Optional ERIC Coversheet — Only for Use with U.S. Department of Education Grantee Submissions

This coversheet should be completed by grantees and added to the PDF of your submission if the information required in this form **is not included on the PDF to be submitted**.

INSTRUCTIONS

- Before beginning submission process, download this PDF coversheet if you will need to provide information not on the PDF.
- Fill in all fields—information in this form **must match** the information on the submitted PDF and add missing information.
- Attach completed coversheet to the PDF you will upload to ERIC [use Adobe Acrobat or other program to combine PDF files]—do not upload the coversheet as a separate document.
- Begin completing submission form at https://eric.ed.gov/submit/ and upload the full-text PDF with attached coversheet when indicated. Your full-text PDF will display in ERIC after the 12-month embargo period.

GRANTEE SUBMISSION REQUIRED FIELDS

Title of article, paper, or other content

Do Top Dogs Rule in Middle School? Evidence on Bullying, Safety, and Belonging

All author name(s) and affiliations on PDF. If more than 6 names, ERIC will complete the list from the submitted PDF.

Last Name, First Name	Academic/Organizational Affiliation	ORCID ID
Schwartz, Amy Ellen	University of Delaware	
Stiefel, Leanna	New York University	
Rothbart, Michah W.	Syracuse University	

Publication/Completion Date—(if *In Press,* enter year accepted or completed) 2016

Check type of content being submitted and complete one of the following in the box below:

- ☑ If article: Name of journal, volume, and issue number if available
- ☐ If paper: Name of conference, date of conference, and place of conference
- ☐ If book chapter: Title of book, page range, publisher name and location
- ☐ If book: Publisher name and location
- ☐ If dissertation: Name of institution, type of degree, and department granting degree

American Educational Research Journal, 53, 5

DOI or URL to published work (if available) https://doi.org/10.3102/0002831216657177

Acknowledgement of Funding— Grantees should check with their grant officer for the preferred wording to acknowledge funding. If the grant officer does not have a preference, grantees can use this suggested wording (adjust wording if multiple grants are to be acknowledged). Fill in Department of Education funding office, grant number, and name of grant recipient institution or organization.

"This work was supported by U.S. Department of Education **[Office name]** Institute of Education Scie through **[Grant number]** R305B1400 to **Institution]** New York Univ. The opinions expressed are those of the authors and do not represent views of the **[Office name]** Institute of Education Scie or the U.S. Department of Education.

Do Top Dogs Rule in Middle School? Evidence on Bullying, Safety, and Belonging

Amy Ellen Schwartz
Syracuse University
Leanna Stiefel
New York University
Michah W. Rothbart
Syracuse University

Recent research finds that grade span affects academic achievement, but only speculates about the mechanisms. In this study, we examine one commonly cited mechanism, the top dog/bottom dog phenomenon, which states that students at the top of a grade span ("top dogs") have better experiences than those at the bottom ("bottom dogs"). Using an instrumental variables strategy introduced in Rockoff & Lockwood (2010) and a longitudinal data set containing student survey data for New York City public middle school students, we estimate the impact of top dog and bottom dog status on bullying, safety, belonging, and academic achievement. This paper provides the first credibly causal evidence that top dog status improves the learning environment and academic achievement. We further find that the top dog effect is strongest in 6th grade and in schools with longer grade spans and that the top dog effect is not explained by new students to a school or student height.

Keywords: Academic Achievement, Bullying, Grade Span, Learning Environment

This is an Accepted Manuscript. The Final Published version appears in the *American Educational Research Journal* (https://doi.org/10.3102/0002831216657177)

Preferred citation:
Schwartz, A. E., Stiefel, L., & Rothbart, M. W. (2016). Do Top Dogs Rule in Middle School? Evidence on Bullying, Safety, and Belonging. *American Educational Research Journal*, 53(5), 1450-1484.

