Published in final edited form as:

Am Educ Res J. 2016 ; 53(5): 1343–1375. doi:10.3102/0002831216667479.

Effects of Extracurricular Participation During Middle School on Academic Motivation and Achievement at Grade 9

Myung Hee Im,

American Institutes for Research

Jan N. Hughes,

Texas A&M University

Qian Cao, and

Texas A&M University

Oi-man Kwok

Texas A&M University

Abstract

We investigated the effect of participating in two domains of extracurricular activities (sports and performance arts/clubs) in Grades 7 and 8 on Grade 9 academic motivation and letter grades, above baseline performance. Participants were 483 students (55% male; 33% Euro-American, 25% African American, and 39% Latino). Propensity score weighting controlled for potential confounders in all analyses. Delayed (Grade 8 only) and continuous participation (Grades 7 and 8) in sports predicted competence beliefs and valuing education; delayed and continuous participation in performance arts/clubs predicted teacher-rated engagement and letter grades. Benefits of participation were similar across gender and ethnicity; however, Latino youth were least likely to participate in extracurricular activities. Implications for reducing ethnic and income disparities in educational attainment are discussed.

Keywords

extracurricular participation; adolescence; academic achievement; motivation; propensity score analysis; ethnicity

An extensive body of research conducted over the past 25 years has documented positive associations between adolescents' participation in school-sponsored extracurricular activities and diverse benefits, including academic success and psychosocial well-being (for reviews, see Farb & Matjasko, 2012; Feldman & Matjasko, 2005). Despite a large body of research documenting benefits of participation, few studies have investigated the benefits of participating in extracurricular activities during the critical middle school years on academic outcomes (for exception, see Fredricks & Eccles, 2008). Furthermore, very few of these studies have used rigorous methods to control for the effect of differences between youth who participate and those who do not participate in extracurricular activities on variables measured *prior* to participation that predict both participation in extracurricular activities and outcomes. Thus, a finding of differences between participation groups may be due to these preexisting confounds rather than to participation. Drawing from bioecological theory and

using propensity score analyses to control for potential confounds, the present study investigates the effects of participation in two broad domains of extracurricular activities (sports and performance arts and clubs) during middle school on academic motivation and achievement in Grade 9.

A Bioecological Perspective on Extracurricular Participation

According to the bioecological model of development (Bronfenbrenner & Morris, 2006), individuals' transactions in particular contexts across time are the proximal drivers of development. The power of proximal processes to influence development depends on the person, the context, and the timing of the transactions. In essence, bioecological models posit that the same experience or transaction may affect developmental processes differently depending on characteristics of the person, characteristics of the context, and the timing of the experience (both in terms of the person's age and the regularity and duration of the experience). Each of these three dimensions is pertinent to understanding effects of extracurricular participation on youths' development.

Activity Context

Youth participate in a variety of contexts (e.g., school, home, communities, neighborhoods, and peer groups) that shape their behaviors, motivation, values, competencies, and views of self and the world. Extracurricular settings are considered important contexts for development, and a youth's transactions within these settings (e.g., interacting with peers and adult leaders, following rules and routines, setting and monitoring performance goals, and confronting and overcoming challenges) are considered proximal drivers of development (Fredricks & Simpkins, 2013).

Diverse extracurricular activities such as sports, music, and clubs share certain features that distinguish them from many of the adolescent's other contexts. For example, extracurricular activities are often structured in ways that facilitate high-quality peer interactions and the development of prosocial friendships (Fredricks & Simpkins, 2012; Simpkins, Vest, Delgado, & Price, 2012). Additionally, youth report a greater emphasis on teamwork and social skills in extracurricular activities, relative to school and other activities, and report more opportunities for initiative, managing emotions, and identity work (Larson, Hansen, & Moneta, 2006).

In addition to differences between extracurricular activity contexts and other contexts, researchers have suggested that different types of activities, such as sports versus music or arts, offer different experiences that may account for differential effects on development (Denault & Poulin, 2009a; Fredricks & Eccles, 2008). For example, in a large sample of 11th graders, students reported experiencing more opportunities for initiative (perseverance and goal achievement), emotional regulation, and teamwork in sports than in other extracurricular activities but fewer opportunities for identity work, positive peer relationships, and experiences that build social capital and prepare youth for college (Larson et al., 2006).

Consistent with a finding of differences in developmental experiences across different activity contexts, researchers have found differences in outcomes associated with participation in sport and non-sport (e.g., academic clubs or performing arts) activities. Specifically, participation in performing arts (e.g., theater, choir, and band) and academic and service clubs is more consistently related to higher grades and academic values than is participation in sports (Denault & Poulin, 2009a; Fredricks & Eccles, 2008). Conversely, participation in sports may be more consistently related to a higher sense of school belonging and closer social ties among students, parents, and schools than is participation in non-sport activities (Broh, 2002; Villarreal, 2013). Sports participation but not participation in performing arts/clubs has also been associated with higher levels of alcohol use and other risky behavior (Eccles & Barber, 1999; Fredricks & Eccles, 2008). These findings may be due to differences in peer group experiences in sports and non-sport activities. For example, Denault and Poulin (2009b) found that boys with behavior problems are more likely to self-select into sport activities than non-sport activities.

Activity Timing

Bioecological theory assumes that the timing of transactions within and across contexts is important. Those transactions that occur more regularly and for a longer period of time have a greater influence of development. Extracurricular activities typically involve regular and frequent participation, which may account, in part, for their effects on development. The current study focuses on the duration of participation. Specifically, duration in the current study refers to the number of years a youth participates in an extracurricular activity. More years of participation is expected to yield more benefits than fewer years due to the time it takes to build meaningful social ties with peers and adult leaders and the skills necessary to perform well. In a review of studies on participation duration, Bohnert, Fredricks, and Randall (2010) concluded that participation for two years yields more benefits than participation for one year and that additional years of participation beyond two may provide further benefits. Importantly, there is a dearth of studies examining effects of consistency or stability of participation. Beginning involvement in an activity one year and not continuing it the following year may signal negative experiences in that activity, which could negatively impact students' liking for and engagement in school. Based on this reasoning, one might expect that students who participate in one year but discontinue participation the second year (discontinued group) would gain fewer benefits from participation than students who do not begin participation until the second year (delayed start group) even though both groups of students participated for an equivalent period of time.

Person Variables

Person variables refer to characteristics such as age, gender, and ethnicity as well as individual differences in physical appearance, mental and emotional resources, and motivation. With respect to age, the same experience may have different developmental impacts at different ages. Although the majority of research on extracurricular participation has occurred at the high school level, the effects of participation may be particularly strong in middle school, when young adolescents are undergoing rapid biological, cognitive, and social changes. For example, early adolescence is a time of increased susceptibility to the

influence of one's peers and an increased desire for autonomy from parents (Brown & Larson, 2009).

The middle school years are of critical importance to students' long-term academic attainment (Alexander, Entwisle, & Kabbani, 2001). The normative decline in students' academic motivation and achievement at the transition to middle school has been explained by a lack of fit between the developmental needs of the youth adolescent and the capacity of the school to meet those needs (Eccles, Wigfield, Midgley, & Reuman, 1993). The transition often brings larger and complex peer ecology, a departmentalized curriculum, and less support from teachers (Eccles & Roeser, 2009). In the face of these challenges, many students disengage from school during the middle school years. The present study aims to understand the effect of extracurricular participation during this period among an ethnically diverse, predominantly low socioeconomic status (SES) sample of youth who entered first grade with low literacy skills. Low literacy skills in first grade are one of the strongest predictors of subsequent academic failure (Sonnenschein, Stapleton, & Benson, 2010).

Research on individual characteristics associated with benefit from extracurricular participation has been restricted primarily to gender and ethnicity. Although boys and girls are equally likely to participate in extracurricular activities, boys are more likely to participate in sports, whereas girls are more likely to participate in performance and fine arts (Denault & Poulin, 2009b; Eccles & Barber, 1999; Fredricks & Eccles, 2008). Furthermore, different factors may predict participation for boys and girls (Denault & Poulin, 2009b; Feldman & Matjasko, 2007). Despite gender differences in predictors and activity contexts, the effects of extracurricular participation are generally similar across gender, with gender moderation inconsistently found (Fredricks & Eccles, 2006, 2008).

With respect to ethnicity, researchers have found that Latino/Hispanic students are less likely to participate in extracurricular activities than are other ethnic groups (Lugaila, 2003; National Center for Education Statistics, 2012; Ream & Rumberger, 2008). Using a national data set, Feldman and Matjasko (2007) reported that of youth in Grades 7 through 12, 36.7% of Latino students did not participate in any extracurricular activity, compared to 21.0% of White/Euro-American students and 25.6% of African American students. Scholars have suggested that the lower levels of participation among Latino youth may be a result of language barriers, a lack of parental awareness of the benefits of participation, and Latino cultural values of duty to the family, which may require youth to be wage earners or help with other family obligations (Peguero, 2010; Shannon, 2006).

