J Youth Adolescence (2016) 45:2260–2277
DOI 10.1007/s10964-016-0508-0

EMPIRICAL RESEARCH

Indirect Effects of Extracurricular Participation on Academic Adjustment Via Perceived Friends' Prosocial Norms

Jan N. Hughes[1] · Qian Cao[1] · Oi-man Kwok[1]

Received: 23 March 2016 / Accepted: 19 May 2016 / Published online: 14 June 2016

Abstract Students who participate in extracurricular activities in middle school exhibit higher levels of academic motivation and achievement, including graduation from high school. However, the mechanisms responsible for these beneficial effects are poorly understood. Guided by the bioecological models of development, this study tested the indirect effects of participation in grade 8 in school sports or performance arts and clubs on grade 9 academic achievement, academic competence beliefs, and school belonging, via adolescents' perceptions of their friends' prosocial norms. Participants were 495 (45 % female) ethnically diverse students (mean age at grade 8 = 13.9 years; SD = .58) who were recruited into a longitudinal study on the basis of below average literacy in grade 1. Using weighted propensity score analyses to control for potential confounders, results of longitudinal SEM found indirect effect of participation in sports, but not of participation in performance arts and clubs, on grade 9 outcomes noted above. Implications of findings for improving educational attainment of at-risk youth are discussed.

Keywords Propensity score analysis · Extracurricular activities · Friends · Academic achievement · Middle school · Longitudinal/developmental

Introduction

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. Dropping out of school is viewed as a gradual process of disengagement from school that begins in elementary grades and increases in secondary school (Alexander et al. 1997; Janosz et al. 2008). Participation in school-based, extracurricular activities such as sports and band are forms of school engagement associated with higher academic motivation and attainment, including school completion (Finn 1989; Fredricks and Eccles 2008). However, there is a dearth of research on the mechanisms responsible for the beneficial effects of extracurricular participation on academic outcomes. An understanding of processes that account for beneficial effects of participation would permit more focused efforts to enhance these effects. The purpose of this study is to test whether changes in perceptions of one's friends' prosocial norms accounts for the effects of participation in grade 8 on grade 9 academic outcomes (i.e., student-perceived academic competency beliefs and school belonging as well as teacher-reported classroom engagement and grades).

Participation in extracurricular activities is common among adolescents in the United States. Approximately 75 % of adolescents in grades 7–12 participate in at least one school-based extracurricular activity (Feldman and Matjasko 2007). Extracurricular activities include a wide range of specific activities, including team and individual

Electronic supplementary material The online version of this article (doi:10.1007/s10964-016-0508-0) contains supplementary material, which is available to authorized users.

✉ Jan N. Hughes
jhughes@tamu.edu

Qian Cao
cqianhome@gmail.com

Oi-man Kwok
omkwok@tamu.edu

[1] Department of Educational Psychology, Texas A&M University, 4225 TAMU, College Station, TX 77843-4225, USA

sports, drama, music, student government, and academic clubs. Consistent with bio-ecological models of development (Bronfenbrenner and Morris 2006), extracurricular activities are viewed as important contexts for development, and the specific experiences occurring within these activities (e.g., interacting with peers and adult leaders, following rules and routines, setting and monitoring goal performance, and confronting and overcoming challenges) are credited with the impact of participation on development (Feldman and Matjasko 2005).

Extracurricular activities such as sports and band are often structured in ways that facilitate high quality peer interactions and the development of friendships (Fredricks and Simpkins 2012). Furthermore, youth state that the opportunity to make friends and to interact with friends is a major reason for participating in sports, band, and other activities (Denault and Poulin 2009b; Simpkins et al. 2012). Importantly, friendships among youth who participate in the same extracurricular context are more likely to be maintained than are friendships with peers who are not co-participants (Schaefer et al. 2011), perhaps due to the regular participation schedules of organized activities (Fredricks and Simpkins 2013).

Furthermore, adolescents who are involved in extracurricular activities are, on average, more academically oriented and prosocial in their orientation than are youth who do not participate in these activities and more likely to report having academically-oriented friendship groups (Fredricks and Eccles 2008). Based on these findings, researchers have suggested that the access to a prosocial peer group that participation provides may account for its academic benefits (Fredricks and Simpkins 2012). Despite the strength of this reasoning, few studies have investigated whether one's friends' characteristics account for effects of extracurricular activity (Blomfield and Barber 2010; Eccles and Barber 1999; Fredricks and Eccles 2005), and no study has employed longitudinal designs that control for prior levels of friends' characteristics or participants' prior performance on the outcomes of interest. Because research documents that participants and nonparticipants differ on a number of demographic, behavioral, and academic variables *prior to participation* (Beal and Crockett 2010; Feldman and Matjasko 2007), longitudinal designs that control for differences in one's friends characteristics and academic functioning that existed *prior* to participation would offer stronger evidence of causal relationships (Denault and Poulin 2009b).

The present study investigates the indirect effects of participation in school-sponsored extracurricular programs in grade 8 on academic outcomes the following year, via its effects on friends' characteristics. Based on the premise that school-sponsored activities are more likely than community-sponsored activities to promote identification with school and its values and norms, including achievement (Finn 1989), the current study focuses on participation in school-sponsored activities. Importantly, the study uses propensity score analyses to equate participant and non-participant groups on a comprehensive set of variables measured prior to participation, thereby reducing potential confounders.

Adolescents' Friends and Academic Functioning

Adolescence is a time of high susceptibility to the influence of one's peers (Brown and Larson 2009). Early adolescents spend increasing amounts of time interacting with peers, increasingly look to friends for validation of their sense of worth, and adopt friends' values and behaviors (Hardway and Fuligni 2006). An extensive body of literature confirms that the similarity between adolescents and their friends on a number of characteristics, including academic orientation and deviant behavior, is a result of both socialization and selection effects (for review see Wentzel et al. 2012). That is, youth both gravitate toward peers whom they perceive as similar to them and as sharing their values and goals (selection effects), and become more similar to their friends over time (socialization effects). In a reciprocal fashion, selection and socialization effects account for the high and increasing similarity between adolescents and the peer groups with whom they interact and identify (Brown and Larson 2009).

Extensive longitudinal research documents effects of one's friends' prosocial and deviant behaviors on adolescents' academic engagement and achievement (Kindermann 2007; Lynch et al. 2013). For example, in a study of middle school students (Berndt and Keefe 1995), students who perceived their friends as positively engaged in school increased in their own positive school involvement from the fall to the spring of the year. Examining the influences of friends' positive engagement (e.g., cooperate with teachers, complete homework and assignments on time) and problem behaviors (smoking cigarettes, lying to one's parents about where you have been or whom you were with) at grade 6 on academic achievement at grade 8, Véronneau and Dishion (2011) found that these two constructs were moderately negatively correlated (−.47). Importantly, friends' positive school engagement predicted improved academic achievement at grade 8, whereas friends' problem behaviors predicted lower academic achievement. In a study of high school students, adolescents whose friends engaged in antisocial behavior were at greater risk of not graduating from school, an effect that was partially mediated by the association between friends' antisocial behavior and the student's lower academic engagement and achievement (Wang and Eccles 2012).

Consistent with channeling theory (Martin et al. 2003), extracurricular activities may socialize a youth's academic

motivation and behavior by channeling them into peer groups that promote identification with school and positive school engagement (Fredricks and Simpkins 2013). The peer experiences in extracurricular activities, relative to experiences in school, involve more teamwork, positive peer interactions, and opportunities for self-expression (Fredricks et al. 2010; Larson et al. 2006). Given these differences in peer processes, and the finding that youth who participate in extracurricular activities tend to be more prosocial and academically oriented than nonparticipants, extracurricular activities are well designed to promote the selection and maintenance of friendships with well-adjusted peers (Fredricks and Simpkins 2013). In addition to promoting friendships, participation creates a peer group culture of shared values and goals (Brown and Larson 2009). Peer groups may influence adolescents' behaviors through reinforcement and punishment for adhering to or rejecting peer group norms (Sage and Kindermann 1999). Recent research on differences in the effects of friends and peer groups suggests that they have similar effects on students' academic engagement, and that the effect of friends with whom one interacts frequently (e.g., in extracurricular activities) may be particularly strong (Kindermann and Skinner 2012).

Middle School and Extracurricular Participation

Despite a large body of research documenting benefits of extracurricular participation at the high school level, few studies have investigated effects of extracurricular participation during the critical middle school grades on students' academic outcomes (see review by Farb and Matjasko 2012; for exception see Chambers and Schreiber 2004; Fredricks and Eccles 2008), and no study has examined the effect of participation on one's friends' characteristics during this developmental period. Because early adolescence is a time of heightened influence of peers (Brown and Larson 2009), and participation in school-based activities becomes widely available in middle school (Fredricks and Eccles 2008, Chambers and Schreiber 2004), it is important to understand the effect of extracurricular participation in middle school on youths' friends' characteristics.