Do Top Dogs Rule in Middle School? Evidence on Bullying, Safety, and Belonging

Introduction

> *Well my story about being bullied all started in 6th grade, I was getting picked on because I was to [sic] tiny for my age. Then I started getting picked on for the guys I would date, they would call me harsh names n [sic] some very bad words, I think everyone knows what im [sic] talking about…then 7th grade year came around and the bullying got worse I had people wanting to fight me and jump me for even trying to talk to kids, I told adults like school officials sadly they did not take action. I even told my parents but they didnt believe me. In 8th grade everything was ok.*
>
> – Angel, NOPLACE4HATE, http://www.noplace4hate.org/real-bullying-stories/

The middle school grades (6th, 7th, and 8th) are widely regarded as a difficult period for students due to hormonal changes, amplified social pressures, and increased bullying (Poiner, 2015). Nationally, a majority of students in 6th, 7th, and 8th grades are separated from students in elementary grades (K-5) and high schools (9-12), perhaps to provide targeted support to students during these difficult years. Recent research, however, finds that transitioning to middle school has a negative effect on student academic achievement (Rockoff & Lockwood, 2010; Schwartz, Stiefel, Rubenstein, & Zabel, 2011; Schwerdt & West, 2013).

One commonly cited potential mechanism for the middle school achievement dip is the top dog/bottom dog (TDBD) phenomenon (Blyth, Simmons, & Bush, 1978; Weiss & Kipnes, 2006; Cook et al., 2008), which occurs when students at the top of a grade span distribution have better experiences than those at the bottom. That is, the oldest students in a school ("top dogs") experience a more favorable school environment (including increased participation in leadership roles and decreased incidence of bullying and fights) than the youngest students in a school ("bottom dogs") who face increased victimization, exposure to delinquent peer influences, and feelings of anonymity. The TDBD hypothesis suggests that when students graduate from an elementary school – where they are top dogs – to a middle school –where they become bottom dogs – they see a dip in quality of school experiences. This TDBD effect may drive the dip in performance observed at the start of middle school. Despite the intuitive appeal, there is little

evidence to support the TDBD hypothesis, much less its effect on academic performance. In this paper, we begin to close this gap, providing causal evidence on the effects of top or bottom dog status on student learning environments and opening the black box of grade span effects. Specifically, we explore effects on bullying, safety, belonging, and academic achievement in the middle school grades, contributing to the evolving literature on mechanisms through which grade span may affect academic performance.

The TDBD hypothesis was first introduced by Blyth et al. (1978) in a study of 622 students attending 14 schools (six K-8 and eight K-6) in a large Midwestern city. The authors found that 7th grade students who were nearly top dogs in a K-8 were more confident, participated in more activities, and felt less anonymous than 7th grade students who were bottom dogs in junior high schools after transitioning from a K-6. This provided suggestive evidence that students in the middle grades could benefit from attending elementary schools with longer grade spans rather than middle schools or junior high schools. Since then, it has been repeatedly observed that students at the top of a grade span distribution are less bullied, safer, and feel more comfortable in school than those at the bottom (Simmons & Blyth, 1987; Weiss & Kipnes, 2006; Byrnes & Ruby, 2007; Cook et al., 2008). While the TDBD hypothesis is widely cited, existing research is best viewed as descriptive and correlational, making little or no effort to disentangle the causal effects from sorting, differences across grades, or other potential confounders. For example, the extent to which the TDBD effect is driven by new students (bottom dogs are almost always new) or student height (bottom dogs are most often shorter than top dogs) is unaddressed in previous literature.

Isolating the effects of top dog and bottom dog status on bullying, safety, and belonging is difficult for two key reasons. First, it is challenging to identify random variation in TDBD

status. On the one hand, some school districts offer only one grade span option to students in each grade, providing no variation to study the TDBD effect and confounding grade effects with the TDBD mechanism. On the other hand, in school districts with a variety of grade spans, grade span might be endogenous to TDBD status. For example, a bullied rising 6th grader in a K-8 school may use 6th grade as an opportunity to transfer to a 6-8 school, making the student a bottom dog. Perhaps this 6th grader is simply more prone to bullying than the average student, which precipitates the transfer and the student's bottom dog status in 6th grade. In this particular example, an OLS estimate of the effect of bottom dog status on bullying would be biased upward. The second key challenge in testing the TDBD hypothesis is the scarcity of student-level data with good measures of bullying, safety, and belonging.