Comparative studies find that all ethnic groups benefit from extracurricular participation, even after adjusting for multiple self-selection factors (Marsh & Kleitman, 2002). The few comparative studies of multiethnic samples of youth report that Latino students benefit more from participation than do African American or Euro-American students (Hull, Kilbourne, Reece, & Huanini, 2008; Villarreal, 2013). For example, using data from a large school district in Oregon, Towe (2012) found a stronger association between extracurricular participation and GPA for Latino than for non-Latino students. Some authors have suggested that extracurricular participation is especially beneficial to Latino students because it fosters

social integration at school, thereby increasing access to non-Latino sources of social capital (Ream & Rumberger, 2008; Simpkins, O'Donnell, Delgado, & Becnel, 2011).

Selection Effects

In addition to the aforementioned substantive gaps in the literature on effects of extracurricular participation, prior studies have not adequately controlled for potential confounds that may account for an association between extracurricular participation and outcomes. Despite an extensive body of research documenting associations between extracurricular participation and academic motivation and achievement, the fact that students choose to participate, or not, in these activities poses a serious obstacle to reaching conclusions regarding a causal role for participation. The crux of the problem is that students are not randomly assigned to participation; rather, students select (or are recruited into) these activities. Furthermore, many students, family, and school variables that are associated with selection into participation versus nonparticipation are also associated with the measured outcomes. For example, student demographic variables (e.g., age, gender, ethnicity, socioeconomic status) and individual characteristics (e.g., interests, competencies, social and behavioral adjustment, academic achievement), family variables (e.g., parent involvement in school and encouragement for participation), and school factors (e.g., size, availability of opportunities) predict whether students participate or not in extracurricular activities as well as the type and intensity of participation (Denault & Poulin, 2009b; Feldman & Matjasko, 2007).

Due to potential selection effects, a finding that participants and nonparticipants differ at some future point on outcomes of interest may tell us little about the *effect* of participation on these outcomes. A few researchers investigating effects of extracurricular participation on development have employed a number of strategies for minimizing selection effects. The most common strategy is the use of covariate analyses in which the effects of a limited number of confounding variables are statistically controlled (Broh, 2002; Fredricks & Eccles, 2006; Gardner, Roth, & Brooks-Gunn, 2008). However, these statistical adjustments can employ only a limited number of observed covariates that may not capture all of the preexisting differences between participants and nonparticipants. Additionally, these statistical adjustments make a number of critical assumptions about the relationship between the covariates and the outcomes across the groups of interest that are rarely tested (Shadish, Cook, & Campbell, 2002).

Prior reviews of the literature on extracurricular participation reflect a consensus that improved methods to minimize threat differences between participation groups prior to participation are needed to move the field forward. In 2005, Feldman and Matjasko concluded that "it is necessary to reduce selection bias to gauge participation's true impact" (p. 202). In their 2012 review, Farb and Matjasko noted that despite repeated calls for better controls for selection bias, "there has not been any movement in this direction" since 2005 (p. 45).

Propensity Score Analysis

Increasingly, social science researchers have employed propensity scores to minimize selection effects in nonexperimental studies. A propensity score is defined as the conditional probability of assignment to a particular “treatment” given a vector of observed covariates (Rosenbaum & Rubin, 1984). Propensity score analysis generates a single index—the propensity score—that summarizes information across potential confounds. Procedures such as matching and weighting can then be used to equate the participant and nonparticipant students on their propensity scores (West et al., 2014). To the degree that equating is achieved on all potential confounders, the propensity score analysis produces an unbiased estimate of the average effect of participation on students. Although one can never be certain that all potential confounders are equated, or balanced, across groups, use of a broad set of covariates known to be associated with the treatment and the outcome minimizes this risk (West et al., 2014). To the authors’ best knowledge, no study has employed propensity score analyses in estimating the effect of extracurricular participation on academic outcomes.

Study Purpose and Research Hypotheses

Given the aforementioned limitations of prior research on effects of extracurricular participation on academic functioning, the present study investigated the role of activity context, duration of participation, and the youth’s gender and ethnicity on effects of participation in extracurricular activities during the middle school grades on academic motivation and achievement. Specifically, the current study investigated the effects of timing of participation across Grades 7 and 8 (i.e., continuous, delayed, discontinued, and no participation) in two broad activity contexts (sports and performance arts/clubs) on students’ Grade 9 academic outcomes, above baseline performance on these outcomes. The potential moderating roles of two person variables, gender and ethnicity, were also investigated.

Based on research suggesting that different activity contexts may influence different aspects of academic motivation and achievement (Blomfield & Barber, 2010; Denault, Poulin, & Pedersen, 2009), we assessed four distinct but related outcomes. Specifically, academic competence beliefs and subjective valuing of academic achievement were selected as outcomes based on extensive evidence of their association with academic effort and attainment (Wigfield, Cambria, & Eccles, 2012). Students’ course grades and teacher-rated behavioral engagement were selected as outcomes based on the strong association between these outcomes at Grade 9 and successful completion of high school and enrollment in postsecondary education (Donegan, 2008; Janosz, Archambault, Morizot, & Pagani, 2008). Importantly, propensity score analysis was employed to strengthen the basis for reaching causal conclusions regarding effects of extracurricular participation on subsequent academic motivation and achievement.

Based on theory and empirical findings discussed previously, we expected boys would participate at a higher rate than girls in sports but that girls would participate at a higher rate than boys in performance arts/clubs. We expected Latino students to participate in both activity domains at a lower rate than African American or Euro-American youth. Relative to students who did not participate in extracurricular activities in Grades 7 or 8 (nonparticipants), we expected students who participated in Grades 7 and 8 (continuous

participants) and students who participated in Grade 8 only (delayed participants) would perform better on Grade 9 outcomes. However, we expected students who participated only in Grade 7 (discontinued participants) in either activity domain would not differ from nonparticipants on Grade 9 outcomes.

With respect to activity context, based on the evidence that participation in performance arts (e.g., band and chorus) and academic clubs is more consistently predictive of academic performance than is sport participation, we expected participation in performing arts/clubs would predict teacher-awarded letter grades and teacher-rated engagement. Based on the reasoning that all school-sponsored activities provide a sense of belonging to school and valuing of school, we expected both activity contexts would predict academic competence beliefs and educational values. With respect to gender and ethnic moderation of participation effects, we expected girls and boys would benefit similarly from participation and that Latino students would benefit more from participation than Euro-American or African American students.

These research questions were pursued with a predominantly low-income sample of youth who were recruited into the current longitudinal study on the basis of academic risk (see Participants section). Although lower achieving and low SES youth are less likely to participate in extracurricular activities, they may be more likely to benefit from participation. For example, participation in extracurricular activities in middle and high school reduced the probability of dropping out of school only for students with low academic and social competence in middle school (Mahoney & Cairns, 1997). A finding of an effect of participation on these outcomes would suggest that increasing participation in middle school is a viable strategy for increasing the educational attainment of students with elevated risk for school failure.

Methods

Overview

Data on students' extracurricular participation was obtained in interviews when students were in Grades 7 and 8. We selected these grades because a wide range of extracurricular activities, including sports, performance arts, and service and academic clubs, were available at these grades on each school campus. Outcomes were assessed at baseline (Year 5 in the longitudinal study, when students were in Grades 4 or 5) and Grade 9. A difference in the grade at which the baseline measure was administered is due to the fact that some students repeated an elementary grade ($n = 156$, 31.5%). For teacher-rated engagement and reading and math achievement, the same outcome measure was used at baseline (Grades 4 or 5) and Grade 9. For teacher-awarded letter grades (which were typically provided by the language arts teacher), the baseline measure was the score on a measure of reading achievement. As described in the following measures section, developmentally appropriate measures of student-perceived academic competence and valuing of education were used at Grade 9 and baseline. Different sources reported on different outcomes: Students reported on their academic competence beliefs and valuing of education, teachers reported on students' letter grades, and reading achievement was assessed on an individually administered test. Bilingual students were interviewed and tested in the language in which they were more

proficient, based on scores on the Woodcock-Muñoz Language Test (Woodcock & Muñoz-Sandoval, 1993), by bilingual examiners.

Participants

Participants were 483 students recruited in the fall of 2000 or 2001 into a larger longitudinal study ($N = 784$) when they were in Grade 1. Data on participation in extracurricular were collected when these students were in Grades 7 and 8 (i.e., years 2007–2009).

Students in the larger longitudinal sample were enrolled in one of three school districts (one urban and two small city districts) in Texas and were selected into the study on the basis of scoring below the median on a district-administered test of literacy administered in the spring of kindergarten or the fall of Grade 1. Based on school records, School District A (student population = 13,558) had an ethnic distribution of 38% White/Euro-American, 37% Latino/Hispanic, 25% African American, and less than 1% other. District B (student population = 24,429) had an ethnic distribution of 35% White/Euro-American, 30% Latino/Hispanic, 30% African American, and 5% other. District C (student population = 7,424) had an ethnic distribution of 67% White/Euro-American, 12% Latino/Hispanic, 12% African American, and 9% other. Additional inclusionary criteria for the larger study included speaking English or Spanish, not receiving special education services other than speech and language services, and not having been previously retained in Grade 1. Of the 1,374 children eligible to participate in the longitudinal study, written parent consent for participation for 5 years was received for 784 (65%). Students with and without consent did not differ on a broad array of variables. Details on recruitment are reported in Hughes, Luo, Kwok, and Loyd (2008).

