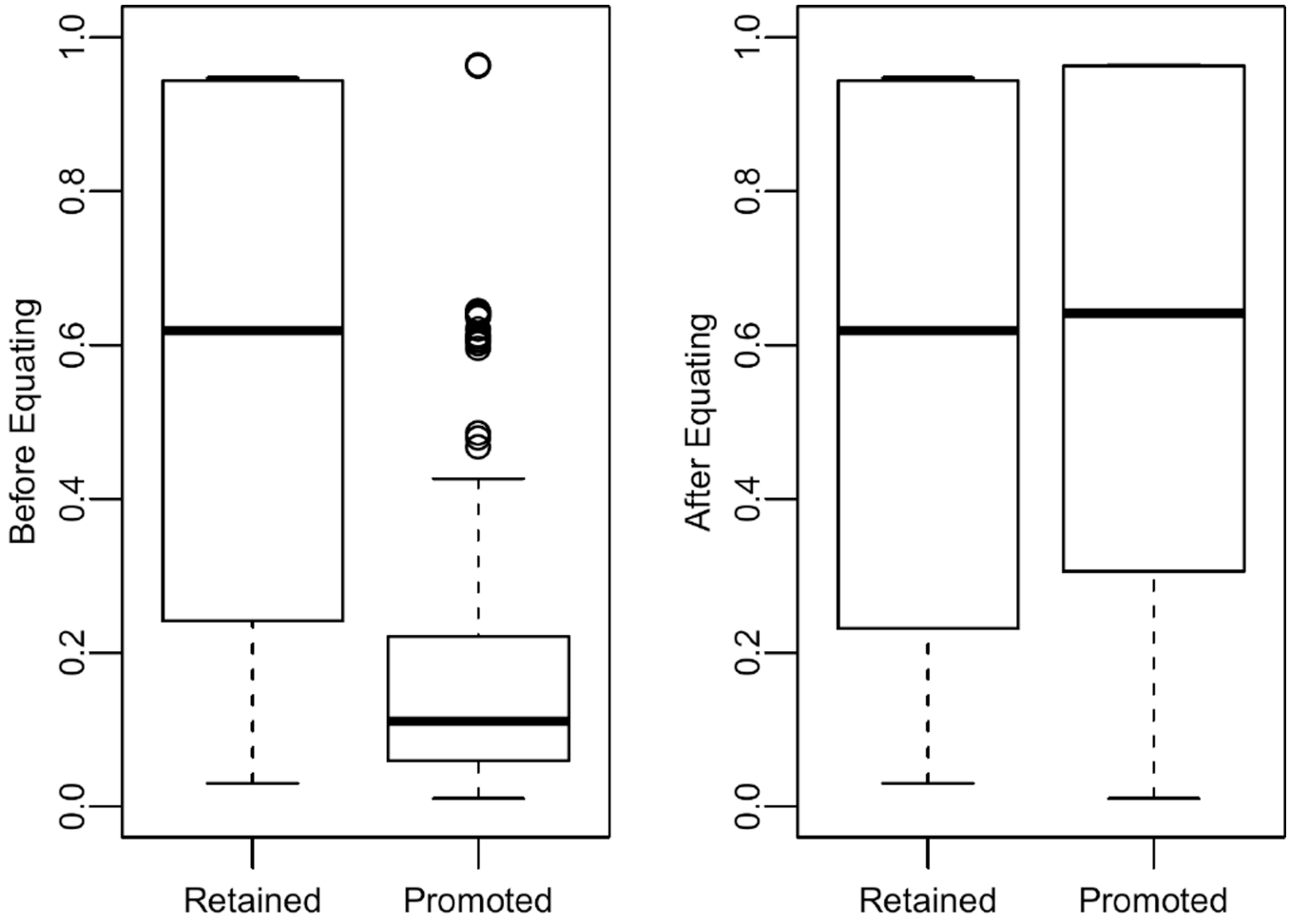

Figure 2.

Boxplots of propensity scores between retained and promoted students before and after propensity score equating using the weighting by the odds method.

Table 1

Absolute Standardized Mean Difference and Variance Ratio of Covariates and Missing Data Pattern

Absolute Standardized Mean Difference	Binary Covariate		Ordinal Covariate		Continuous Covariate		Missing Data Pattern	
	Before	After	Before	After	Before	After	Before	After
SMD ≤ 0.05	26.32%	21.05%	0.00%	0.00%	6.67%	15.56%	31.82%	2.27%
0.050 < SMD ≤ 0.100	10.53%	21.05%	0.00%	0.00%	8.89%	11.11%	31.82%	6.82%
0.100 < SMD ≤ 0.125	5.26%	10.53%	0.00%	33.33%	4.44%	11.11%	20.45%	2.27%
0.125 < SMD ≤ 0.175	15.79%	10.53%	0.00%	0.00%	11.11%	15.56%	6.82%	0.00%
SMD > 0.175	42.11%	36.84%	100.00%	66.67%	68.89%	46.67%	9.09%	88.64%
Variance Ratio	**Binary Covariate**		**Ordinal Covariate**		**Continuous Covariate**		**Missing Data Pattern**	
	Before	After	Before	After	Before	After	Before	After
VR ≤ 0.50	0.00%	15.79%	0.00%	0.00%	0.00%	4.44%	0.00%	34.09%
0.50 < VR ≤ 0.80	21.05%	15.79%	0.00%	33.33%	11.11%	28.89%	27.27%	27.27%
0.80 < VR ≤ 1.25	47.37%	52.63%	66.67%	33.33%	57.78%	44.44%	47.73%	18.18%
1.25 < VR ≤ 2.00	31.58%	15.79%	33.33%	33.33%	28.89%	22.22%	11.36%	15.91%
VR > 2.00	0.00%	0.00%	0.00%	0.00%	2.22%	0.00%	13.64%	4.55%

Note. SMD is absolute standardized mean difference. VR is variance ratio. Before is before propensity score equating (pre-existing difference). After is after propensity score equating using the weighting by the odds method.