This article explores the TDBD hypothesis, using relatively new longitudinal data on students and schools in the nation's largest school district, New York City (NYC), including student-level responses to questions on the NYC School Survey regarding student experiences and the school environment. We estimate the effects of TDBD status on bullying, safety, and belonging using two cohorts of NYC public middle school students, including roughly 90,000 6th-8th grade students in over 500 schools. These rich data allow us to use both student fixed effects and a variety of student characteristics to control for potential differences between students. Further, we aim to address potential endogeneity of TDBD status using an instrumental variables strategy previously employed to study grade span and student mobility (Rockoff and Lockwood, 2010; Schwartz, Stiefel, & Cordes, 2016). The intuition is as follows. The problem with the causal interpretation of correlations between outcomes and TDBD status is that students effectively choose top dog/bottom dog status by enrolling in a school with a particular grade span in a given grade. Thus, TDBD status reflects the combined effect of a student's grade and

school grade span. But, as discussed in Rockoff & Lockwood (2010), grade span in 3rd grade is plausibly exogenous to TDBD status in 6th, 7th, and 8th grades because grade span choice for 3rd grade occurs long before middle school and is not likely to reflect anticipation of the middle school learning environment three to five years later. Grade span in 3rd grade is, however, highly predictive of TDBD status in middle school grades. Thus, grade spans of students' 3rd grade schools can be used as a set of instruments for top dog/bottom dog status in 6th, 7th, and 8th grades.

In this paper, then, we estimate the effects of TDBD status by examining student reports of the learning environment as students are promoted through and transfer between schools. We estimate how students' experiences and academic achievement change as they move from bottom to top dog or from top to bottom dog. In brief, we find that top dog status improves a student's learning environment. Top dogs are *less likely* to report bullying, fights, and gang activity, and *more likely* to report feeling safe and welcome in school than bottom dogs due to their top dog status. In contrast, bottom dogs report *higher* rates of bullying, fighting, and gang activity, and *lower* rates of safety and belonging than top and middle dogs, although the bottom dog results are sensitive to the inclusion of student fixed effects and addressing selection into TDBD status. We note that there is greater variation in TDBD status among 6th graders than among 7th and 8th graders. Thus, we explore possible heterogeneity by grade of the top dog effect, finding a larger top dog effect in 6th grade than 8th grade. We test whether the top dog effect is bigger in schools with longer grade spans and find the top dog effect is increasing in "heap size." We also dig deeper into the causes of the TDBD effect, showing results are not driven by new students or student height. While new students and shorter students have more negative perceptions of the school learning environment, there is an independent top dog effect

even once conditioning on these plausible moderators. We further note that other plausible explanations of the TDBD effect in 6th, 7th, and 8th grades, such as differences in financial resources, class sizes or use of subject-specific teachers are unlikely, because these do not vary greatly for students in the middle school grades no matter what NYC school students attend (Rockoff and Lockwood, 2010). Moreover, we show that top dog status leads to improved academic achievement and provide correlational evidence that this could operate through improved student experiences for top dogs. These findings suggest that longer grade spans (such as K-8 schools) may better serve students in the middle grades compared to other popular grade spans (e.g., 6-8) due to the benefits of top dog status in longer grade spans.

The paper is organized as follows. In Section II, we review the literature on the TDBD phenomenon as it relates to grade span. In Section III, we discuss our methods, including our data and measures, descriptive statistics and analytic approach. In Section IV we summarize our results, and we discuss our conclusions in Section V.

Literature Review: TDBD and Grade Span

Recent evidence suggests that student academic performance in middle school grades is shaped by school grade span (Rockoff & Lockwood, 2010; Schwartz et al., 2011; Schwerdt & West, 2013). There are at least three plausible explanations for the why grade span matters. Grade span could affect a student's (1) top dog/bottom dog status, (2) mobility, due to school transitions, and (3) school and student characteristics, including school/cohort size, class size, subject-specific teachers, funding per pupil, and student motivation and self-concept. We address a gap in the literature, providing insight into how student experiences are shaped by grade span and complementing existing research exploring other potential mechanisms, for example

mobility, school characteristics such as cohort or school size and classroom environments, and student motivation and self-concept (Alspaugh, 1998; Bloom & Unterman, 2012; Cordes, Schwartz, & Stiefel, 2014; Carolan, Weiss, & Matthews, 2015; Eccles, Lord, & Midgley, 1991; Eccles & Midgley, 1989; Eccles, Wigfield, Midgley, Reuman, Iver, & Feldlaufer, 1993; Howley, 2002; Lee & Smith, 1997; Napier, 2008; Offenberg, 2001; Powell, Farrar, & Cohen, 1985; Rubenstein, Schwartz, Stiefel, & Zabel, 2009).