At the end of the first five years of participation in the study, parental consent for continued participation was received for 569 of the 784 participants. Almost all nonconsent was due to nonresponse. Attrition analyses found no differences between participants and nonparticipants on a wide range of variables assessed when students were in first grade, including gender, parent education level, literacy scores, reading and math achievement, IQ, ethnicity, and bilingual status. Participants (54.7% male) were 12.56 years of age ($SD = .37$) at Grade 7, 65.9% were economically disadvantaged based on income eligibility for free or reduced lunch, and 41.5% of parents' highest level of educational attainment was a high school diploma or less. The ethnic composition of the sample was 33.1% Euro-American, 25.3% African American, 38.5% Latino/Hispanic (of whom 32.0% were enrolled in bilingual education at Grade 5), and 3.1% Other. Participants were enrolled in 64 schools during Grade 7 and 72 schools in Grade 9. The increase in the number of schools reflects the increased geographical dispersal of the sample over time. At Grade 7, participants' mean reading age-standard scores from the Woodcock-Johnson III (Woodcock, McGrew, & Mather, 2001) or its Spanish-language equivalent (Batería III Woodcock-Munoz; Woodcock, Muñoz-Sandoval, McGrew, Mather, & Schrank, 2004) was 96.44 ($SD = 14.13$).

Measures

Extracurricular Participation—In individual interviews at school, the interviewer asked students to indicate if they participated that year in each of four activity contexts: (a) sports;

(b) performance arts, fine arts, or music; (c) academic clubs; and (d) other school activities such as student council, newspaper, or service activities. For each activity category, students were given examples of activities that fit that category (e.g., examples of sports activities included football, baseball, cheerleading, pep squad, and tennis). Students were asked to name the specific activity in which they engaged, and their answers were written verbatim. Verbatim responses were coded into activity domain by two graduate assistants, whose agreement was 98%. Students were asked to report only activities that were sponsored by the school and not part of the regular school day. Thus, if students were in a drama class but did not participate in drama activities outside class time (e.g., practices or performances after school hours), that activity was not counted. Based on the relatively small number of students participating in academic clubs and other school activities such as student council and prior research finding similar profiles of participants across these activities (Feldman & Matjasko, 2007), these activities were combined with performance arts and music into a performance arts/clubs category. Although data on intensity of participation were collected, for the present study, participation in each broad activity category (sports and performance arts/clubs) was defined as a dichotomous variable.

Four different duration patterns in participation across two consecutive grades (i.e., Grades 7 and 8) were possible: (1) continuous: students participated for both years; (2) discontinued: students participated only in Grade 7; (3) delayed: students participated only in Grade 8; and (4) nonparticipation: student did not participate in either year. Nonparticipants were the reference group for the three contrasts.

Competence Beliefs and Valuing of Educational Attainment—At Grade 9, students completed the 11-item Academic Competence and Effort Beliefs Scale ($\alpha = .89$) and the 10-item Value of Education Scale of the Motivation for Education Attainment Questionnaire ($\alpha = .85$) (masked), a multidimensional measure of motivation to complete high school and pursue postsecondary education. Example Academic Competence and Effort Beliefs items include “I am on track to graduate from high school” and “Nothing will get in the way of my going to college.” Example Value of Education items include “If I work hard in school, I will get a better job than the kids who don’t try hard” and “School is not that important for future success” (reverse scored). The scale has demonstrated good construct and criterion-related validity in an at-risk sample of Grade 9 students (Cham, West, Hughes, and Im, 2015).

At baseline, students’ academic competence beliefs and educational values were assessed with the Competence Beliefs and Subjective Task Values Questionnaire (Wigfield et al., 1997). Five items assess competence beliefs in each subject (i.e., reading and math). Specifically, children were asked how good they were in that subject, how good they were relative to the other things they do, how good they were relative to other children, how well they expected to do in the future in that subject, and how good they thought they would be at learning something new in that subject. Students indicated their response to each item by pointing to a thermometer numbered 1 to 30. The endpoint and midpoint of each scale was labeled with a verbal descriptor of the meaning of that scale (e.g., with 1 indicating *not at all good*, 15 indicating *ok*, and 30 indicating *one of the best*). Children rated their subjective valuing of reading and math by indicating how interesting/fun each subject was, how

important they thought being good in each subject was compared to other activities, and how useful they thought each subject was using a similar 1 to 30 scale. Wigfield et al. (1997) reported that children's reports on this measure were moderately correlated with teacher and parent report of competencies for children as young as third grade. Based on moderate correlations between reading and math competency scores ($r = .30$) and reading and math subjective valuing scores ($r = .54$), a mean academic competence beliefs score and a mean academic valuing score were computed for reading and math.

Letter Grades—Students' language arts teachers were asked to report the letter grade (from A to F, with A = 4 and F = 0) that the student received in his or her class for the most recent grading period. Language arts was selected because all students take language arts in Grade 9. In a few cases (7%), when a language arts teacher was not available to report on students' grades, another teacher who knew the student well reported on the student's grades in his or her class.

Teacher-Rated Classroom Engagement—The same teacher who reported on students' grades also rated students' classroom engagement using an 11-item questionnaire adapted from Skinner, Zimmer-Gembeck, and Connell (1998). Items assess effort, persistence, concentration, and interest. Example items include: "tries hard to do well in school," "participates in class discussion," "pays attention in class," and "just wants to learn only what he/she has to in school" (reverse scored). Teachers were asked to indicate the extent to which each statement was true on a 1 (*not true at all*) to 4 (*very true*) scale. The scale demonstrates good factorial validity (masked) and internal consistency (α at baseline and Grade 9 was .92 and .91, respectively).

Reading Achievement—The Woodcock-Johnson III Tests of Achievement (WJ-III; Woodcock et al., 2001) is an individually administered measure of academic achievement for individuals ages 2 to adulthood. The WJ-III Broad Reading W Scores, which are based on the Letter-Word Identification, Reading Fluency, and Passage Comprehension subtests, were used. Extensive studies document the reliability and construct validity of the WJ-III (Woodcock et al., 2001). Spanish language-dominant children were administered the Bateria III, the equivalent Spanish version of the WJ-III (Woodcock et al., 2004), by bilingual examiners.

Covariates for Propensity Score Analysis—A total of 44 covariates (potential confounders), all of which were measured in Grades 4 or 5, prior to opportunity to participate in middle school activities, were used to estimate the propensity scores of students who did and did not participate in extracurricular activities (see propensity score estimation under Study 1 and Study 2 in the following). These 44 covariates (listed in Table 1) were selected to be as comprehensive as possible, including variables that have been shown in prior research to be associated with extracurricular participation and with measures of academic motivation and achievement. These variables were assessed with direct child testing and interviews (e.g., measures of language proficiency, academic achievement, perceived teacher-student support, perceived competence beliefs in reading and math, value of reading and math, and perceived social acceptance), teacher questionnaires (e.g.,

behavioral, academic, and social functioning), parent questionnaires (e.g., family demographics, educational aspirations, and child behavioral and social functioning), and school records (e.g., child ethnicity, age, gender, and bilingual class placement).

Data Analytic Procedure

Propensity Score Analysis

As discussed in the introduction, the first step in propensity score analysis is to estimate each student's propensity score for each participation category (i.e., conditional probability of being in each participation category) given the student's scores on the covariates. The second step is to equate the estimated propensity scores' distributions between the participants and nonparticipants. The third step is to check the balance of the distributions of the set of 44 covariates between the participating and nonparticipating students. Each step is described in greater detail in the following.

Estimation—We estimated two propensity scores depending on two activity domains, sports and performance arts/clubs. The propensity score for the sports was the probability of a student's participating in sports in Grade 7 versus not participating in sports. The propensity score for the performance arts/club activity was the probability of a student participating in performance arts/clubs in Grade 7 versus not participating in performance arts/clubs in Grade 7.

To estimate propensity scores, we used the random forests method (Breiman, 2001) under the R package version 3.1.0 (Strobl, Boulesteix, Kneib, Augustin, & Zeileis, 2008). According to previous studies (Lee, Lessler, & Stuart, 2010), the random forests method reduces bias in the estimate of the effect of a treatment (i.e., extracurricular participation) on outcomes by identifying complex and nonlinear relationships of covariates with students' participation status.