Extracurricular Activity Type and Benefits of Participation

The typical middle and high school offers a range of extracurricular activities, including various team and individual sports, performance arts such as music and arts, and academic or service clubs (Chambers and Schreiber 2004). Although students who participate in each type of activity tend to have higher levels of prosocial behavior and school engagement, relative to non-participants (Eccles and Barber 1999; Mahoney 2000), the specific benefits of participation may differ across activity types. For example, participation in performance arts (e.g., theater, choir, band) and academic and service clubs is more consistently related to higher grades and academic values than is participation in sports (Denault and Poulin 2009a; Fredricks and Eccles 2005, 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). Sports participation has also been associated with higher levels of alcohol use and other risky behavior (Denault and Poulin 2009a; Fredricks and Eccles 2005).

Untangling Selection and Socialization Effects

Students are not randomly assigned to participation; rather, students select, or are recruited into, these activities. Multiple student, family, and school variables are associated with selection into participation versus non-participation, and many of these variables are also associated with the measured outcomes (Feldman and Matjasko 2007). For example, students' educational aspirations and friends' characteristics predict future participation in extracurricular participation as well as students' subsequent educational attainment (Beal and Crockett 2010).

Due to potential selection effects, a finding that participants and non-participants differ at some future point on an outcome (e.g., friends' characteristics, educational aspirations, or academic achievement) may be due to these pre-existing differences rather than to participation. The most common strategy to minimize selection effects has been covariate analyses, in which the effects of a limited number of potential confounders are statistically controlled (Fredricks and Eccles 2006). However, these statistical adjustments can employ a limited number of observed covariates that may not capture all of the pre-existing differences between participants and non-participants and present additional statistical challenges (Shadish, Cook, and Campbell 2002).

Propensity Score Analysis

A propensity score is defined as the conditional probability of receiving treatment (in our case extracurricular participation), given a vector of observed covariates (Rosenbaum and Rubin 1984). Propensity scores generate a single index-the propensity score-that summarizes information across the covariates (i.e., potential confounders). The selection of a comprehensive set of covariates is essential to the effectiveness of propensity score analysis in removing bias (Hong and Raudenbush 2006). Procedures

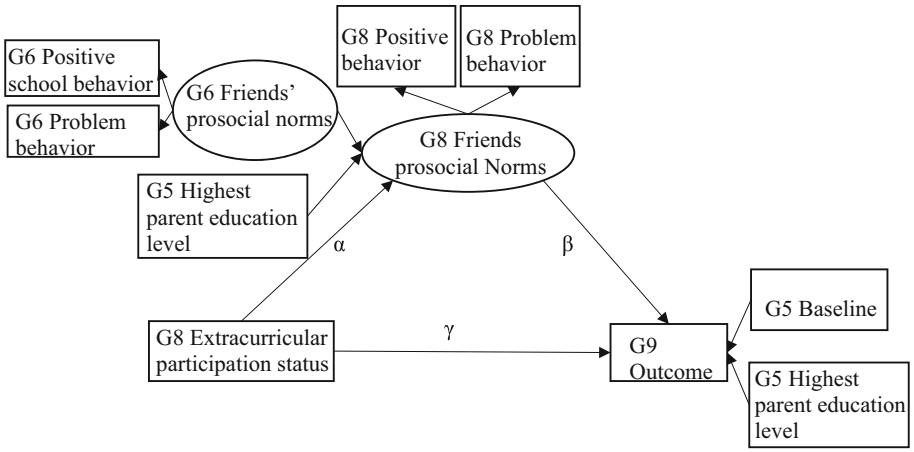

Fig. 1 The hypothesized mediation model. Extracurricular participation status refers to one of two activity domains (i.e., sports and performance arts or clubs) in grade 8. Separate analyses were conducted for each of four grade 9 outcome variables: competence beliefs, education value, behavioral engagement, and course grades. All analyses controlled for the baseline measures of perceived friends' prosocial norms in grade 6 and baseline measure of the outcome variables in grade 5. For the sports model only, grade 5 highest parent education level was a covariate for perceived friends' prosocial norms and all outcomes. G5, G6, and G8 are grade 5, 6, and 8, respectfully

such as matching and weighting can then be used to equate the treatment group (i.e., participating students) and control group (i.e., non-participating students) on their propensity scores (West et al. 2014). Given successful equating is achieved on all confounding variables, the propensity score analysis produces an unbiased estimate of the average effect of participation on students.

To the authors' knowledge, only one published study has employed propensity score analysis to test the effect of extracurricular participation on students' academic outcomes. Utilizing the same longitudinal sample as the current study, Im et al. (2015) used propensity score analysis to investigate the effect of two broad domains of extracurricular activities (i.e., sports and performance arts or clubs) in grades 7 and 8 on students' grade 9 academic motivation and achievement. Participation in sports predicted students' grade 9 competence beliefs and valuing of education, and participation in performance arts and clubs predicted students' grade 9 competence beliefs and teacher-rated classroom engagement and letter grades. The current study extends the Im et al. study by investigating whether changes in adolescents' perceived friends' prosocial norms account for the effects of participation. An understanding of the role of friends in accounting for the benefits of extracurricular participation would have implications for enhancing benefits of participation.

Gender Differences

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 and Poulin 2009b; Eccles and Barber 1999; Fredricks and Eccles 2008). Despite gender differences in activity contexts, the effects of extracurricular participation are generally similar for boys and girls (Im et al. 2015; Fredricks and Eccles 2006, 2008). Girls and boys also differ in their friends' school engagement. Although girls' friends are more positively engaged in school than are boys' friends (Véronneau and Dishion 2010; Wang and Eccles 2012), the effects of friends' school engagement on academic and behavioral outcomes are generally similar for boys and girls (Cotter and Smokowski 2016; Véronneau et al. 2010). This study is the first to test gender differences in the mechanisms responsible for effects of participation on academic outcomes.

Hypotheses

Based on the preceding theoretical considerations and empirical findings, we test a model positing indirect effects of participation in two broad domains of school-based extracurricular activities (sports and performance arts and clubs) in grade 8 on grade 9 academic outcomes (i.e., competence beliefs and valuing of education, teacher- rated classroom engagement, and academic grades), via effects of participation on friends' prosocial norms. The hypothesized model is depicted in Fig. 1. By equating participant groups at baseline on a large number of measured covariates that are associated with participation in sports or performance arts or clubs and the outcome variables, and controlling for students' baseline performance on both the

outcome and the mediator (i.e., perceived friends' prosocial norms), the study provides a strong basis for estimating the hypothesized indirect effects. Based on lack of prior research, analyses of gender moderation of the hypothesized indirect effects are exploratory.

Participants were recruited into a longitudinal sample when in first grade based on academic risk and are ethnically diverse and predominantly from low-income families. Although academically at-risk youth and low income youth are less likely to participate in extracurricular activities than are lower-risk youth (Fredricks and Simpkins 2012), the positive effects of participation may be stronger for high risk youth for whom other opportunities to form positive connections with the school and between home and may be less available (Mahoney and Cairns 1997; Mahoney 2000). Thus study findings may inform policies designed to reduce income and ethnic disparities in educational attainment.

Methods

Participants

Participants were 495 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 activities were collected when these students were in grade 8 (typically academic year 2008–2009 or 2009–2010). 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, 37 % 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 = 7424) had an ethnic distribution of 67 % White, 12 % Hispanic, 12 % African American, and 9 % other. Additional inclusionary criteria for the larger study included speaking English or Spanish and, not receiving special education services other than speech and language services, and not having been previously retained in grade 1.

At the end of the first 5 years of participation in the study, parental consent for continued participation was received for 569 of the 784 participants. Almost all non-consent was due to non-response. Of these 569 who re-consented at Year 5, 495 met criteria for inclusion in the current study, which included data on participation status at grade 8 and at least one outcome measure at grade 9. Attrition analyses found no differences between the 495 participants and the 289 attrited participants 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. Additional attrition analyses found no evidence of selective attrition between grade 5 and grade 9. The 495 participants (54.7 % male) were 13.9 years of age ($SD = 0.58$) at grade 8; 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 % Hispanic (of whom 32.0 % were enrolled in bilingual education at grade 5), and 3.1 % Other. At grade 8, participants' mean reading age-standard scores from the Woodcock-Johnson III (Woodcock et al. 2001) or its Spanish-language equivalent (Batería III Woodcock-Munoz (Woodcock et al. 2004) was 97.09 ($SD = 14.68$). Participants were enrolled in 69 schools during grade 8 and 72 schools in grade 9.