Schwartz et al. (2011) suggest that grade span may affect student achievement because grade span changes the social environment and particularly TDBD status. Longer grade spans, for example, extend students' opportunity to be among the oldest in a school, while students entering middle schools transition from being the oldest in elementary school to the youngest in middle school. In addition, Cook et al. (2008) suggest that middle school entry exposes 6th or 7th graders to older peers who can serve as negative influences and hinder academic performance. Moreover, Blyth et al. (1978) hypothesize top dogs in schools with longer grade spans benefit from delaying school transitions, experiencing more welcoming school environments and having greater opportunity to be nearly top dogs in their school. Conversely, new students may be particularly vulnerable to bullying and poorer perceptions of the learning environment and bottom dogs are almost always new (Pellegrini, Long, Solberg, Roseth, Dupuis, Bohn & Hicke, 2010).

The TDBD effect may also depend upon student grade due to developmental differences across middle school grades, such as differences in prefrontal cortex development, pubescent physical maturation, and development of greater ability to think abstractly (Eccles, 1999; Fuster, 2002; Keating, 1990; Yurgelun-Todd, 2007). For example, 8th graders may have matured to be better equipped to serve as school leaders than 6th graders, resulting in larger top dog effects

among 8th graders than 6th graders. Alternatively, one might believe that 6th grade top dog effects are greater than in 8th grade, because 8th graders have greater ability to situate themselves in social contexts and may maintain better perceptions of the learning environment even without top dog status. Most middle school configurations give 8th graders top dog status, which may indicate policymakers' implicit preferences to give 8th graders a boost to their learning environment (rather than, for example, including 9th graders in middle school or beginning high school in 8th grade).

Others have found that short students are more susceptible to bullying (Borg, 1999; Voss & Mulligan, 2000). Students 6th through 8th grades grow at a rapid rate and differences in timing of growth spirts during early adolescents might matter (Eccles, 1999). Thus, the TDBD effect may depend on student height.

Despite being widely cited in the grade span literature, there is little empirical evidence on the TDBD phenomenon *per se*. Blyth et al. (1978) and Simmons and Blyth (1987) are notable exceptions, providing evidence from a longitudinal study in 14 Milwaukee public schools in the 1970s. They compare student responses to surveys in 6th and 7th grades in an effort to disentangle the confounding effects of timing of middle school entrance and TDBD status. They find negative effects of bottom dog status on female students. These effects, however, may reflect differences across students rather than the position of students in a school. For example, parents of unhappy students attending K-8 schools may use the summer before 7th grade as an opportunity to change school districts while allowing their children to begin a junior high school on time. This is explained in Cook et al. (2008, p. 108):

> [I]t is possible that any correlation between school grade span and the measured infraction rate reflects nonrandom sorting of students. Parents may choose where to live or whether to keep their children in the public schools based, in part, on the configuration of grades. That sort of selection process may influence the characteristics of the student body in ways not necessarily reflected in observed indicators.

School choice could bias an estimate of the effect of TDBD status on student experiences, overestimating the TDBD effect in the previous example because only the relatively happy 7th grade students stay enrolled long enough to become top dogs in 8th grade. As a result, previous research on the TDBD phenomenon falls short of establishing a causal relationship because it does not convincingly address the plausible endogeneity of enrollment decisions over time.

Methods

Data

We use richly detailed, longitudinal, student-level data from the NYC Department of Education (NYCDOE). These data include indicators for eligibility for free or reduced price lunch, ethnicity/race, gender, English language learner status, English is primary home language, special education status, borough of residence, birth date, height, weight, and New York State standardized examination test scores. Student height and weight are collected annually as a part of NYC's Fitnessgram (over 85% of middle school students are assessed). We link these data to student responses on the NYC School Survey. We match students to school attended and use school-level longitudinal data on enrollment by grade and grades served (grade span) from the New York State School Report Cards (SRC).

NYC School Survey data include student responses to more than 60 questions regarding student experiences, school environment, and non-academic outcomes. Starting in the 2006-07 (2007) school year, the NYCDOE distributed the NYC School Survey to all students in grades 6-12. From 2008-2011 (our sample period), about 90% of general education and part-time special education middle school students respond to the survey each year, with response rates varying marginally by question.