Equating—With the estimated propensity score, we employed weighting approach using the odds method (Hirano, Imbens, & Ridder, 2003) under R package. In this method, each student who belonged to a particular participation category was given a weight of 1.00, whereas nonparticipants were given a weight of $\frac{\hat{\pi}}{1-\hat{\pi}}$, where $\hat{\pi}$ is the nonparticipating student's estimated propensity score. This weighting procedure reflects survey sampling weighting procedures to estimate parameters and their associated standard errors (Asparouhov, 2005). With this weighting by the odds method, one can estimate the effect of activity participation for students who did not actually participate compared to closely equated students who participated (Schafer & Kang, 2008).

Covariates Distribution Balance—To evaluate the effectiveness of the propensity score equating procedure, we checked the balance of the distributions of the completely observed values of the covariates across participants' groups and the missing data pattern of the covariates between the participating and nonparticipating students (Rosenbaum & Rubin, 1984). Specifically, using the weighted propensity scores, we calculated the absolute standardized mean difference (SMD) of the 44 covariates between participation groups

(Rubin, 2001; Stuart, 2010). An SMD score of 0 indicates perfect balance (i.e., no difference in SMD).

Effect of Extracurricular Participation on Students' Outcomes

Analysis of covariance (ANCOVA) was employed to examine the hypothesized model using *Mplus* version 7.2 (Muthén & Muthén, 1998–2014). We used students' weighted propensity scores with WEIGHT function. In data analysis, we calculated intraclass correlation (ICC) for school since school characteristics such as size is associated with participation (Feldman & Matjasko, 2007) to determine the necessity of accounting for the data dependency. We then used TYPE=COMPLEX in combination with the CLUSTER function (i.e., school at Grade 9) to take into account the data dependency (i.e., students nested within schools).

The effects of duration (timing) of extracurricular participation by activity domain in Grades 7 and 8 were estimated separately for each of the four Grade 9 academic outcomes (competence beliefs, valuing of education, teacher-rated engagement, and teacher-awarded letter grades), controlling for prior performance on the outcomes. Tests of the effects of sport duration on outcomes used students' weighted propensity score for participating in sports in Grade 7. Tests of the effects of performance arts/clubs' duration used students' weighted propensity score for participating in performance arts/clubs in Grade 7. We created four categories representing groups with different duration patterns: continuous, delayed, discontinued, and nonparticipation. We then created the three dummy variables (i.e., continuous, delayed, and discontinued), using the nonparticipation group as the reference group (see Measures section for more details), which was assigned a value of 0. Each of the other groups was given a value of 1 on the dummy variable that contrasted it with the reference group in the analysis and a value of 0 on the other dummy variables. With this dummy variable coding approach, effects can be directly interpreted as differences in effects of each pattern of duration (i.e., continuous, delayed, and discontinued participation) relative to effect of the reference group (i.e., nonparticipation) in that domain. For example, the effect of continuous participation equals the effect of continuous participation *minus* the effect of no participation. When more than one participation pattern had a significant effect on an outcome, we then conducted post hoc tests to determine if one pattern had a stronger effect than another pattern. For the post hoc tests, we used the Wald test.

We first allowed the relation to vary between the two effects and each outcome separately. Next, we imposed equality constraints on these relations. Next, using multiple group analysis, we investigated whether student's gender (boys and girls) or ethnicity (Euro-American, African American, and Latino students) moderated the effects of different duration pattern and Grade 9 outcomes.

Results

Descriptive Statistics

Table 2 presents the number and percentage of students participating in each broad activity domain for Grades 7 and 8 and the number of students in each duration pattern, separately by gender and ethnicity. Table 3 presents descriptive statistics for the outcomes in the

hypothesized models. The variables were screened for non-normality and extreme values. None of the variables used in the analyses exhibited levels of skewness or kurtosis associated with problematic tests of fit or standard errors (West, Finch, & Curran, 1995). Each baseline score measured at Year 5 was predictive of the corresponding outcome at Grade 9 (range, .15–.29). All outcomes were positively correlated with each other. The associations between education belief and academic competence belief and between teacher-rated engagement and letter grade were strong (.52 and .65, respectively).

Covariates Balance Across Participating and Nonparticipating Groups

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, parent report, and teacher report). Figure 1 shows the balance measures before and after weighting the sample. After weighting by the odds method together with the random forests propensity scores, the balance improved the distributions of propensity scores and the distributions of the observed values of the covariates between the participating and nonparticipating students. Furthermore, after propensity score equating, the covariates had lower standard mean differences. Specifically, SMDs of covariates selected on the basis of high correlations with the outcomes ranged from 0 to 0.23, indicating good balance (Ho, Imai, King, & Stuart, 2007). Additional details on covariate balance are available from the corresponding author. We concluded that the weighting procedure successfully equated the participating and nonparticipating students on the set of 44 covariates measured in or prior to Year 5 (before any student had begun middle school).

Gender and Ethnic Differences in Participation

In responding to our first hypothesis, we tested gender and ethnic differences in participation in sports and performance arts/clubs as well as in non-participation at either Grade 7 or Grade 8. Because results were similar across grades in terms of direction and magnitude, findings only for Grade 8 only are detailed here.

Gender Differences—As expected, girls were less likely to participate in sports than boys (z -ratio = 22.78, p = .006) but were more likely than boys to participate in performance art/clubs (z -ratio = 2.71, p = .007). Although boys and girls differed in what activity they selected, they did not differ in whether they were involved in any extracurricular activities or not.

Ethnic Differences—Also as expected, Latino students were less likely than African American (z -ratio = -2.14 , p = .032) or Euro-American (z -ratio = -3.16 , p = .002) students to participate in sports. African American and Euro-Americans did not differ in their rate of participation in sports. In the domain of performance arts/clubs, Latino students were less likely to participate than Euro-American students (z -ratio = -2.12 , p = .034), whose participation was similar to that of African Americans. When considering ethnic differences in status as participating or not in any extracurricular activity, Latino were less likely to participate than African American (z -ratio = -2.07 , p = .038) or Euro-American youth (z -ratio = -3.68 , p < .001), who did not differ from each other.

Effect of Duration of Participation on Outcomes

First, at the school level we calculated the intraclass correlation for each Grade 9 outcome. The ICCs for competence belief, valuing of education, teacher-rated classroom engagement, and teacher-awarded letter grade were .07, .00, .03, and .04, respectively. Thus, all analyses took into account the nonindependent data structure (i.e., students nested within schools).

We then tested the effect of the different duration patterns across Grades 7 and 8 on outcomes at Grade 9, separately by the activity domain (i.e., sports or performance arts/clubs). The hypothesized ANCOVA models we examined with *Mplus* are the saturated path models (with degrees of freedom of zero) in the structural equation modeling (SEM) framework. Therefore, we do not report the fit indices due to perfect fit for all models in the study. Table 4 shows the standardized parameter estimates, corresponding standard error, and *p* value for each effect. Given that the reference group was the nonparticipation group, effects can be directly interpreted as differences in effects of each pattern of duration (i.e., continuous, delayed, and discontinued participation) relative to nonparticipation in that domain. Results for duration patterns will be presented separately for each activity category.

Continuous participation in sports had a significant positive effect on Grade 9 academic competence beliefs ($\beta = .14$, $SE = .04$, $p < .001$) and valuing of education ($\beta = .16$, $SE = .05$, $p = .003$). A significant positive effect of delayed sports participation was found only for Grade 9 valuing of education ($\beta = .15$, $SE = .05$, $p = .004$). Finally, we found no effect of discontinued participation in sports relative to no participation.

Turning to performance arts/clubs, continuous participation had a significant positive effect on Grade 9 competence belief ($\beta = .17$, $SE = .03$, $p < .001$), teacher-rated classroom engagement ($\beta = .21$, $SE = .05$, $p < .001$), and teacher-awarded letter grades ($\beta = .16$, $SE = .04$, $p < .001$). Results of delayed performance arts/clubs' participation parallel those for continuous participation, with effects found for Grade 9 competence belief ($\beta = .10$, $SE = .04$, $p = .019$), teacher-rated classroom engagement ($\beta = .10$, $SE = .04$, $p = .020$), and teacher-awarded letter grades ($\beta = .06$, $SE = .03$, $p = .020$). A significant positive effect was found for discontinued participation in performance arts/clubs relative to no participation in performance arts/clubs on academic competence beliefs only ($\beta = .09$, $SE = .04$, $p = .040$).

Gender and Ethnic Moderation of Effect of Duration Patterns—Using multiple group analysis, in separate analyses we tested the potential gender and ethnic moderating effects in the relation between the pattern of duration of participation and each Grade 9 outcome. We first allowed the relations between each pattern of participation and outcomes to vary across student gender and ethnicity (i.e., relaxed model) and then imposed equality constraints on these relations (i.e., constrained model), sequentially. We compared the two competing nested models (i.e., relaxed vs. constrained) at the significance of $\alpha = .05$ using Wald tests. The null hypothesis of the Wald test is that the constrained model (indicating no moderation effect) fits the data equally well as the relaxed model (indicating the existence of moderation effect). We tested all possible comparisons between groups (i.e., male vs. female for testing gender moderation; Euro-American vs. Latino; Euro-American vs. African American; Latino vs. African American for testing ethnicity moderation) on each outcome in two activity domains. Additional details on the results of Wald tests for testing gender and

ethnicity moderation effect of duration patterns of participation on each of grade 9 outcomes are available from the corresponding author.