Assessment Overview

All student-report measures and reading achievement were assessed in individual interviews at school between November and April of the given year. 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 and Muñoz-Sandoval 1993). Extracurricular participation was assessed in grade 8, and perceived friends' prosocial norms were assessed in grade 6 and 8. Covariates used in the propensity score analyses were assessed in grade 4 or 5.

Outcomes were assessed at baseline (grade 5) and grade 9. 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 and math achievement were assessed on an individually administered test. Teacher- reported data were obtained in the Spring of the year via questionnaires that were mailed to teachers and for which teachers received a small payment. Student-reported data and measures of reading achievement were obtained in individual interviews and testing sessions at school between November and April. For teacher-rated engagement and reading achievement, the outcome measure was the same at baseline and grade 9. For teacher-awarded grades (which were typically provided by the language teacher), the baseline measure was the score on the 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 baseline and grade 9.

Measures

Extracurricular Participation

In individual interviews at school in the Spring of grade 8, the interviewer asked students to indicate if they participated that year in each of four school-sponsored activity contexts: (a) sports; (b) performance 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). 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 and Matjasko 2007), these activities were combined with performance arts and music into a performance arts or clubs category. Participation in each broad activity category (sports and performance arts or clubs) was defined as a dichotomous variable (0 = did not participate and 1 = did participate).

Competence Beliefs and Valuing of Educational Attainment

At grade 9, students completed the 11-item Academic Competence and Effort Beliefs Subscale ($\alpha = .89$) and the 10-item Value of Education Subscale ($\alpha = .85$) of the Motivation for Education Attainment Questionnaire (Cham et al. 2014), a multi-dimensional measure of motivation to complete high school and pursue post-secondary 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 demonstrated good construct and criterion-related validity in an at-risk sample of grade 9 students (Cham et al. 2014).

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). Example items include "How good are you in reading?" and "How good are you in reading, relative to other children?" Youth responded on a 1-30 thermometer (1 = "not at all good and 30 = indicating "one of the best"). Youth rated their subjective valuing of reading and math by indicating how interesting, fun, or important each subject was, using a similar scale. The scale has demonstrated good criterion validity (Wigfield et al. 1997). 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 for Math.

Course Grades

Students' language arts teachers were asked to report the 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. When a language arts teacher was not available (7 % of cases), 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 et al. (1998). Example items include "tries hard to do well in school"; "participates in class discussion"; and pays attention in class." Teachers indicated 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 (Hughes et al. 2014). The internal consistency reliabilities(α) at baseline and grade 9 were .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 Batería III, the equivalent Spanish version of the WJ III (Woodcock et al. 2004) by bilingual examiners.

Perceived Friends' Positive School Behavior, Problem Behaviors, and Prosocial Norms

Students were individually interviewed and asked to name up to eight peers with whom they spend time outside of the

classroom (e.g., at lunch, before or after school, in the neighborhood). Students could name peers who did not attend their grade or school. For each peer named, students answered questions describing the peer's friendship status (i.e., close friend or "someone you just spend time with") and the peer's positive school behaviors and problem behaviors. Based on research finding stronger peer effects for peers identified as close friends compared to peers with whom one spends time but does not share a close relationship (for a review, see Berndt et al. 1999), only data concerning characteristics of close friends were used in the current analysis.

At grade 6 and grade 8, the number of close friends nominated was 3.12 (SD = 1.99), and 3.40 (SD = 1.79), respectively. The student answered "yes" or "no" to 3 questions describing positive school behaviors for each close friend (i.e., "Does friend plan to go to college?"; "Does friend get along with teachers and other adults?"; "Is the friend doing well in school?") and 4 questions describing problem behaviors of each close friend (i.e., "Does friend regularly smoke or chew tobacco?"; "Is friend often out on the town at night?"; "Has friend ever been caught by the police?"; and "Has the friend ever skipped school?"). For each positive school behavior and problem behavior item, the score was the mean percentage of one's close friends who engaged in that behavior; scores could range from 0 to 1.0. A positive school behavior score was calculated as the mean of the three positive school behavior items, and a problem behavior score was calculated as the mean of the four negative involvement items.

Measures of friends' positive school behaviors and problem behaviors are strongly negatively correlated with each other, and both predict changes in adolescents' academic achievement (Véronneau and Dishion 2011). Furthermore, problem behaviors may be viewed as indicators of disaffection from school or amotivation, which many researchers view as the opposite of positive school involvement (Green-Demers et al. 2008). Thus, to obtain a single, broad index of friends' prosocial norms, a latent perceived friends' prosocial norms construct was developed with two factors: friends' positive school behaviors and friends' problem behavior (see "Results" section for more details on the latent construct).

Covariates for Propensity Score Analysis

Forty-six covariates (potential confounders), all of which were measured in grade 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 in grade 8. These 46 covariates (listed in "Appendix" section) 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 academic functioning. 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, perceived social acceptance); teacher questionnaires (e.g., behavioral, academic, and social functioning); parent questionnaires (e.g., family demographics, educational aspirations, child behavioral and social functioning); and school records (e.g., child ethnicity, age, and gender, bilingual class placement).

Propensity Score Analysis

The first step in propensity score analysis is to estimate each student's propensity score for each of the two participation categories (i.e., conditional probability of participating in sports and the conditional probability of participating in performance arts or club category), given the student's scores on the 46 covariates. The second step is to equate the estimated propensity score distributions between participants and non-participants for each of the two extracurricular activity domains, separately. The third step is to check the balance of the distribution of the 46 covariates between participants and non-participants for each of the two extracurricular activity domains. Each step is described in detail below.

Propensity Score Estimation

We estimated two sets of propensity scores depending on two extracurricular activity domains: sports and performance arts or clubs. Specifically, the propensity score for sports was the probability of a student participating in sports in grade 8 versus not participating in sports. A total of 239 students participated in sports and 256 did not participate in sports. The propensity score for performance arts or clubs was the probability of a student participating in performance arts or clubs in grade 8 versus not participating in performance arts or clubs. A total of 164 students participated in performance arts/clubs, and 331 did not participate in performance arts/clubs. Propensity scores were estimated using the random forest method (Breiman 2001), with the R package version 3.2.0 (Strobl et al. 2008), which automatically identifies complex and nonlinear relationship of covariates with a treatment status (i.e., extracurricular participation in this study), thereby reducing bias in the estimate of the effect of a treatment on outcomes (Lee et al. 2010).

Equating

To equate participant and non-participant groups on the covariates, we applied the odds method (Schafer and Kang 2008)) to weight the propensity scores. Specifically, students who participated in a given activity domain were given a weight of 1, and those who did not participate in that domain were given a weight of $\frac{\widehat{\pi}}{1-\widehat{\pi}}$ ($\widehat{\pi}$ is estimated propensity score for non-participating students). The weighting by odds method can estimate the participation effect for students who did not actually participate compared to closely equated students who did participate. Figure 2 shows the boxplots of propensity scores between participating and non-participating students in the two activity domains (i.e., sports, performance arts or clubs) before and after equating. As expected, compared with the propensity score distributions prior to weighting, after equating the distributions were balanced in terms of the means and distributions of the propensity scores in participating and non-participating groups.

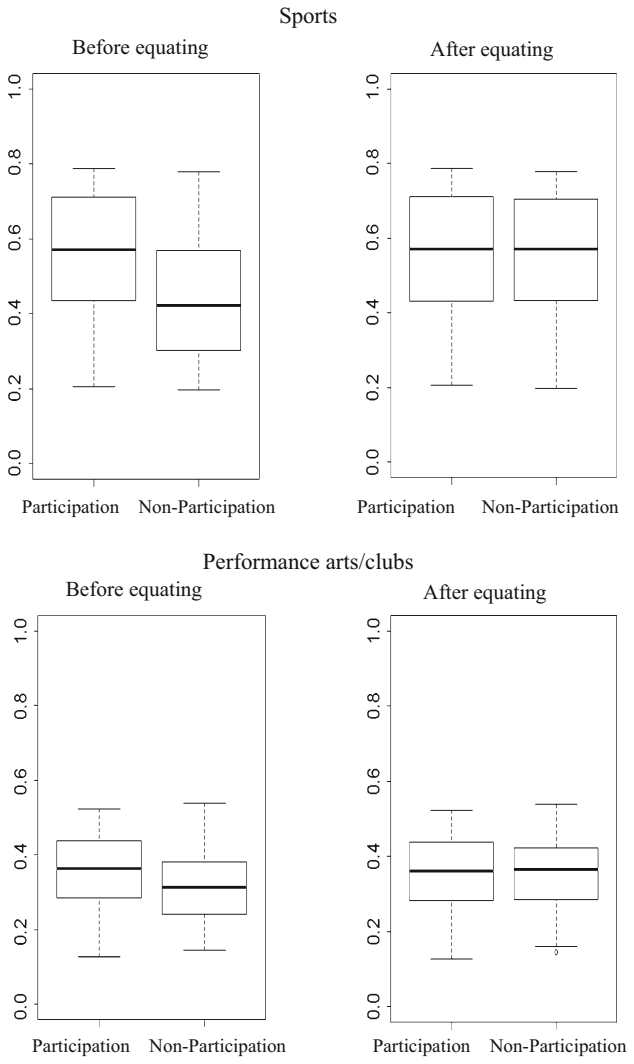

Fig. 2 Boxplot of propensity score between participation and non-participation in sport and performance arts or clubs before and after propensity score equating using odds method