Our sample includes two cohorts of NYC public school general education and part-time special education students who entered 6th grade for the first time in 2008 or 2009, attended NYC public schools in 3rd grade, made standard academic progress from grades 6-8, and responded to the NYC School Survey in at least one year.[1] The sample spans a four year period from 2008 through 2011, and each student is observed for three years. All together, our sample includes more than 500 unique schools, and 90,000 unique students.

Measures

The main outcome variables are student reports of experiences in school. While there are many measures of school learning environment in the NYC School Survey, we rely on components that most closely match the outcomes studied in the TDBD literature – bullying, safety, and belonging. Following previous research, we use student-level survey data to construct measures of student experiences. As shown in Table 1, the measures include student reports of the frequency of school bullying, fights, and gang activity; how frequently a student stays home due to feeling unsafe; whether a student feels safe in school hallways, bathrooms, and locker rooms; and the extent to which a student feels he or she is known by school adults and welcome in school. Survey respondents provide answers on a scale of 1 to 4 for each question. We use indicator variables for each measure of bullying, safety, and belonging, which take a value of 1 if the student reports the activity happens more frequently ("all of the time" or "most of the time") or the student reports agreement ("agree" or "strongly agree") and a value of 0 if the activity happens less frequently ("some of the time" or "never") or the student reports disagreement ("disagree" or "strongly disagree").[2]

<Table 1 near here>

We construct seven binary outcome variables in all, which we term *Bullying*, *Fights*, *Gangs*, *StayHome*, *SafeSchool*, *Known*, and *Welcome* (see Table 1). Previous researchers have used the same measures of bullying, safety, and belonging and have found that they provide apt measures of the school learning environment (Lacoe, 2013; Schwartz, Stiefel, & Wiswall, 2016). For example, in order to assess the construct validity of bullying measures, Lacoe (2013) compares student survey responses to school-level administrative measures of school violence. She finds that student responses to these questions are highly correlated in the expected direction with violence reported annually through the New York State Violent and Disruptive Incident Reporting (VADIR) system (Lacoe, 2013, fig. 3).

Still, to address potential weaknesses in individual measures of the learning environment, we use multiple survey questions, exploring whether the TDBD effect is consistently observed across multiple measures of bullying, safety, and belonging.[3] While answers to individual survey questions are likely imperfect, collectively they provide a clear picture of student perceptions of the school environment. For example, our bullying measure asks whether "students threaten or bully other students at school" and, therefore, does not directly identify whether students are actually bullied themselves (student self-reports such as "I have been bullied at school this year" or "I have bullied others at school this year" are the most commonly used measure in the bullying literature and peer nominations are a distant second; see Nakamoto & Schwartz, 2010; Kim & Leventhal, 2008). Some NYC School Survey questions on student safety, however, ask about individual experiences (for example, "I stay home because I don't feel safe at school" and "I am safe in the hallways, bathrooms, and locker rooms at my school"). Previous research suggests these measures collectively offer a strong indication of student perceptions of bullying and safety in school (see Kim & Leventhal, 2008).

Our main independent variable is student TDBD status, defining top dogs as students at the top of a grade span and bottom dogs as those at the bottom of a grade span. "Middle dogs" are enrolled in any grade between the top and bottom (e.g., 7th grade students are middle dogs in a 6-8 school).

There are a wide variety of school grade spans available to middle school students in NYC during this period. Of these, five major school types (6-8, 6-12, 5-8, K-6, K-8) each enroll more than 2% of the 6th grade population, and more than a dozen others enroll small shares of students. Importantly, these grade spans provide within-grade variation in TDBD status. For example, 6th grade students would be top dogs in a K-6, but bottom dogs in a 6-8. Of equal importance, these different grade spans are unlikely to exhibit significant resource and staffing differences. NYC funding for educating 6th, 7th, and 8th grade students is the same regardless of grade span and NYC students in these grades generally have subject-specific teachers in both environments rather than a single teacher for all subjects. Moreover, class sizes are similar when comparing students in the same grades (Rockoff and Lockwood, 2010). Taken in sum, NYC provides a unique context for which within-grade variation in TDBD status can be observed, but for which it is unlikely that staffing and resource differences between elementary and middle schools would explain differences in performance (Rockoff and Lockwood, 2010).