Gender moderation: According to Wald tests (with degrees of freedom of one for all Wald tests), one significant gender moderation effect in sport ($\chi^2 = 5.68$, $p = .017$ for valuing of education) of the 12 tests and two significant gender moderation effects in performance arts/clubs (for teacher-rated classroom engagement, $\chi^2 = 12.10$, $p < .001$; for teacher-awarded letter grade, $\chi^2 = 3.97$, $p = .046$) of the 12 tests were found. Specifically, in sports, the effect of discontinued participation on valuing of education was positively significant only for girls ($\beta = .16$, $SE = .08$, $p = .040$). In performance arts/clubs, for teacher-rated classroom engagement, the effect of discontinued participation was positively significant only for boys ($\beta = .12$, $SE = .03$, $p < .001$). For teacher-awarded letter grades, the effect of delayed participation was positively significant only for boys ($\beta = .18$, $SE = .07$, $p = .007$).

Ethnic moderation: For ethnic moderation analyses, we tested 36 pairwise comparisons (4 outcomes $\times$ 3 pairwise among three ethnic groups $\times$ 3 duration patterns) in each activity domain. Based on Wald tests for testing ethnic moderation effect, one significant effect in sport (for teacher-awarded letter grades $\chi^2 = 10.76$, $p = .001$) of the 36 tests and eight significant effects in performance arts/clubs (for simple presentation, these results of Wald tests were not reported here) of the 36 tests were found.

Turning to performance arts/clubs, for competence belief, the effect of discontinued participation was positively significant for Latino students ($\beta = .14$, $SE = .06$, $p = .018$) and African American students ($\beta = .326$, $SE = .045$, $p < .001$) while the effect of discontinued participation was negatively significant for Euro-American students ($\beta = 2.12$, $SE = .05$, $p = .017$). For valuing of education, the effect of continuous participation was positive and significant only for Latino students ($\beta = .10$, $SE = .03$, $p = .001$) whereas the effect of discontinued participation was negative and significant only for Euro-American students ($\beta = 2.15$, $SE = .05$, $p = .004$). For teacher-rated classroom engagement, the effect of discontinued participation was positive and significant only for African American students ($\beta = .16$, $SE = .06$, $p = .009$), whereas the effect of delayed participation was positively significant only for Latino students ($\beta = .22$, $SE = .09$, $p = .011$).

Discussion

The present study is the first to employ the weighted propensity score analyses to control for students' probability of participating in school-sponsored extracurricular activities. Given the frequent caveat that individual, family, and school characteristics associated with participation profiles may account for observed associations between participation and outcomes, it is surprising that propensity score analyses has not previously been used in this body of research. When propensity scores are calculated on a large number of variables associated with the predictor and the outcome, as the case in the present study, propensity score analyses employing weights or matching closely mimic the results of an experimental study (West et al., 2014). Although we cannot definitively rule out the possibility that preexisting differences on some unmeasured variable potentially accounts for the results, this possibility is greatly reduced given the breadth of the baseline assessment and the

balance achieved on the covariates across participation categories. The present study also controlled for students' baseline scores on each outcome measure, thereby substantially increasing the internal validity of the effects detected in this study. Thus, results provide strong support for the conclusion that participation in school-sponsored sports and performance arts or clubs in middle school has a positive impact on students' academic motivation and achievement at the transition to high school. Furthermore, the broad activity domain (sports or performance arts/clubs) and duration of participation are differentially predictive of academic outcomes.

Activity Domain

Sports participation predicted competence beliefs and educational value, whereas participation in performance arts/clubs predicted competence beliefs, teacher-rated classroom engagement, and teacher-awarded grades. The finding that both activity domains predicted an increase in academic competence beliefs suggests that extracurricular activities provide a context in which students can meet and overcome challenges and increase one's skill level, thereby building confidence. The finding that performance arts/clubs but not sports predicts achievement is consistent with prior research and with the proposition that the social context for band and academic clubs provides greater reinforcement for academic achievement than is the case for sports (Denault & Poulin, 2009a; Fredricks & Eccles, 2008). Importantly, the current study's methodology provides a stronger basis for concluding that the association between activity context and achievement is due to differences in participation context rather than differences in student and family characteristics associated with selection into different contexts. Notably, we calculated two propensity scores based on the two broad activity domains (i.e., the propensity to participate in sports and the propensity to participate in performance arts/clubs). The decision to create separate propensity scores was based on prior research finding differences in characteristics associated with participation in sports and non-sports (Feldman & Matjasko, 2007). For example, in a study of Canadian youth, Denault and Poulin (2009b) reported that number of problem behaviors at Grade 6 was positively associated with sports participation but not with performance arts/clubs at Grade 7. Of interest, in the current study, the correlation between propensity scores for sports participation and propensity scores for performance arts/clubs was not statistically significant ($r = -.07$). That is, students who select participation in sports and students who select participation in performance arts/clubs differ on a range of student and family characteristics prior to participation. Failure to take these differences into account may lead to erroneous conclusions regarding differential benefits associated with different types of activities.

Duration of Participation

Turning to patterns of participation, although continuous participation in an activity domain was most consistently associated with benefits, students who began participation in an activity domain in Grade 8 (delayed pattern) accrued nearly as many benefits as did students who began participation in Grade 7 and continued in Grade 8. Conversely, with the exception of competence beliefs for performance arts/clubs, students who began a participation domain in Grade 7 but quit in Grade 8 did not differ from nonparticipants on Grade 9 outcomes. Thus, the number of years of participation is less important than the

pattern of years (i.e., delayed pattern was more helpful than discontinued pattern even though each pattern involves one year of participation). Unfortunately, the present study does not have information on the reasons students discontinued participation. Students who discontinue may have had negative experiences in the activity or may have not had the opportunity to participate due to participation requirements such as grades or tryout performance.

Gender

As expected, girls and boys are equally likely to participate in extracurricular activities, but boys are more likely to select sports participation than girls, and girls are more likely than boys to select performance arts/clubs. Benefits of continuous participation in each activity domain were similar for girls and boys. Gender moderated the effect of three noncontinuous participation patterns. Because we proffered no hypotheses regarding gender moderation and ran 12 tests for each activity domain (for a total of 24 tests), these three significant effects may be due to chance and should be interpreted with considerable caution. Further complicating interpretation of gender moderation of noncontinuous patterns of participation is a lack of consideration in these analyses for more complex patterns of participation, including switching from one domain to another.

Ethnicity

Consistent with prior research, Latino students in our sample had a lower rate of participation in extracurricular activities than Euro-American or African American youth. African American students were less likely than Euro-American youth to participate in performance arts/clubs but did not differ from Euro-American youth in sports participation. The lower rate of participation among Latino students in extracurricular activities may be due to a number of factors, including (a) less parental understanding of the academic benefits of participation and a corresponding lack of support for participation; (b) the value of obligation to the family within the Latino culture, which may lead students to work or take care of younger siblings instead of participating in what may be viewed as a social activity; and (c) practical constraints such as transportation and expenses associated with participation (Garner, 2013).

We expected that extracurricular participation would be more beneficial for Latino than for Euro-American or African American youth. Limited support for this hypothesis was found.

Specifically, in the domain of performance arts/clubs, discontinued participation in performance arts/clubs predicted higher competence beliefs for Latino and African American youth but lower competence beliefs for Euro-American youth. Delayed participation positively predicted teacher-rated engagement only for Latino youth. Finally, continuous participation predicted higher valuing of education for Latinos and lower valuing of education for Euro-Americans. In summary, Latino students are least likely to participate in extracurricular activities but as likely as or more likely than other ethnic groups to benefit from participation. Extracurricular activities are key social groups with which adolescents identify (Simpkins et al., 2011). Furthermore, consistent with social identity complexity theory, such participation predicts greater cross-ethnic friendships and more positive

intergroup attitudes (Knifsend & Juvenon, 2014). The present study suggests that Latino students are deprived of participation opportunities, which likely contributes to low social integration at school and a lower sense of school belonging, both strong predictors of school engagement and completion of high school (Janosz et al., 2008).

Study Limitations and Future Directions

Results need to be interpreted in the context of study limitations. First, because participants in the study were selected on the basis of academic risk when they entered first grade, results may not generalize to samples entering school with above average literacy skills. The sample is also predominantly low SES. Because low SES and below average literacy skills in first grade predict subsequent academic failure, including dropping out of school, the current sample is an important one for understanding the role of extracurricular participation in reducing academic failure and improving low educational attainment.