Balance Check

To evaluate the effectiveness of the propensity score equating on reducing differences between participation groups on the covariates, we checked the balance of the distribution of the 46 covariates and their corresponding missing data pattern between participants and non-participants in each activity domain (Rosenbaum and Rubin 1984). Specifically, we calculated the absolute standardized mean difference (SMD) across participant groups on the 46 covariates and missing data patterns. A SMD of 0 indicates perfect balance. If any covariate indicates a substantial lack of balance (i.e., SMD $>$ .25), we further controlled for the potential confounding effect by this covariate by including it in the model (Ho et al. 2007). To reduce the number of statistical tests, thereby inflating type 1 error rate, we created two composite outcome variables for the SMD analyses. Specifically, given the strong correlations between grade 9 education value and academic competence belief (.65), we created a composite variable labeled academic beliefs, computed by averaging education value and achievement competence beliefs. Similarly, given the strong correlation between the grade 9 teacher-rated engagement and letter grade (.52), we created a composite variable labeled academic achievement by averaging teacher-rated engagement and letter grade.

Table 1 (panel 1 and panel 2) reports SMDs before and after propensity score weighting for sports and for performance arts or clubs for the 10 covariates with the strongest correlation with each of the composite outcomes, academic beliefs and academic achievement. Panel 1 in Table 1 shows that after propensity weighting, none of the 10 covariates with the highest correlation with academic beliefs differed between participation groups for sports or for performance arts or clubs by more than 0.25, indicating good balance (Ho et al. 2007). Panel 2 in Table 1 shows that after propensity weighting, youth participating in sports and those not participating in sports continued to differ only on the highest educational level of any adult in the household. For the domain of performance arts or clubs, no covariate had an SMD above 0.25.

Additionally, we assessed the degree to which the propensity weighting procedure also reduced differences between participating groups on variables most highly correlated with the hypothesized mediator, perceived friends' prosocial norms, prior to weighting (Panel 3 of Table 1). Only highest adult educational level in the home

Table 1 Absolute standardized mean difference (SMD) before and after balance check

Covariates	Sport		Performance arts or clubs	
	SMD		SMD	
	Before	After	Before	After
Panel I: (sort by correlation with academic belief)				
Teacher-student warmth_s	.201	.052	.021	.001
Subjective valuing of reading achievement_s	.264	.108	.165	.169
Subjective overall value of academic achievement_s	.298	.084	.197	.215
Educational aspirations for student_p	.156	.143	.402	.195
Student's hyperactivity_p	.057	.142	.135	.078
School belonging_s	.307	.108	.019	.031
Overall academic competence_s	.352	.090	.168	.158
Reading competence beliefs_s	.246	.098	.155	.121
Student's conduct problems_p	.101	.196	.225	.126
Student's ethnicity (African American vs. White)	.093	.053	.037	.017
Panel II: (sort by correlation with academic achievement)				
Highest adult educational level in household_p	.422	.251	.426	.139
Educational aspirations for student_t	.198	.158	.151	.008
Classroom engagement_t	.18	.154	.143	.056
Educational aspirations for student_p	.156	.143	.402	.195
Student's economic disadvantaged status	.211	.138	.324	.101
Student's hyperactivity_t	.077	.135	.204	.106
Student's conduct problems_p	.101	.196	.225	.126
School level reading achievement score Time 8	.049	.168	.116	.013
Woodcock-Johnson III Broad Math W Score	.217	.149	.366	.246
School level math achievement score at Time 8	.026	.159	.129	.008
Panel III: (sort by correlation with friend engagement)				
Student's economic disadvantaged status	.211	.138	.324	.101
School level reading achievement score Time 8	.049	.168	.116	.013
School level math achievement score at Time 8	.026	.159	.129	.008
Highest adult educational level in household_p	.422	.251	.426	.139
Student's gender	.277	.172	.257	.060
Student's ethnicity (Hispanic vs. White)	.283	.117	.207	.090
Classroom engagement_t	.180	.154	.143	.056
Teacher-student warmth	.201	.052	.021	.001
Home school relationship (alliance)	.174	.141	.076	.042
Friends' prosocial norms	.372	.184	.040	.047

had a SMD above 0.25. Thus, even though the propensity score weighting procedure was based on extracurricular participation status, the weighting procedure also reduced baseline differences on covariates most highly correlated with youths' perceived friends' prosocial norms.

Only 2 out of 46 (i.e., 4 %) of the missing data patterns in sports had a SMD above 0.25 (both SMD = .26), which may be a result of the low missing data rates for these two covariates (i.e., 4.4 and 4.2 % respectively). None of the missing data patterns in performance arts or clubs had a SMD above .25. Also, after weighting, the propensity scores for the participant and non-participant groups had good overlap. Based on these balance and overlap checks, we concluded that the propensity weighting procedure effectively equated participation groups on missing data patterns and on the most important covariates, with the exception of highest parent educational level for the sports participation model. To reduce the potential confounding effect by the "highest parent educational level", we included this covariate as a predictor of all grade 9 outcomes in the sports models only. We also included highest parent educational level as a predictor of grade 8 perceived

friends' prosocial norms in the mediation models for both sports and performance arts or clubs.

Results

Descriptive and Preliminary Results

Table 2 presents the correlations and descriptive statistics for the variables in the mediation models. All variables were checked for skewness and kurtosis, and none of the variables showed violation of the normality assumption based on commonly used cutoff values (i.e., skewness ≤ 2 and kurtosis ≤ 7; West, Finch, and Curran 1995). Perceived friends' positive school behaviors at grade 6 and grade 8 were moderately correlated ($r = .28$), as were friends' problem behavior at grade 6 and grade 8 ($r = .26$). Each baseline score at grade 5 was significantly correlated with the corresponding outcome at grade 9.

The measurement model for friends' prosocial norms in grades 6 and 8 was first examined. The perceived friends' prosocial norms latent factor consisted of two indicators at each assessment grade: friends' positive school behaviors and problem behaviors. The model was analyzed with M*plus* version 7.2 (Muthén and Muthén 1998–2014). The initial measurement model did not provide an adequate fit to the data. After correlating the residuals between grade 6 and grade 8 problem behaviors, the revised model resulted in a good fit (i.e., $\chi^2(1) = .29$, $p > .05$, RMSEA = 0, CFI = 1, SRMR = .008). The standardized path coefficients for the two indicators were .57 (SE = .21) and .50 (SE = .18) for positive and problems behaviors, respectively, at grade 6. At grade 8 the standardized path coefficients for positive and problem behaviors were .73 (SE = .27) and .50 (SE = .18), respectively. All path coefficients were significant at $p < .01$.

Indirect Effects Model

We analyzed our data using M*plus* version 7.2 (Muthén and Muthén 1998–2014) with the WEIGHT function to take into account the weighted propensity score, along with the TYPE = COMPLEX function to take into account the potential data dependency in our data (i.e., student nested within schools). The missingness in the dataset was handled with full information likelihood (FIML) parameter estimator (Enders 2010). In all the analyses, we controlled for the effect of perceived friends' prosocial norms at grade 6 on friends' prosocial norms at grade 8, and the effect of the grade 5 measure of each outcome on the grade 9 outcome. The bootstrap confidence interval method was employed to test the indirect effect, which is recommended when dealing with small sample size and small indirect effects (Cheung 2007). Specifically, we adopted the CINTERVAL (bcbootstrap) function (with 2000 iterations) in M*plus,* which produced the bias-corrected bootstrap confidence intervals, providing an accurate estimate of the indirect effect. Model fit for the indirect models in both sports and performance arts or clubs was good (i.e., RMSEA range from .00 to .04, CFI range from .96 to 1.00, SRMR range from .03 to .05).