"Heap size" captures grade span length, measured as the number of grades underneath top dogs in the same school. For example, the heap size of a 6-8 school is two. We use the natural logarithm of heap size, because additional grades in schools with long grade spans affect the relative position of students less than those with short grade spans.[4]

We also distinguish "new dogs," who have just enrolled in a new school (most, but not all, of whom are also bottom dogs); and "big dogs," who are tall compared to other students in

their school. Students are new dogs if they attended a different school – in NYC or elsewhere -- in the preceding school year and are "returning" otherwise. New bottom dogs begin a new school at a standard time, but new middle dogs enter a new school midway through the normal grade progression.

Student height is measured once a year as a part of NYC's Fitnessgram program and is available for over 85% of the sample. We measure students' relative height in a school as "zHeight," which is the z-score of height calculated using the mean and standard deviation of height in the student's school in that year. Thus, a student with a zHeight of 1, for example, is one standard deviation taller than the mean student in his or her school; the school mean of z-height is zero and the standard deviation is one. We also measure students' academic achievement on statewide English Language Arts (ELA) and math exams, standardizing scores to mean zero and standard deviation of one (z-scores) based on NYC test takers in a given grade and year.

Descriptive Statistics

Table 2 shows descriptive statistics for the most common grade spans for 6th grade students in 2008 (Table S1 in the online version of this journal shows descriptive statistics for uncommon grade spans). A majority of 6th grade students attend 6-8 schools (and are bottom dogs) and there are significant differences in student characteristics across grade spans. For example, 6th graders in 6-8 schools (bottom dogs) are less likely to be free lunch eligible or part-time special education and more likely to be white or Asian than 6th graders in K-6 (top dogs) and K-8 schools (middle dogs). Students in 6-8 schools are relatively advantaged on observed characteristics compared to students in other grade spans.

<Table 2 near here>

Figure 1 shows 6th grade student responses (2008 6th graders) to survey questions regarding bullying, safety, and belonging for the most common middle school grade spans – those comprising at least 2% of each grade's population citywide (Figure 1 offers binary indicators that are consistent with our empirical strategy; details on all four responses to each question are in Table S2 with 7th and 8th grader responses available upon request of the authors). Among 6th graders, top dogs report feeling safer, less bullying, and greater belonging than bottom and middle dogs. For example, among 6th graders, the top dogs (in K-6 schools) are less likely to report students threaten or bully other students at school most or all of the time (25.8%) than the bottom dogs (in 6-8 or 6-12 schools, reporting 31.0% and 33.0%, respectively). Similarly, on average, middle dogs report better experiences than bottom dogs. Among 6th graders, middle dogs (in 5-8 or K-8 schools), are less likely to report there is gang activity always or most of the time (17.3% and 15.9%, respectively) than the bottom dogs (18.2% and 19.4% for 6-8 and 6-12, respectively).

<Figure 1 near here>

Comparing students in different grades within the same grade span, top dogs again report better experiences than bottom dogs. Among students attending 6-8 schools, a greater share of top dogs (80.2% of 8th graders) report they agree (or strongly agree) that most of the teachers, counselors, school leaders, and other adults in school know who they are compared to bottom dogs (73.1% of 6th graders, see Table S2). As another example, top dogs in K-8 schools (8th grade) are less likely to report they (strongly) agree that students get into physical fights at their

school (24.8%) than K-8 middle dogs (37.8% and 32.4% of 6th and 7th graders report fights, respectively). Taken together, the descriptive statistics are consistent with the TDBD hypothesis. We turn next to impact estimates using regression analysis.

Analytic Strategy

Baseline Model

Our baseline model links bullying, safety, belonging, and academic performance to TDBD status as well as a set of student characteristics, and grade, year, and student fixed effects, as follows:

(1) $BSB_{ist} = \beta_0 + \beta_1 TD_{ist} + \beta_2 BD_{ist} + \beta_3 G_{ist} + \beta_4 X_{it} + \alpha_i + \mu_{ist}$

where BSB_{ist} is a bullying, safety, or belonging outcome for student i in school s in year t; TD_{ist} is a variable indicating top dog status; BD_{ist} is variable indicating bottom dog status; G_{ist} is a series of binary variables indicating if student i is in grade 6, 7, or 8; X_{it} is a vector of time-varying student characteristics (English language proficiency, free or reduced-price lunch, special education participation, and borough of residence) and, in models excluding student fixed effects, X also includes time-invariant student characteristics (gender, ethnicity, if English is spoken at home); α_i is a student fixed effect; and μ_{ist} is an error term. Standard errors are clustered at the school-year level in order to correct for correlations among students who are sharing a school and have the same TDBD status.