Second, the investigation of ethnic differences in benefits associated with extracurricular participation is hindered by failure to take into account the heterogeneity that exists within each ethnic group. Previous studies have identified important differences within Latino groups associated with extracurricular participation, including generational status, language spoken in home, cultural orientation or acculturation, and SES (Peguero, 2010; Simpkins et al., 2011). School-level factors, such as the ethnic composition of the school, may also moderate the effect of extracurricular participation (masked). Future studies conducted within Latino samples are necessary to clarify which Latino youth are most likely to benefit from extracurricular participation, in what school and activity contexts.

Third, our categorization of activities as sports or performance arts/clubs does not capture variations in the transactions occurring within specific sports (e.g., football vs. tennis) or performance arts/clubs (e.g., band vs. student council). As observed by Guest and McRee (2008), the links between activity contexts and adjustment are likely explained by social factors such as relationships, identities, and norms within specific activity settings more so than the content of an activity itself. Unfortunately, the present study is not capable of identifying those transactions occurring in activity contexts that may account for observed benefits.

Fourth, our category of discontinued participation (i.e., participating in a particular activity domain in Grade 7 and not continuing in that activity domain in Grade 8) does not distinguish between students who switched activity domains and students who were not involved in any activity domain in Grade 8. Approximately 20% of students in the discontinued groups switched activity domains. Future studies with larger, multiethnic samples are needed to distinguish between students who discontinue one activity domain but continue to be involved in another activity domain.

Fifth, budget limitations allowed us to assess teacher ratings from only one teacher, typically the language arts teacher. Aggregation of ratings from multiple teachers of subject matter courses common to all students in these grades would enhance the reliability of the teacher ratings.

Implications for Policy and Practice

Despite study limitations, the “big picture” is clear: Extracurricular participation in Grades 7 and 8 or only in Grade 8 in middle school promotes academic motivation and achievement for at-risk youth. Furthermore, these benefits are generally similar for boys and girls and for members of different ethnic groups. Because the current sample is predominantly low SES and ethnically diverse and entered first grade with below average literacy skills, study findings suggest that policies and practices that increase extracurricular participation during the critical middle school years are likely to improve school success for students at risk for poor educational attainment. Schools differ in a number of ways that may expand or limit opportunities for participation, including the number of activities offered, the financial costs to students associated with participation, academic requirements for eligibility to participate, and the level of competition for inclusion. It is important for schools and families to view participation as an educational asset rather than an expendable option, thereby increasing opportunities for participation by all students.

The finding of lower participation among Latino youth is of considerable concern given the importance of participation to improving school engagement and school completion among all ethnic groups (Donegan, 2008). Although Latinos were less likely to participate in extracurricular activities than other ethnicities, participation had an equal or, in some cases, larger impact on Latino youths’ academic competence beliefs, valuing of education, teacher-rated engagement, and letter grades. Results of a qualitative study (Garner, 2013) of middle school Latino students suggested that family obligations such as working to help with family finances and caring for younger children, lack of parent understanding of the academic benefits of extracurricular activities and a corresponding lack of support for participation, and practical constraints such as transportation and expenses serve as barriers to participation. Studies have found that parent encouragement of participation is particularly important to Latino students’ decisions to participate in extracurricular activities (Dunn, Kinney, & Hofferth, 2003; Shannon, 2006). Thus, specific outreach to Latino students and their parents, emphasizing the connection between participation and school success, and remedies to lessen practical constraints such as transportation may increase participation among this segment of the population. Given the heterogeneity within the Latino group on variables that may affect participation, including immigrant generational status and English language proficiency (Peguero, 2010), schools are advised to employ focus groups and surveys to gain a better understanding of factors that influence participation in their particular community. For example, there may be particular activities, such as soccer, that are not available at a particular school but that, if offered, might be an attractive option for Latino students in that school community.

Conclusion

Using propensity score analyses to remove selection bias, the study found that participation in extracurricular activities in middle school promotes positive school identities, behavioral engagement in the classroom, and letter grades at Grade 9, above students’ performance on the same or similar measures administered at Grade 5. Participation in sports and performance arts/clubs are associated with different outcomes; consequently, students who

participate in both sports and performance arts/clubs likely experience the broadest benefits. Students who begin an activity context in Grade 8 gain benefits comparable to students who begin an activity in Grade 7 and continue into Grade 8. Positive school identities, behavioral engagement in classroom, and course grades at the beginning of high school are highly predictive of ultimate success in completing high school (Donegan, 2008; Janosz et al., 2008; Legault, Green-Demers, & Pelletier, 2006). Given the at-risk nature of the sample, findings suggest the potential impact of extracurricular participation in middle school on reducing dropout rates and ethnic and income disparities in educational attainment.

Acknowledgments

This research was funded by grant HD 039367 from the National Institute for Child Health and Human Development awarded to Jan N. Hughes.

Appendix

Table A1

Absolute Standardized Mean Difference (SMD) of Selected Covariates

Covariates	Study 1				Study 2	
	Sports		Performance Arts/Clubs		Breadth	
	Before	After	Before	After	Before	After
Classroom behavioral engagement	.20	.18	.15	.04	.17	.18
Educational aspirations for student	.18	.21	.17	.07	.25	.17
Woodcock-Johnson III Broad Math W score	.21	.15	.30	.16	.34	.22
Home-school relationship (alliance)	.18	.14	.02	.00	.10	.06
Teacher-student warmth	.21	.04	.01	.02	.18	.04
Woodcock-Johnson III Broad Reading W score	.05	.01	.21	.09	.16	.10
Subjective valuing of reading achievement	.18	.05	.13	.11	.31	.20
Subjective overall valuing of achievement	.28	.07	.12	.10	.32	.15
Locus of control	.19	.07	.09	.05	.23	.02
Overall scholastic competence (Harter Scale)	.15	.07	.16	.12	.42	.18
Student's prosocial behaviors	.07	.09	.06	.01	.01	.04
Student's ethnicity (Hispanic vs. White)	.20	.11	.15	.14	.35	.13
Student's conduct problems	.03	.03	.15	.03	.05	.05
Student's economic disadvantage status	.05	.01	.25	.14	.35	.21
Student's gender	.30	.16	.25	.07	.03	.00
Student's hyperactivity-teacher	.09	.18	.14	.02	.18	.21
Student's conduct problems-parent	.06	.19	.09	.02	.18	.23

Note. SMD is absolute standardized mean difference. Before is before propensity score equating (preexisting 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 outcomes (without propensity score correction).

Biographies

Myung Hee IM, PhD, is a researcher at the American Institute for Research, 1000 Thomas J. Jefferson Street NW, Washington, DC 20007; myunghee.im@gmail.com. Her research

interests include measurement invariance and latent growth analysis under multilevel modeling and structural equation modeling. She has conducted large-scale longitudinal data analyses to investigate the impact of grade retention on students' academic achievement and engagement.

Jan N. Hughes, PhD, is professor emerita at Texas A&M University. Her research interests include peer relationships, teacher-student relationships, social and emotional development, and academic achievement. She has conducted large-scale longitudinal studies to investigate risk and protective processes in children and adolescents.

Qian Cao is a doctoral student in the Research, Measurement, and Statistics Program in the Department of Educational Psychology at Texas A&M University. She holds the position of graduate research assistant. Her research interests include longitudinal data analyses, structural equation modeling, and measurement invariance.

Oi-Man Kwok, PhD, holds the position of professor in the Department of Educational Psychology at Texas A&M University and teaches in the Learning Sciences Program. His research interests include structural equation modeling and longitudinal data analysis.

References

- Alexander KL, Entwisle DR, Kabbani NS. The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record*. 2001; 103:760–822. DOI: 10.1111/0161-4681.00134
- Asparouhov T. Sampling weights in latent variable modeling. *Structural Equation Modeling*. 2005; 12:411–434. DOI: 10.1207/s15328007sem1203_4
- Blomfield C, Barber B. Australian adolescents' extracurricular activity participation and positive development: Is the relationship mediated by peer attributes? *Australian Journal of Educational & Developmental Psychology*. 2010; 10:114–128.
- Bohnert A, Fredricks J, Randall E. Capturing unique dimensions of youth organized activity involvement: Theoretical and methodological considerations. *Review of Educational Research*. 2010; 80:576–610. DOI: 10.3102/00346543533
- Breiman L. Random forests. *Machine Learning*. 2001; 45:5–32. DOI: 10.1023/A:1010933404324
- Broh BA. Linking extracurricular programming to academic achievement: Who benefits and why? *Sociology of Education*. 2002; 75:69–91.
- Bronfenbrenner U, Morris PA. The bioecological model of human development. In: Bronson MB, editor *Self-regulation in early childhood: Nature and nurture*. New York, NY: Guilford Press; 2006. 793–828.
- Brown BB, Larson J. Contextual influences on adolescent development. In: Lerner RM, Steinberg L, editors *Handbook of adolescent psychology*. 3rd. Vol. 2. Hoboken, NJ: Wiley; 2009. 74–103.
- Cham H, West SG, Hughes JN, Im M. Effect of retention in elementary grades on grade 9 motivation for educational attainment. *Journal of School Psychology*. 2015; 53:7–24. [PubMed: 25636258]
- Denault A, Poulin F. Intensity and breadth of participation in organized activities during the adolescent years: Multiple associations with youth outcomes. *Journal of Youth and Adolescence*. 2009a; 38:1199–1213. DOI: 10.1007/s10964-009-9437-5 [PubMed: 19669900]
- Denault A, Poulin F. Predictors of adolescent participation in organized activities: A five-year longitudinal study. *Journal of Research on Adolescence*. 2009b; 19:287–311. DOI: 10.1111/j.1532-7795.2009.00597.x
- Denault A, Poulin F, Pedersen S. Intensity of participation in organized youth activities during the high school years: Longitudinal associations with adjustment. *Applied Developmental Science*. 2009; 13:74–87. DOI: 10.1080/10888690902801459