Sports

As shown in Panel A of Table 3, the direct effect of sport participation on perceived friends' prosocial norms was significant ($\alpha = .11$) except in the engagement model, in which case the effect was marginally significant ($\alpha = .09$; $p < .10$). The direct effect of perceived friends' prosocial norms was statistically significant on competence beliefs ($\beta = .34$), value of education ($\beta = .29$), and course grades ($\beta = .25$), but not for classroom engagement, which was marginally significant ($\beta = .18$, $p < .10$). The direct effect of sport participation, after taking into account the indirect pathway, was statistically significant only for the value of education ($\gamma = .14$). Additionally, we examined the bias-corrected bootstrap confidence intervals for all the indirect effects and found significant indirect effects for the effect of sport participation on competence beliefs ($\alpha\beta = .04$, 95 % CI = [.01, .16]), value of education ($\alpha\beta = .04$, 95 % CI = [.01, .12]), course grade ($\alpha\beta = .07$, 95 % CI = [.02, .30]), and engagement ($\alpha\beta = .02$, 95 % CI = [.002, .17]).

Performance Arts or Clubs

As seen in Panel B of Table 3, the direct effect of performance arts or clubs on perceived friends' prosocial norms was significant only in the engagement model ($\alpha = .09$). The direct effects of perceived friends' prosocial norms on all outcomes were significant ($\beta = .19$ for competence beliefs, .20 for valuing of education, .19 for letter grades, and .13 for engagement, respectively). The direct effects of performance arts or club were statistically significant on competence beliefs ($\gamma = .11$), course grades ($\gamma = .10$) and engagement ($\gamma = .15$). The bias-corrected bootstrap confidence intervals test showed that none of the proposed indirect effects of performance arts or clubs was significant. Furthermore, the direct effects of performance arts or clubs participation on competence beliefs ($\gamma = .11$), course grades ($\gamma = .10$) and classroom engagement ($\gamma = .15$) were significant when the hypothesized indirect pathways were included. These results suggested that perceived friends' prosocial norms did not account for the effect of performance arts or clubs participation on the outcome variables.

Table 2 Correlations and descriptive statistics for outcomes at grade 9 and baseline variables of outcomes at grade 5 (N = 495)

		1	2	3	4	5	6	7	8	9	10	11	12	13
Correlations														
1	S_Prosocial 8	–												
2	S_Antisocial 8 (Reversed)	.37**	–											
3	S_Prosocial 6	.28**	.08	–										
4	S_Antisocial 6 (Reversed)	.17**	.26**	.29**	–									
5	S_Valuing of Education 9	.26**	.22**	.05	.07	–								
6	S_Ach. Compet. Belief 9	.29**	.24**	.03	.05	.52**	–							
7	T_Engagement 9	.21**	.24**	.07	.16**	.22**	.27**	–						
8	T_Course Grade 9	.21**	.25**	.04	.13*	.17**	.20**	.65**	–					
9	S_Value of Read & Math 5	.11*	.07	.09*	.02	.16**	.20**	.03	.08	–				
10	S_Comp Read & Math 5	.07	.03	.06	−.02	.12*	.21**	−.01	.07	.65**	–			
11	T_Engagement 5	.11*	.16**	.06	.06	.06	.09	.29**	.25**	.15**	.12*	–		
12	Reading Achievement 5	.02	.05	.05	.004	.05	.13**	.13*	.19**	.14**	.25**	.33**	–	
13	Male	−.15**	−.14**	−.08	−.12*	−.08	−.13**	−.21**	−.17**	−.02	.06	−.19**	−.08	–
Descriptive statistics														
	Mean	.89	.91	.9	.96	3.94	4.09	2.71	2.42	23.73	21.8	2.8	95.76	–
	SD	.17	.15	.17	.09	.65	.61	.64	1.3	4.7	4.69	.68	13.71	–
	Missing (%)	2	2	5	5	6	6	22	24	4	4	19	4	0

T_ is teacher reported. S_ is student-reported. Number at the end of each variable indicates grade. Ach. Comp = Achievement Competence; Comp = Competence Belief

* $p < .05$; ** $p < .01$

Table 3 Standardized effects of hypothesized mediation model

	Competence beliefs			Valuing of education			Course grades			Engagement		
	Est.	S.E.	95 %CI[b]	Est.	S.E.	95 %CI[b]	Est.	S.E.	95 %CI[b]	Est.	S.E.	95 %CI[b]
Panel A: Sport												
Sport → Friends (α)	.11**	.05		.11**	.05		.11*	.05		.09	.05	
Friend → Outcome (β)	.34**	.12		.29**	.05		.25**	.07		.18	.09	
Sport → Outcome (γ)	.06	.04		.14**	.03		−.02	.04		−.03	.06	
Indirect effect[a] (αβ)	.04	.03	[.01, .16]	.04	.02	[.01, .12]	.07	.08	[.02, .30]	.02	.01	[.002, .17]
Panel B: Performance arts or clubs												
Perf → Friends (α)	.06	.06		.06	.05		.07	.05		.09*	.04	
Friend → Outcome (β)	.19*	.08		.20*	.08		.19**	.08		.13**	.03	
Perf → Outcome (γ)	.11*	.05		−.01	.03		.10*	.05		.15*	.07	
Indirect effect[a] (αβ)	.01	.02	[−.01, .05]	.02	.02	[−.01, .05]	.03	.05	[−.01, .11]	.01	.01	[−.001, .03]

Estimates are standardized. SE is the corresponding standard error. TYPE = COMPLEX with school ID at Grade 9 for CLUSTER function was used to adjust for the data dependency. [a] Based on the unstandardized coefficients with bootstrap method; [b] A 95 % CI without covering zero indicates a statistically significant indirect effect. Friend = perceived friends' prosocial norms; Perf = Performance Arts or clubs

* $p < .05$; ** $p < .01$

Gender Moderation Effect

We also examined the potential gender moderation effect in the mediation model. We first allowed each path (i.e., α, β, γ) to be freely estimated so they could vary across student gender (relaxed model) and then restricted one path to be equal across gender (restricted model). We then compared the two competing nested models one at a time using the Satorra–Bentler test. Only one of the gender moderation effects was significant: the direct effect of performance arts or clubs on course grade was significantly different between male and female students ($\chi^2(1) = 4.02$, $p < .05$). Specifically, the positive effect of performance arts or clubs on course grades was significant only for male students ($\beta = .62$, SE = 0.22, $p < 0.01$).

Supplementary Analyses: Participation in Both Contexts

Of the 239 youth who participated in Sports and the 164 who participated in Performance Arts, 80 participated in both activity domains. In order to determine if the effect of each domain differed when combined with the other domain, we conducted two supplementary analyses. Specifically, we investigated an effect of sports alone versus both (sports plus performance arts) and an effect of performance arts alone versus both. These analyses paralleled the primary analyses (see supplementary files). For Sports alone versus both, no difference in the indirect effect was found for any of the four outcomes. Similar results were found for performance arts alone versus both. Thus, we concluded that results for the sports and for the performance arts analyses were not contaminated by inclusion of students who participated in both domains.

Discussion

Results of this study provide the strongest evidence to date that participation in sports in middle school leads to increased levels of prosocial norms of one's friends. This support, in turn, accounts for increases in the youth's academic competency beliefs, sense of school belonging, course grades, and classroom behavioral engagement. Middle school is typically a time of decreased academic support from peers, especially among ethnic minority and low income youth (Im et al. 2016), which characterizes the current sample. Findings suggest that participation in sports may stem the normative decline in perceived support from friends for prosocial norms, thereby improving youth's engagement in school and, ultimately, educational attainment.

An indirect effect of sports participation, via changes in perceived friends' prosocial norms, was found for each of the four outcomes. Specifically, after equating students who did and did not participate in sports in grade 8 on a wide range of covariates associated with sports participation and with the outcomes, sports participation predicted higher levels of perceived friends' prosocial norms in grade 8, controlling for grade 6 perceived friends' prosocial norms. In turn, perceived friends' prosocial norms in grade 8 predicted all outcomes except teacher-rated classroom engagement. In contrast, participation in performance arts and clubs did not predict changes in perceived friends' prosocial norms, and friends' prosocial norms did not account for the direct effects of performance arts/clubs on outcomes.

These results differ from those of Fredricks and Eccles (2005), who found prosocial peers mediated effects of participation in performance arts and clubs, but not participation in sports, on students' positive identification with school. Differences in findings may be due to differences in study methodology. The Fredricks and Eccles study employed a cross-sectional research design. Appropriately, these authors cautioned that the cross sectional nature of their data cannot rule out the conclusion that their findings are the results of selection effects. The longitudinal design of the current study and successful equating of participation groups on a comprehensive set of relevant covariates reduces the possibility that results are due to unmeasured confounds.