Even conditional on observed student characteristics and student fixed effects, students may still sort into TDBD status. That is, students may choose their school environment by enrolling in a grade span in a given grade and effectively selecting to be top or bottom dog in a particular grade (as described previously, this may bias estimates of the TDBD effect).[5] As outlined earlier, to address possible endogeneity of TDBD status, we use grade spans of students'

3rd grade schools as instruments for TD and BD status, following Rockoff and Lockwood (2010) and Schwerdt and West (2013). TDBD status and grade span are closely related in any year, but it is unlikely that grade span three to five years before an observation reflects current middle school learning environment, except through its effect on TDBD status. To be specific, we instrument for TD_{ist} and BD_{ist} (using two-stage least squares) in model (1) using a vector of binary variables indicating if student *i* (who attends school s) is exposed to a particular grade span in 3rd grade.[6]

Our key coefficients are β_1 and β_2, which capture the effect of top dog and bottom dog status on student perceptions of the learning environment, respectively. For some of these outcomes, such as *Bullying*, *Fights*, *Gangs*, and *StayHome*, negative coefficients reflect a better learning environment, indicating students are less likely to report these negative conditions. For the other outcomes, including *SafeSchool*, *Known*, and *Welcome*, positive coefficients indicate a better learning environment. Negative β_1 (β_2) coefficients in *Bullying*, *StayHome*, *Fights*, and *Gangs* models indicate top dogs (bottom dogs) reported less bullying, staying home because they feel unsafe, fights, and gangs compared to middle dogs. If β_1 equals -0.01 in the *Fights* model, for example, it indicates that top dogs are one percentage point less likely to report frequent physical fights at school (all or most of the time) compared to middle dogs.[7] Positive β_1 (β_2) coefficients in *SafeSchool*, *Known*, and *Welcome* models indicate top dogs (bottom dogs) reported feeling safer, more known, and more welcome in school compared to middle dogs. If β_1 equals 0.01 in the *SafeSchool* model, for example, it indicates that top dogs are one percentage point more likely to report (strongly) agreeing that they feel safe in the hallways, locker rooms, and bathrooms at school compared to middle dogs. We henceforth report estimates as "less

likely" or "more likely" for the models that estimate the effect of top dog (or bottom dog) status on *Bullying* (and so on).

Heterogeneity by Grade and Grade Span Length

We explore effect heterogeneity to analyze when students obtain the greatest boost from top dog status. We assess heterogeneity of TDBD effects by grade, as follows:

(2) $BSB_{ist} = \beta_0 + \beta_1{}^{G}TD_{ist} + \beta_2{}^{G}BD_{ist} + \beta_3 G_{ist} + \beta_4 X_{it} + \alpha_i + \mu_{ist}$

where all variables are as previously defined and superscript "*G*" indicates that student grade is interacted with the variable. (For example, ${}^{6}TD_{ist}$ takes a value of 1 if a 6th grade student attends a school in which 6th graders are top dogs, -- such as K-6 and 3-6 schools -- and takes a value of 0 otherwise.) This model will shed light on questions such as whether the benefits of top dog status are greater in 8th grade or 6th grade. The coefficients in the vector β_1 (β_2) capture the relationship between grade and top dog (bottom dog) status.

As suggested previously, the size of the TDBD effect may also depend on grade span length (heap size). We estimate a model in which we interact TDBD status with heap size to test whether longer grade spans (bigger heaps) are associated with larger TDBD effects. In the heap size models, β_1 and β_2 are estimated coefficients of the interaction effect of heap size and TDBD status (similar to the interaction effects outlined previously for grade and TDBD status).[8]

Digging Deeper into the TDBD Effect

One explanation for the TDBD phenomenon is that it merely reflects bottom dog difficulties adjusting to a new school. Thus, the estimated bottom dog effect may reflect the relative disadvantage of being new to a school rather than bottom dog status *per se*. Conversely,

one might instead argue that being new to a school is a component of bottom dog status. We estimate how much of the TDBD effect can be explained by new student status using variation in student entry into new schools. Importantly, middle dogs are often new students as well. For example, a student who "graduates" from a K-6 and enters a 6-8 school as a 7th grader is a new middle dog. To explore new student status as a possible explanation for the TDBD effect, we interact TDBD status with new student status to test the extent to which new student status explains the TDBD effects.