- Donegan B. The linchpin year. *Educational Leadership*. 2008; 65:54–57. DOI: 10.1177/1555458908317893
- Dunn JS, Kinney DA, Hofferth SL. Parental ideologies and children's afterschool activities. *American Behavioral Scientist*. 2003; 46:1359–1386. DOI: 10.1177/0002764203046010006
- Eccles JS, Barber BL. Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research*. 1999; 14:10–43. DOI: 10.1177/0743558499141003
- Eccles JS, Roeser RW. Schools, academic motivation, and stage-environment fit. In: Lerner RM, Steinberg L, editors *Handbook of adolescent psychology*. 3rd. Hoboken, NJ: John Wiley & Sons; 2009. 404–434.
- Eccles JS, Wigfield A, Midgley C, Reuman D. Negative effects of traditional middle schools on students' motivation. *Elementary School Journal*. 1993; 93:553–574. DOI: 10.1086/461740
- Farb AF, Matjasko JL. Recent advances in research on school-based extracurricular activities and adolescent development. *Developmental Review*. 2012; 32:1–48. DOI: 10.1016/j.dr.2011.10.001
- Feldman AF, Matjasko JL. The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research*. 2005; 75:159–210. DOI: 10.3102/00346543075002159
- Feldman AF, Matjasko JL. Profiles and portfolios of adolescent school-based extracurricular activity participation. *Journal of Adolescence*. 2007; 30:313–332. DOI: 10.1016/j.adolescence.2006.03.004 [PubMed: 16678248]
- Fredricks JA, Eccles JS. Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations. *Developmental Psychology*. 2006; 42:698–713. DOI: 10.1037/0012-1649.42.4.698 [PubMed: 16802902]
- Fredricks JA, Eccles JS. Participation in extracurricular activities in the middle school years: Are there developmental benefits for African American and European American youth? *Journal of Youth and Adolescence*. 2008; 37:1029–1043. DOI: 10.1007/s10964-008-9309-4
- Fredricks JA, Simpkins SD. Promoting positive youth development through organized after-school activities: Taking a closer look at participation of ethnic minority youth. *Child Development Perspectives*. 2012; 6:280–287. DOI: 10.1111/j.1750-8606.2011.00206.x
- Fredricks JA, Simpkins SD. Organized out-of-school activities and peer relationships: Theoretical perspectives and previous research. In: Fredricks JA, Simpkins SD, editors *Organized Out-of-School Activities: Settings for Peer Relationships*. New Directions for Child and Adolescent Development. Vol. 140. 2013. 1–17.
- Gardner M, Roth J, Brooks-Gunn J. Adolescents' participation in organized activities and developmental success 2 and 8 years after high school: Do sponsorship, duration, and intensity matter? *Developmental Psychology*. 2008; 44:811–830. DOI: 10.1037/0012-1649.44.3.814
- Garner JF. A phenomenology of nonparticipation in extracurricular activities by Hispanic middle school students. *Dissertation Abstracts International Section A: Humanities and Social Sciences*. 2013 Order No. AAI3516819
- Guest AM, McRee N. A school-level analysis of adolescent extracurricular activity, delinquency, and depression: The importance of situational context. *Journal of Youth and Adolescence*. 2008; 38:51–62. DOI: 10.1007/s10964-008-9279-6 [PubMed: 19636791]
- Hirano K, Imbens GW, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*. 2003; 71:1161–1189. DOI: 10.1111/1468-0262.00442
- Ho DE, Imai K, King G, Stuart EA. Matching as nonparametric pre-processing for reducing model dependence in parametric causal inference. *Political Analysis*. 2007; 15:199–236. DOI: 10.1093/pan/mpl013
- 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. DOI: 10.1037/0022-0663.100.1.1 [PubMed: 19578558]
- Hull P, Kilbourne B, Reece M, Husaini B. Community involvement and adolescent mental health: Moderating effects of race/ethnicity and neighborhood disadvantage. *Journal of Community Psychology*. 2008; 36:534–551. DOI: 10.1002/jcop.20253

- 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. DOI: 10.1111/j.1540-4560.2008.00546.x
- Knifsend CA, Juvonen J. Social identity complexity, cross-ethnic friendships, and intergroup attitudes in urban middle school. *Child Development*. 2014; 85:709–721. DOI: 10.1111/cdev.12157 [PubMed: 24032401]
- Larson RW, Hansen DM, Moneta G. Differing profiles of developmental experiences across types of organized youth activities. *Developmental Psychology*. 2006; 42:849–863. DOI: 10.1037/0012-1649.42.5.849 [PubMed: 16953691]
- Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. *Statistics in Medicine*. 2010; 29:337–346. DOI: 10.1002/sim.3782 [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. DOI: 10.1037/0022-0663.98.3.567
- Lugaila TA. A child's day: 2000 (selected indicators of child well-being). 2003. Retrieved from <https://www.census.gov/prod/2003pubs/p70-89.pdf>
- Mahoney JL, Cairns RB. Do extracurricular activities protect against early school dropout? *Developmental Psychology*. 1997; 33:241–253. DOI: 10.1037/0012-1649.33.2.241 [PubMed: 9147833]
- Marsh HW, Kleitman S. Extracurricular school activities: The good, the bad, and the nonlinear. *Harvard Educational Review*. 2002; 72:464–514. DOI: 10.17763/haer.72.4.051388703v7v7736
- Muthén LK, Muthén BO. Mplus user's guide. 7th. Los Angeles, CA: Muthén & Muthén; 1998–2014.
- National Center for Education Statistics. Digest of educational statistics. 2012. Retrieved from http://nces.ed.gov/programs/digest/2012menu_tables.asp
- Peguerro AA. A profile of Latino school-based extracurricular activity involvement. *Journal of Latinos and Education*. 2010; 9:60–71. DOI: 10.1080/15348430903253076
- Ream RK, Rumberger RW. Student engagement, peer social capital, and school dropout among Mexican American and non-Latino White students. *Sociology of Education*. 2008; 81:109–139. DOI: 10.1177/003804070808100201
- Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*. 1984; 79(387):516–524. DOI: 10.1080/01621459.1984.10478078
- 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. DOI: 10.1023/A:1020363010465
- Schafer JL, Kang J. Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods*. 2008; 13:279–313. DOI: 10.1037/a0014268 [PubMed: 19071996]
- Shadish WR, Cook TD, Campbell DT. Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton-Mifflin; 2002.
- Shannon CS. Parents' messages about the role of extracurricular and unstructured leisure activities: Adolescent' perceptions. *Journal of Leisure Research*. 2006; 38:398–420.
- Simpkins SD, O'Donnell M, Delgado MY, Becnel JN. Latino adolescents' participation in extracurricular activities: How important are family resources and cultural orientation? *Applied Developmental Science*. 2011; 15:37–50. DOI: 10.1080/10888691.2011.538618
- Simpkins SD, Vest AE, Delgado MY, Price CD. Do school friends participate in similar extracurricular activities?: Examining the moderating role of race/ethnicity and age. *Journal of Leisure Research*. 2012; 44:332–352.
- Skinner EA, Zimmer-Gembeck MJ, Connell JP. Individual differences and the development of perceived control. *Monographs of the Society for Research in Child Development*. 1998; 63:1–231. DOI: 10.2307/1166220
- Sonnenschein S, Stapleton LM, Benson A. The relation between the type and amount of instruction and growth in children's reading competencies. *American Educational Research Journal*. 2010; 47:358–389. DOI: 10.3102/0002831209349215

- Strobl C, Boulesteix AL, Kneib T, Augustin T, Zeileis A. Conditional variable importance for random forests. *BMC Bioinformatics*. 2008; 9doi: 10.1186/1471-2105-9-307
- Stuart EA. Matching methods for causal inference: A review and a look forward. *Statistical Science*. 2010; 25:1–21. DOI: 10.1214/09-STS313 [PubMed: 20871802]
- Towe CR. Hispanic students' connection to school: The relation between extracurricular participation and grade point average. *Dissertation Abstracts International Section A: Humanities and Social Sciences*. 2012:3709.
- Villarreal V. Characteristics and consequences of extracurricular activity participation of Hispanic middle school students. *Dissertation Abstracts International Section A: Humanities and Social Sciences*. 2013 (Order No. AAI3532241).
- 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; 82:906–919. DOI: 10.1037/a0036387 [PubMed: 24708350]
- West SG, Finch JF, Curran PJ. Structural equation models with non-normal variables: Problems and remedies. In: Hoyle RH, editor *Structural equation modeling: Concepts, issues, and applications*. Thousand Oaks, CA: Sage Publications; 1995. 56–75.
- 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. 463–478.
- Wigfield A, Eccles JS, Yoon KS, Harold RD, Arbretton AJA, Freedman-Doan C, Blumenfeld PC. Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*. 1997; 89:451–469. DOI: 10.1037/0022-0663.89.3.451
- Woodcock RW, McGrew KS, Mather N. *Woodcock-Johnson III tests of achievement*. Riverside, CA: Riverside Publishing; 2001.
- Woodcock RW, Muñoz-Sandoval AF. *Woodcock-Muñoz language survey, Spanish form*. Riverside, CA: Riverside Publishing; 1993.
- Woodcock RW, Muñoz-Sandoval AF, McGrew K, Mather N, Schrank F. *Bateria III Woodcock-Munoz*. Riverside, CA: Riverside Publishing; 2004.