Different Pathways for Different Activity Domains

A positive effect of sports, but not performance arts, on perceived friends' prosocial norms at grade 8 (above and beyond friends' prosocial norms at grade 6) may be due to a greater emphasis on team work in sports, relative to performance arts activities and clubs. In football and basketball, coordination between players is essential. For example, in basketball, one player passes the ball to an open player who can then make the goal. Activities that involve this type of team work and interdependence lead to stronger social bonds and shared norms among group members (Levine et al. 1998). Competition with other teams may also strengthen group cohesion, which increases conformity to group norms (O'Reilly and Caldwell 1985). Because these bonds are formed in the context of a school-sponsored activity, participation may result in a more positive identification with school and its values, as was the case in the current study. In turn, students with a more positive identification with school tend to engage in fewer problem behaviors and to have higher academic motivation and achievement (Wang and Holcombe 2010).

The effect of participation in performance arts and clubs on academic outcomes is not explained by perceived friends' prosocial norms. Whereas participation in these activities may not influence perceived friend support for prosocial norms, it may influence other aspects of a youth's peer experiences, such as the amount of time the youth interacts with friends at school, which may enhance students' identification with and commitment to school (Finn 1989). The positive effects of participation in performance arts and clubs on competence beliefs, classroom engagement, and grades may also be a result of other assets associated with participation (e.g., access to supportive adult leaders and opportunities to develop academically-relevant skills).

Performance arts and clubs may also provide youth with a safe and accepting peer group, thereby increasing students' sense of acceptance and well-being. Performance arts activities such as band may have a stronger focus on cooperative rather than individual rewards. For example, typically all members of a band participate in performances, and individual "star players" are less salient than they are in sports. Cooperative reward structures, compared to more individual or competitive reward structures, promote more mutual help and assistance and more positive emotional experiences (Johnson and Johnson 2005). Future studies are needed to identify the specific assets associated with performance arts and clubs that account for its positive effect on academic functioning.

Despite gender differences in extracurricular participation domain gender did not moderate the effect of participation on one's perceived friends' prosocial norms, or the effect of one's perceived friends' prosocial norms on academic outcomes. Thus, the finding that friends' prosocial norms accounts for the positive effects of sports on academic outcomes holds for both boys and girls.

Study Strengths and Limitations

The study has several strengths, including the use of propensity score analyses to reduce the potential for unobserved confounds to account for the observed associations. By ruling out pre-existing differences between youth who choose to participate or not to participate in a given activity domain, this study provides the strongest evidence to date of an effect of participation on academic outcomes. Secondly, the study employed a prospective design that identified the indirect effect of participation, via its direct effect on one's friends' prosocial norms. Third, because the sample is ethnically diverse and predominantly low SES, findings are of considerable relevance to educators and policy makers concerned with reducing ethnic and SES disparities in educational achievement.

Despite these strengths, the study also is limited in several ways that warrant caution in interpreting findings. First, our measure of perceived friends' prosocial norms does not specify whether one's friends were co-participants in extracurricular activities. This information is of potential interest, because benefits of extracurricular participation in middle school or high school may be stronger when members of one's social network also participate (Mahoney 2000). However, knowledge of whether friends were co-participants is not critical to the argument that participation in extracurricular activities influences one's friendships. Youth typically navigate multiple, overlapping peer contexts (neighborhood, school, church, community clubs and sports). It is reasonable to expect that participation in extracurricular activities channels youth to positive experiences with prosocial youth who are engaged in school. These peer experiences influence a youth's academic values and behavior; consequently, these values and behaviors influence friendship choices both within and outside the particular participation context. In other words, as a youth identifies positively with school norms and prosocial

values, the youth is likely to seek affiliation with others who share these values (Véronneau et al. 2010), both within and outside of the specific extracurricular activity.

Second, our participation categories are broad. With respect to sports activities, distinguishing between team sports and individual sports may have led to more nuanced findings. For example, previous research has found that team sports may be more highly associated with group integration and liking for team members than individual sports (Brawley et al. 1987) and more highly associated with risk taking behavior such as alcohol use (Blomfield and Barber 2010). Based on national norms (National Federation of State High School Associations 2015) as well as interviews with guidance counselors in the participating schools, participation in team sports (primarily football, basketball and baseball among boys and volleyball, basketball, and soccer among girls) is much more common than is participation in individual sports. Thus, our results may not generalize to individual sports. Similarly, the category of performance arts and clubs includes a wide variety of activities that likely differ in the specific activity context. For example, the website for one middle school lists 55 non-sport clubs and activities from which youth may elect to participate. It is likely that this heterogeneity in specific activities is associated with differing assets and outcomes. Given likely differences between types of activities in peer experiences, it is important for future studies with larger samples to test differences in outcomes associated with different specific sport and non-sport activities.

Third, our measure of perceived friends' prosocial norms is based on self-report, and adolescents tend to over-estimate similarity between their characteristics and those of their friends (Prinstein and Wang 2005). However, adolescents' perceptions of their peers' academic achievement, engagement, and motivation predict their own academic outcomes, both concurrently (Lynch et al. 2013) and over time (Altermatt and Pomerantz 2005). Thus, a youth's perceptions of friends' characteristics may be at least as important as friends' actual attitudes and behaviors.

Finally, because students in the current sample were recruited into a larger longitudinal study in first grade on the basis of scoring below the median for their school district on a test of early literacy, the current findings may not generalize to students entering school with above average academic readiness. Future studies with students representing the entire range of academic risk are necessary to determine if level of risk moderates study findings.

Conclusion

The current study's findings suggest that sports participation, but not participation in other extracurricular activities, may buffer low SES youth from normative declines in levels of friend support for academic achievement, thereby accounting for the academic benefits of participation. These findings challenge earlier findings that sports participation may contribute to an increase in problem behaviors such as smoking (Fredricks and Eccles 2005). Differences in study methodology may account for differences in findings, as the present study minimized the possibility that preexisting differences between students who participate in sports and do not participate in sports would account for outcomes.

Findings also suggest the potential benefits of policies and practices that encourage all students to participate in extracurricular activities and remove barriers to participation, including financial and transportation barriers. Schools are also encouraged to increase opportunities for students of all athletic ability levels to participate in sports. Whereas school coaches report that building teamwork, goal setting, civic engagement, and other "life skills" is integral to their role (Smith and Smoll 2008), they receive little training in empirically-supported strategies for accomplishing these goals. Researchers are encouraged to identify such strategies and develop interventions to assist coaches in implementing them, thereby enhancing the beneficial effects of sports participation on academic outcomes.

Authors contributions JH conceived, designed, and directed the study, including collection of data, and took the primary role in preparation of the manuscript. QC conducted the data analyses, participated in the interpretation of the data, and wrote the results section. OK consulted on the data analysis. All authors read and approved the final manuscript.

Funding This Study was funded by the National Institute of Child Health and Human Development (Grant HD039367) awarded to the first author.

Conflict of interest The authors report no conflict of interests.

Human and Animals Rights The study was conducted in compliance with ethical standards. All procedures performed in this study involving human participants were in accordance with the ethical standards of the American Psychological Association and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Ethical Approval All procedures of the study including data collection and analyses were approved by the Institutional Review Board of Texas A&M University to ensure the proper protection of human subjects.

Informed Consent Informed consent was obtained by the first author from all individual participants included in the study.

Appendix

See Table 4.