Table 2

Absolute Standardized Mean Difference and Variance Ratio of Selected Covariates

Covariates	Dimension	Source	SMD		VR	
			Before	After	Before	After
Student's ethnicity (African-American vs. other)	Demographics	Archival	0.27	0.26	1.07	1.19
Student's gender	Demographics	Archival	0.32	0.10	1.45	1.01
Household highest level of education	Demographics	Parent	0.34	0.27	1.14	1.30
Student's composite district literacy test score	Performance	Performance	0.66	0.19	2.50	1.92
Student's prosocial behaviors	Behavior	Classmates	0.36	0.26	1.31	1.53
Student's prosocial behaviors	Behavior	Parent	0.16	0.02	1.11	1.02
Student's ADHD behaviors	Behavior	Parent	0.40	0.04	1.32	1.14
Student's ADHD behaviors	Behavior	Teacher	0.45	0.28	1.11	1.15
Student's conduct problems	Behavior	Parent	0.34	0.03	1.53	1.10
Student's school engagement	Behavior	Teacher	0.61	0.15	1.03	1.22
Student's academic self-efficacy	Motivation	Student	0.13	0.06	1.04	1.29
Teacher's educational aspirations for student	Motivation	Teacher	0.71	0.12	1.17	1.16
Student's liking by peers	Social	Classmates	0.32	0.31	1.01	1.37
Student's social preference by peers	Social	Classmates	0.39	0.13	1.04	1.06
Student's conscientiousness personality	Personality	Teacher	0.57	0.12	1.09	1.26
Parent's perceived teachers responsibility for child's learning and behavior	Parent involvement	Parent	0.03	0.14	1.17	1.07
Parent's perceived parent-teacher shared responsibilities	Parent involvement	Parent	0.09	0.09	1.56	1.48
Parent's satisfaction with home-school communication	Parent involvement	Parent	0.08	0.14	1.16	1.09
Home-school relationship (alliance)	Home-school relationship	Teacher	0.34	0.12	1.29	1.40
Home-school relationship (parent's involvement)	Home-school relationship	Teacher	0.17	0.10	1.10	1.66

Note. SMD is absolute standardized mean difference. VR is variance ratio. Before is before propensity score equating (pre-existing difference). After is after propensity score equating using the weighting by the odds method. The covariates were selected based on their high correlations with the general motivation factor (without propensity score correction).

Table 3

χ^2 Test Statistic, RMSEA, χ^2 Difference Test of Measurement Invariance Models

Measurement Invariance Model		χ^2	*df*	RMSEA [90% CI]	χ^2 Difference Test			
					Comparison	Test Statistic	*df*	*p*
(1)	Factor Pattern	1320.33	880	0.05 [0.04, 0.05]				
(2)	Factor Loadings + (1)	1375.13	929	0.05 [0.04, 0.05]	(1)	82.63	49	.002
(3)	Latent Intercepts + (2)	1395.37	957	0.04 [0.04, 0.05]	(2)	64.41	28	< .001
(4)	Measurement Error Variances and Covariances + (3)	1472.60	992	0.05 [0.04, 0.05]	(3)	180.65	35	< .001
(5)	Factor Variances + (4)	1480.85	996	0.05 [0.04, 0.05]	(4)	9.14	4	.06

Note. RMSEA is the root mean square error of approximation. 90% CI is the 90% confidence interval. MLMV estimation was used; the test statistic for the χ^2 difference test is calculated using procedures described by Asparouhov and Muthén (2006; 2010).

Table 4

Unstandardized Factor Loadings, Factor Means, and Factor Variances of the Bifactor Model for Motivation for Education Attainment

Item	Unstandardized Factor Loadings			
	Teacher Educational Expectations	Peer Aspirations	Value of Education	General Educational Motivation
1				0.41
2				0.30
3				0.55
4				0.35
5			0.60	0.65
6				0.22
7				0.73
8				0.35
9				0.54
10				−0.10
11				0.16
12				0.05
13	0.54			0.25
14	0.56			0.42
15	0.26			0.29
16	0.47			0.16
17	0.18			0.65
18		0.90		0.55
19		0.64		0.48
20		0.31		0.53
21		0.15		0.31
22		0.62		0.60
23		0.36		0.52
24			0.68	0.30
25			0.38	0.49
26			0.89	0.27
27			0.25	0.45
28			0.30	0.47
29			0.08	0.53
30			0.28	0.50
31			0.06	0.27
32			0.17	0.38
Factor Mean (Retained / Promoted)	0.68 / 0.00 **	0.37 / 0.00 **	1.00 / 0.00 **	−0.25 / 0.00
Factor Variance (Retained / Promoted)	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00	1.00 / 1.00

Note. The factor means of the promoted students were set to 0. The factor variances of the retained and promoted students were set to 1.00. The factor means represent standardized differences (Cohen's *d*). See Appendix A for specific items.

* $p < .05$,

** $p < .01$.