Figure 1.

The boxplots of propensity scores between participation and nonparticipation students before and after propensity score equating using the weighted odds method. parti = participation group; non-parti = nonparticipating group.

Table 1

List of Covariates for Propensity Score Analyses

No.	Covariates	Source
1.	Student's gender	Archival
2.	Student's ethnicity (Hispanic vs. White)	Archival
3.	Student's ethnicity (African American vs. White)	Archival
4.	Student enrolled in bilingual class	Archival
5.	Student enrolled in bilingual class in first grade	Archival
6.	Student's economic disadvantaged status	Archival
7.	Highest adult employment level in the household	Parent
8.	Highest adult educational level in household	Parent
9.	Educational aspirations for student	Parent
10.	Student's emotional symptoms	Parent
11.	Student's conduct problems	Parent
12.	Student's hyperactivity	Parent
13.	Student's peer problems	Parent
14.	Student's prosocial behaviors	Parent
15.	Student's dominant language	Performance
16.	Woodcock-Johnson III Broad Reading W Score	Performance
17.	Woodcock-Johnson III Broad Math W Score	Performance
18.	Teacher-student warmth	Student
19.	Math competence beliefs	Student
20.	Reading competence beliefs	Student
21.	Sports competence beliefs	Student
22.	Subjective valuing of math achievement	Student
23.	Subjective valuing of reading achievement	Student
24.	Overall academic competence	Student
25.	Subjective overall valuing of achievement	Student
26.	Subjective valuing of sports	Student
27.	Overall scholastic competence (Harter Scale)	Student
28.	Social competence (Harter Scale)	Student
29.	Athletic competence (Harter Scale)	Student
30.	School belonging	Student
31.	Classroom performance approach goal structure	Student
32.	Classroom performance avoidance goal structure	Student
33.	Friends' prosocial orientation	Student
34.	Student's victimization in classroom	Student
35.	Locus of control	Student
36.	Peer affiliation structured activities	Student
37.	Educational aspirations for student	Teacher
38.	Academic performance in classroom	Teacher
39.	Student's conduct problems	Teacher

No.	Covariates	Source
40.	Student's hyperactivity	Teacher
41.	Student's prosocial behavior	Teacher
42.	Home-school relationship (alliance)	Teacher
43.	Parental school involvement	Teacher
44.	Classroom behavioral engagement	Teacher

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2Frequency of Extracurricular Participation Status by Gender and Ethnicity ($N = 483$)

	Sport						Performing Art/Clubs					
	Grade			Duration Patterns			Grade			Duration Patterns		
	7	8	Con.	Delay.	Discon.	None	7	8	Con.	Delay.	Discon.	None
African American, n	57	63	37	26	20	39	30	39	14	25	16	67
Percentage ^a ($n = 122$)	47	52	30	21	16	32	25	32	11	20	13	55
Latino, n	79	73	40	33	39	74	46	51	19	32	27	108
Percentage ^a ($n = 186$)	42	39	22	18	21	40	25	27	10	17	15	58
Euro-American, n	98	90	74	16	24	46	45	61	25	36	20	79
Percentage ^a ($n = 160$)	61	56	46	10	15	29	28	38	16	23	13	49
Other, n	8	6	5	1	3	6	6	10	5	5	1	4
Percentage ^a ($n = 15$)	53	40	33	7	20	40	40	67	33	33	7	27
Total n	242	232	156	76	86	165	127	161	63	98	64	258
Percentage ($n = 483$)	50	48	32	16	18	34	26	33	13	20	13	53
Girls, n	95	90	57	33	38	91	62	87	32	55	30	102
Percentage ^b ($n = 219$)	43	41%	26%	15%	17%	42%	28%	40%	15%	25%	14%	47%
Boys, n	147	142	99	43	48	74	65	74	31	43	34	156
Percentage ^b	56	54	38	16	18	28	25	28	12	16	13	59
Total n	242	232	156	76	86	165	127	161	63	98	64	258
Percentage ($n = 483$)	50	48	32	16	18	34	26	33	13	20	13	53

Note. Duration patterns are the four different patterns in participation across Grades 7 and 8. Con. = students participated continuously for both years; Delay. = students participated continuously only in Grade 8; Discon. = students participated discontinuously only in Grade 7; and None = student did not participate in either year.

^aProportion of participation in each ethnicity.

^bProportion of participation in each gender.

Correlations and Descriptive Statistics for Outcomes at Grade 9 and Baseline Variables of Outcomes at Grade 5 ($N = 483$)**Table 3**

Correlation	1	2	3	4	5	6	7	8
1 S_Competence Belief 9	—							
2 S_Valuing of Education 9	0.52*	—						
3 T_Classroom Engagement 9	0.26*	0.22*	—					
4 T_Letter Grades 9	0.18*	0.17*	0.65*	—				
5 S_Competence Belief 5	0.21*	0.12*	0.00	0.09	—			
6 S_Valuing of Education 5	0.20*	0.15*	0.03	0.09	0.65*	—		
7 T_Engagement 5	0.08	0.05	0.29*	0.23*	0.12*	0.16*	—	
8 Achievement 5	0.12*	0.04	0.12*	0.18*	0.24*	0.14*	0.31*	—
Descriptive statistics								
Mean	4.09	3.94	2.72	2.43	21.79	23.71	2.81	95.90
Standard deviation	0.60	0.65	0.64	1.29	4.69	4.73	0.68	13.67
Missing (%)	6	5	22	24	4	4	19	4

Note. Number at the end of each variable indicates the student's grade when the variable was measured. Achievement is Woodcock-Johnson III Broad Reading score. The correlations are bivariate correlations based on listwise deletion. S_ = student reported; T_ = teacher reported.

* $p < .05$ (two-tailed).

Standardized Effect of Different Pattern of Duration by Activity Type During Middle School (Grades 7 and 8) on Grade 9 Outcomes

Table 4

Outcome Effect	S Competence Belief 9			S Valuing of Education 9			T Classroom Engagement 9			T Letter Grades 9		
	Est.	SE	p	Est.	SE	p	Est.	SE	p	Est.	SE	p
Sport												
Continuous (<i>n</i> = 156)	.14*	.04	<.001	.16*	.05	.003	.04	.07	.505	.07	.05	.200
Delayed (<i>n</i> = 76)	.09	.05	.057	.15*	.05	.004	.05	.04	.262	.06	.05	.169
Discontinued (<i>n</i> = 86)	.05	.03	.073	.05	.06	.346	.00	.06	.942	-.01	.05	.909
Performance arts/clubs												
Continuous (<i>n</i> = 63)	.17*	.03	<.001	.00	.04	.970	.21*	.05	<.001	.16*	.04	<.001
Delayed (<i>n</i> = 98)	.10*	.04	.019	.03	.06	.530	.10*	.04	.020	.06*	.03	.020
Discontinued (<i>n</i> = 64)	.09*	.04	.040	-.01	.05	.870	.01	.03	.750	-.03	.04	.460

Note. Number at the end of each variable indicates the student's grade when the variable was measured. Competence belief is academic competence belief. Classroom engagement is student's behavioral classroom engagement. Letter grade is student's letter grade indicating higher value as higher score. Parameter indicates the standardized mean difference from reference group (i.e., nonparticipation group, *n* = 165 for sport and *n* = 258 for performance arts/clubs). We used a weighted regression model with the weighted propensity score. We controlled for the corresponding baseline variable on each of the outcomes. TYPE=COMPLEX with school ID at Grade 9 for CLUSTER function was used to adjust the data dependency. LISTWISE deletion was utilized for the use of WEIGHT function. S_ = student-reported; T_ = teacher-reported; Est. = standardized parameter estimates; SE = the corresponding standardized standard error.

* *p* < .05.