Table 4 Covariates Used in in Propensity Score Analyses

Covariates
Source: Archival
1 Student's gender
2 Student's ethnicity (Hispanic vs. White)
3 Student's ethnicity (African American vs. White)
4 Student enrolled in bilingual class grade 4
5 Student enrolled in bilingual class in grade 1
6 Student's economic disadvantaged status
7 School size Grade 8
8 School level math achievement score Grade 8
9 School level reading achievement score Grade 8
Source: Parent
10 Highest adult employment level in the household
11 Highest adult educational level in household
12 Highest educational level expected for student
13 Student's emotional symptoms (Strengths and Difficulties Questionnaire; Goodman 2001)
14 Student's conduct problems (Strengths and Difficulties Questionnaire; Goodman 2001)
15 Student's hyperactivity (Strengths and Difficulties Questionnaire; Goodman 2001)
16 Student's peer problems(Strengths and Difficulties Questionnaire Goodman 2001)
Source: Teacher
17 Highest educational level expected for student
18 Academic performance in classroom (mean rating of reading, math, and overall academic performance relative to grade level expectancies.
19 Student's conduct problems (Strengths and Difficulties Questionnaire; Goodman 2001)
20 Student's hyperactivity (Strengths and Difficulties Questionnaire; Goodman 2001)
21 Student's prosocial behavior (Strengths and Difficulties Questionnaire; Goodman 2001)
22 Home school relationship: alliance (Parent Involvement in Early Years; Hill and Hughes 2007)
23 Parental school involvement (Parent Involvement in Early Years; Hill and Hughes 2007)
24 Classroom behavioral engagement (Student Engagement Questionnaire; Chen et al. 2010))
Source: Student
25 Teacher-student warmth (Teacher Network of Relationships Inventory; Hughes et al. 2008)
26 Math competence beliefs (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
27 Reading competence beliefs (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
28 Sports competence beliefs (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
29 Subjective valuing of math achievement (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
30 Subjective valuing of reading achievement (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
31 Overall academic competence (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
32 Subjective overall valuing of academic achievement (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
33 Subjective valuing of sports (Competence Beliefs and Subjective Task Values; Wigfield et al. 1997)
34 Overall scholastic competence (Self Perception Profile for Children; Harter 1985)
35 Social competence (Self Perception Profile for Children; Harter 1985)
36 Athletic competence (Self Perception Profile for Children; Harter 1985)
37 School belonging (Psychological Sense of School Membership; Goodman 2001)
38 Classroom performance approach goal structure (Student Perception of Classroom Goals; Midgley et al. 2000)
39 Classroom performance avoidance goal structure (Student Perception of Classroom Goals; Midgley et al. 2000)
40 Friends' prosocial orientation (student reported percentage of friends with prosocial behaviors)
41 Student-reported victimization in classroom
42 Locus of control (Student Perception of Control Questionnaire; Skinner et al. 1998)
43 Peer affiliation: structured activities. Student report of participating in structured activities with friends
Source: Performance
44 Student's dominant language (Woodcock and Muñoz-Sandoval 1993)
45 Woodcock-Johnson III Broad Reading W Score (Woodcock et al. 2001)
46 Woodcock-Johnson III Broad Math W Score (Woodcock et al. 2001

All measures were assessed at Grade 4 or 5 unless otherwise indicated. Additional information on these variables is available from the first author

References

Alexander, K. L., Entwisle, D. R., & Horsey, C. S. (1997). From first grade forward: Early foundations of high school dropout. *Sociology of Education, 70*, 87–107. doi:10.2307/2673158.

Altermatt, E., & Pomerantz, E. M. (2005). The implications of having high-achieving versus low-achieving friends: A longitudinal analysis. *Social Development, 14*, 61–81. doi:10.1111/j.1467-9507.2005.00291.x.

Beal, S. J., & Crockett, L. J. (2010). Adolescents' occupational and educational aspirations and expectations: Links to high school activities and adult educational attainment. *Developmental Psychology, 46*, 258–265. doi:10.1037/a0017416.

Berndt, T. J., Hawkins, J. A., & Jiao, Z. (1999). Influences of friends and friendships on adjustment to junior high school. *Merrill-Palmer Quarterly, 45*, 13–41.

Berndt, T. J., & Keefe, K. (1995). Friends' influence on adolescents' adjustment to school. *Child Development, 66*, 1312–1329. doi:10.1111/j.1467-8624.1995.tb00937.x.

Blomfield, C., & Barber, B. (2010). Australian adolescents' extracurricular activity participation and positive development: Is the relationship mediated by peer attributes? *Australian Journal of Educational & Developmental Psychology, 10*, 114–128. www.newcastle.edu.au/journal/ajedp.

Brawley, L. R., Carron, A. V., & Widmeyer, W. N. (1987). Assessing the cohesion of teams: Validity of the Group Environment Questionnaire. *Journal of Sport Psychology, 9*, 275–294.

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. doi:10.1023/A:1010933404324.

Broh, B. A. (2002). Linking extracurricular programming to academic achievement: Who benefits and why?. *Sociology of Education, 75*, 69–91. http://soe.sagepub.com/.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In M. B. Bronson (Ed.), *Self-regulation in early childhood: Nature and nurture* (pp. 793–828). New York, NY: Guilford Press.

Brown, B. B., & Larson, J. (2009). Contextual influences on adolescent development. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (3rd ed., Vol. 2, pp. 74–103). Hoboken, NJ: Wiley.

Cham, H., Hughes, J. N., West, S. G., & Im, M. H. (2014). Assessment of adolescents' motivation for educational attainment. *Psychological Assessment, 26*, 642–659. doi:10.1037/a0036213.

Chambers, E. A., & Schreiber, J. B. (2004). Girls' academic achievement: Varying associations of extracurricular activities. *Gender and Education, 16*, 327–346. doi:10.1080/09540250042000251470.

Chen, Q., Hughes, J. N., Liew, J., & Kwok, O. (2010). Joint contributions of peer acceptance and peer academic reputation to achievement in academically at-risk children: Mediating processes. *Journal of Applied Developmental Psychology, 31*, 448–459. doi:10.1016/j.appdev.2010.09.001.

Cheung, M. W. (2007). Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 227–246. doi:10.1080/10705510709336745.

Cotter, K. L., & Smokowski, P. R. (2016). Perceived peer delinquency and externalizing behavior among rural youth: The role of descriptive norms and internalizing symptoms. *Journal of Youth and Adolescence, 45*, 520–531. doi:10.1007/s10964-015-0382-1.

Denault, A., & Poulin, F. (2009a). Intensity and breadth of participation in organized activities during the adolescent years: Multiple associations with youth outcomes. *Journal of Youth and Adolescence, 38*, 1199–1213. doi:10.1007/s10964-009-9437-5.

Denault, A., & Poulin, F. (2009b). Predictors of adolescent participation in organized activities: A five-year longitudinal study. *Journal of Research on Adolescence, 19*, 287–311. doi:10.1111/j.1532-7795.2009.00597.x.

Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: What kind of extracurricular involvement matters? *Journal of Adolescent Research, 14*, 10–43. doi:10.1177/0743558499141003.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: The Guilford Press.

Farb, A. F., & Matjasko, J. L. (2012). Recent advances in research on school-based extracurricular activities and adolescent development. *Developmental Review, 32*, 1–48. doi:10.1016/j.dr.2011.10.001.

Feldman, A. F., & Matjasko, J. L. (2005). The role of school-based extracurricular activities in adolescent development: A comprehensive review and future directions. *Review of Educational Research, 75*, 159–210. doi:10.3102/00346543075002159.

Feldman, A. F., & Matjasko, J. L. (2007). Profiles and portfolios of adolescent school-based extracurricular activity participation. *Journal of Adolescence, 30*, 313–332. doi:10.1016/j.adolescence.2006.03.004.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*, 117–142.

Fredricks, J. A., & Eccles, J. S. (2005). Developmental benefits of extracurricular involvement: Do peer characteristics mediate the link between activities and youth outcomes? *Journal of Youth and Adolescence, 34*, 507–520. doi:10.1007/s10964-005-8933.

Fredricks, J. A., & Eccles, J. S. (2006). Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations. *Developmental Psychology, 42*, 698–713. doi:10.1037/0012-1649.42.4.698.

Fredricks, J. A., & Eccles, J. S. (2008). 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, 37*, 1029–1043. doi:10.1007/s10964-008-9309-4.

Fredricks, J. A., Hackett, K., & Bregman, A. (2010). Participation in boys and girls clubs: Motivation and stage environment fit. *Journal of Community Psychology, 38*, 369–385. doi:10.1002/jcop.20369.

Fredricks, J. A., & Simpkins, S. D. (2012). Promoting positive youth development through organized after-school activities: Taking a closer look at participation of ethnic minority youth. *Child Development Perspectives, 6*, 280–287. doi:10.1111/j.1750-8606.2011.00206.x.

Fredricks, J. A., & Simpkins, S. D. (2013). Organized out-of-school activities and peer relationships: Theoretical perspectives and previous research. *New Directions for Child and Adolescent Development, 2013*(140), 1–17. doi:10.1002/cad.20034.

Goodman, R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of Child Psychology and Psychiatry, 40*, 791–799. doi:10.1097/00004583-200111000-00015.

Green-Demers, I., Legault, L., Pelletier, D., & Pelletier, L. G. (2008). Factorial invariance of the academic amotivation inventory (AAI) across gender and grade in a sample of Canadian high school students. *Educational and Psychological Measurement, 68*, 862–880. doi:10.1177/0013164407313368.

Hardway, C., & Fuligni, A. J. (2006). Dimensions of family connectedness among adolescents with Mexican, Chinese, and European backgrounds. *Developmental Psychology, 42*, 1246–1258. doi:10.1037/0012-1649.42.6.1246.

Harter, S. (1985). *The self-perception profile for children: A revision of the perceived competence scale for children*. Denver, CO: University of Denver.

Hill, C. R., & Hughes, J. N. (2007). Further evidence of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. *School Psychology Quarterly, 22*, 380–406. doi:10.1037/1045-3830.22.3.380.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199–236. doi:10.1093/pan/mpl013.

Hong, G., & Raudenbush, S. W. (2006). Evaluating kindergarten retention policy: A case study of causal inference for multilevel observational data. *Journal of the American Statistical Association, 101*, 901–910. doi:10.1198/016214506000000447.

Hughes, J. N., Im, M., Cham, H., Kwok, O. M., & West, S. G. (2014). Latino students' school engagement at the transition to middle school: Role of bilingual education and school ethnic context. *Journal of Research on Adolescence, 26*, 642–659. doi:10.1111/jora.12142.

Hughes, J. N., Luo, W., Kwok, O., & Loyd, L. K. (2008). Teacher-student support, effortful engagement, and achievement: A 3-year longitudinal study. *Journal of Educational Psychology, 100*, 1–14.

Im, M., Hughes, J. N., Cao, Q., & Kwok, O. M. (2015). *Effects of Extracurricular Participation during Middle School on Academic Motivation and Achievement at Grade 9*. Manuscript under review, available from Jan Hughes, 4225 TAMU, College Station, TX, 77843-4225.

Im, M., Hughes, J. N., & West, S. G. (2016). Effect of trajectories of friends' and parents' school involvement on adolescents' engagement and achievement. Early View. *Journal of Research on Adolescence*. doi:10.1111/jora.12247.

Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of Social Issues, 64*(1), 21–40. doi:10.1111/j.1540-4560.2008.00546.x.

Johnson, D. W., & Johnson, R. T. (2005). New developments in social interdependence theory. *Genetic, Social, and General Psychology Monographs, 131*(4), 285–358. doi:10.3200/MONO.131.4.285-358.

Kindermann, T. A. (2007). Effects of naturally existing peer groups on changes in academic engagement in a cohort of sixth graders. *Child Development, 78*, 1186–1203. doi:10.1111/j.1467-8624.2007.01060.x.

Kindermann, T. A., & Skinner, E. A. (2012). Will the real peer group please stand up? A "tensegrity" approach to examining the synergistic influences of peer groups and friendship networks on academic development. In A. M. Ryan & G. Ladd (Eds.), *Peer relationships and adjustment at school* (pp. 51–78). Charlotte, NC: Information Age Publishing.

Larson, R. W., Hansen, D. M., & Moneta, G. (2006). Differing profiles of developmental experiences across types of organized youth activities. *Developmental Psychology, 42*, 849–863. doi:10.1037/0012-1649.42.5.849.

Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine, 29*, 337–346. doi:10.1002/sim.3782.

Levine, J. M., Moreland, R. L., & Ryan, C. S. (1998). Group socialization and intergroup relations. In C. Sedikides, J. Schopler, & C. A. Insk (Eds.), *Intergroup cognition and intergroup behavior*. Mahwah, NJ: Lawrence Erlbaum Associates.

Lynch, A. D., Lerner, R. M., & Leventhal, T. (2013). Adolescent academic achievement and school engagement: An examination of the role of school-wide peer culture. *Journal of Youth and Adolescence, 42*, 6–19. doi:10.1007/s10964-012-9833-0.

Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development, 71*, 502–516. doi:10.1111/1467-8624.00160.

Mahoney, J. L., & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology, 33*, 241–253. doi:10.1037/0012-1649.33.2.241.

Martin, T. F., White, J. M., & Perlman, D. (2003). Religious socialization: A test of the channeling hypothesis of parental influence on adolescent faith maturity. *Journal of Adolescent Research, 18*, 169–187. doi:10.1177/0743558402250349.

Midgley, C., Maehr, M. L., Hurda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Manual for the patterns of adaptive learning scales*. Ann Arbor, MI: University of Michigan.

Muthén, L. K., & Muthén, B. O. (1998–2014). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

National Federation of State High School Associations. (2015). *2014–2015 High School Athletics Participation Survey*. Retrieved December 4, 2016 from http://www.nfhs.org/ParticipationStatics/ParticipationStatics.aspx/.

O'Reilly, C., & Caldwell, D. F. (1985). The impact of normative social influence and cohesiveness on task perceptions and attitudes: A social-information processing approach. *Journal of Occupational Psychology, 58*, 193–206. http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)2044-8325/issues.

Prinstein, M. J., & Wang, S. S. (2005). False consensus and adolescent peer contagion: Examining discrepancies between perception and actual reported levels of friends' deviant and health risk behaviors. *Journal of Abnormal Child Psychology, 33*, 293–306. doi:10.1007/s10802-005-3566-4.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*(387), 516–524. doi:10.1080/01621459.1984.10478078.

Sage, N. A., & Kindermann, T. A. (1999). Peer networks, behavior contingencies, and children's engagement in the classroom. *Merrill-Palmer Quarterly, 45*, 143–171.

Schaefer, D. R., Simpkins, S. D., Vest, A. E., & Price, C. D. (2011). The contribution of extracurricular activities to adolescent friendships: New insights through social network analysis. *Developmental Psychology, 47*, 1141–1152. doi:10.1037/a0024091.

Schafer, J. L., & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods, 13*, 279–313. doi:10.1037/a0014268.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Simpkins, S. D., Vest, A. E., Delgado, M. Y., & Price, C. D. (2012). Do school friends participate in similar extracurricular activities?: Examining the moderating role of Race/Ethnicity and age. *Journal of Leisure Research, 44*, 332–352. http://js.sagamorepub.com/jlr.

Skinner, E. A., Zimmer-Gembeck, M. J., & Connell, J. P. (1998). Individual differences and the development of perceived control. *Monographs of the Society for Research in Child Development, 63*, 1–231. doi:10.2307/1166220.

Smith, R. E., & Smoll, F. L. (2008). Coach-mediated team building in youth sports. *Journal of Applied Sport Psychology, 9*, 114–132. doi:10.1080/10413209708415387.

Strobl, C., Boulesteix, A. L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics, 9*, 307. doi:10.1186/1471-2105-9-307.

Véronneau, M. H., & Dishion, T. J. (2010). Predicting change in early adolescent problem behavior in the middle school years: A

mesosystemic perspective on parenting and peer experiences. *Journal of Abnormal Child Psychology, 38*, 1125–1137. doi:10.1007/s10802-010-9431-0.

Véronneau, M. H., & Dishion, T. J. (2011). Middle school friendships and academic achievement in early adolescence: A longitudinal analysis. *The Journal of Early Adolescence, 31*, 99–124. doi:10.1177/0272431610384485.

Véronneau, M., Vitaro, F., Brendgen, M., Dishion, T. J., & Tremblay, R. E. (2010). Transactional analysis of the reciprocal links between peer experiences and academic achievement from middle childhood to early adolescence. *Developmental Psychology, 46*, 773–790. doi:10.1037/a0019816.

Wang, M., & Eccles, J. S. (2012). Social support matters: Longitudinal effects of social support on three dimensions of school engagement from middle to high school. *Child Development, 83*, 877–895. doi:10.1111/j.1467-8624.2012.01745.x.

Wang, M. T., & Holcombe, R. (2010). Adolescents' perceptions of classroom environment, school engagement, and academic achievement. *American Educational Research Journal, 47*, 633–662. doi:10.3102/0002831209361209.

Wentzel, K. R., Donlan, A., & Morrison, D. (2012). Peer relationships and social-motivational processes. In A. M. Ryan & G. Ladd (Eds.), *Peer relationships and adjustment at school* (pp. 79–107). Charlotte, NC: Information Age Publishing.

West, S. G., Cham, H., Thoemmes, F., Renneberg, B., Schulze, J., & Weiler, M. (2014). Propensity scores as a basis for equating groups: Basic principles and application in clinical treatment outcome research. *Journal of Consulting and Clinical Psychology, 82*, 906–919. doi:10.1037/a0036387.

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Thousand Oaks, CA: Sage Publications.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*, 451–469. doi:10.1037/0022-0663.89.3.451.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III tests of achievement*. Riverside, CA: Riverside Publishing.

Woodcock, R. W., & Muñoz-Sandoval, A. F. (1993). *Woodcock-Muñoz language survey*. Riverside, CA: Riverside Publishing.

Woodcock, R. W., Muñoz-Sandoval, A. F., McGrew, K., Mather, N., & Schrank, F. (2004). *Bateria III Woodcock-Munoz*. Riverside, CA: Riverside Publishing.

Jan N. Hughes, Ph.D. is Professor Emerita in the Department of Educational Psychology at Texas A&M University. She obtained her doctorate in Educational Psychology. Her research Interests include peer relationships, teacher-student relationships, social and emotional development, and academic achievement.

Qian Cao, M.S. 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, Ph.D. holds the position of Professor in the Department of Educational Psychology at Texas A&M University. He obtained his doctorate in the quantitative methods program in the Department of Psychology at Arizona State University. His research interests include structural equation modeling and longitudinal data analysis.