Another feature of the TDBD phenomenon is that top dogs tend to be taller than bottom dogs, and, as discussed previously, taller students may have better experiences. Top dogs are in higher grades than bottom dogs within the same school and are usually taller, especially in middle school, when students are still growing. The TDBD effect could be driven or exacerbated by student height, as being relatively tall might be a component of top dog status (height could be an explanation or a mechanism). We explore this in two ways. First, we explore the moderating effect of height by interacting TDBD status with relative height in a school. Second, we explore the potential mediating effect by including height as a control variable. We use both zHeight and zHeight-squared to allow for possible non-linearity in the relationship between height and perceptions of the school environment (since there could be distinct advantages to being of average height). Here, we restrict our analyses to the 85% of the sample with height data and re-estimate our preferred model.

Effects on Academic Achievement

Finally, we explore the impact of TDBD status on student academic achievement, contributing to the growing grade span literature. We estimate the *causal* relationship between

TDBD status and academic performance, changing the outcome variables in model (2) from BSB_{ist} to $TEST_{ist}$ – where $TEST_{ist}$ is the math or ELA exam z-score for student *i* in school *s* and in year *t*, using the same vector of instrumental variables and student fixed effects. The coefficients in the vector β_1 (β_2) now capture the relationship between grade, top dog (bottom dog) status, and academic achievement.

As discussed previously, TDBD status could affect academic achievement through students' perceptions of the school learning environment. Unfortunately, we cannot causally assess the effects of bullying, safety, and belonging on academic achievement. To do so, we would need to estimate student test score, y, as a function of *both* TDBD status and student learning environment. In such a model, however, OLS estimates will be biased because student learning environment, TDBD status, and academic achievement are endogenous. We cannot use the same instrumental variables strategy used elsewhere in this paper to resolve the endogeneity problem, because we would need two unique sets of instruments, one for TDBD status and another for school learning environment. But we have only one set of instruments (the vector of variables reflecting grade span of 3rd grade school). Instead of causal estimates, we offer *descriptive* evidence on the role perceptions of the school environment play in academic achievement, which will complement the *causal* evidence on the impact of TDBD status on both student experiences and academic achievement. Using a student fixed effects model, we estimate the relationship between bullying, safety, and belonging and academic achievement. These models are specified as follows:

(3) $TEST_{ist} = \gamma_0 + \gamma_1 BSB_{ist} + \gamma_2 X_{it} + \alpha_i + \mu_{ist}$

where $TEST_{ist}$ is the math or ELA exam z-score for student *i* in school *s* in year *t*; and all other variables are as previously defined. Results from these models provide descriptive evidence of

the relationship between bullying, safety and belonging, and academic achievement in 6th, 7th, and 8th grades but do not provide causal evidence that the impact of TDBD status on academic achievement operates directly through perceptions of the school learning environment. For example, it is also plausible that achievement affects student perceptions of the learning environment and TDBD's effect on bullying, safety, and belonging operates through declines in academic achievement (consistent with reduced academic self-concept found in Eccles et al., 1991). Still, these estimates may show that the grade span effect on academic achievement could possibly operate through the TDBD effect on bullying, safety, and belonging.

Results

Baseline Results

Table 3 shows OLS, IV, and IV with student fixed effects estimates of the TDBD effect in the middle school grades. In the OLS model (Panel A), we estimate that top dogs are less likely to report that there is gang activity in their schools (1.8 percentage point decrease in probability of reporting *Gangs*), and more likely to report feeling safe in hallways, locker rooms, and bathrooms (2.8 percentage point increase in probability of reporting *SafeSchool*) and that they are known (1.5 percentage point increase in probability of reporting *Known*) than middle dogs. In the OLS model, we further estimate that bottom dogs are more likely to report *Bullying* (2.3 percentage points), *Gangs* (3.9 percentage points), and *StayHome* (1.0 percentage point increase in the probability of reporting staying home from school because he or she feels unsafe), and less likely to report *SafeSchool* (-4.0 percentage points), *Known* (-3.7 percentage points), and *Welcome* (-1.2 percentage points) than middle dogs. Top dogs also are less likely to report *Bullying*, *Gangs* and *StayHome*, but more likely to report *SafeSchool*, *Known* and *Welcome* than

bottom dogs, which we estimate by subtracting the coefficients on bottom dog status from the coefficients on top dog status and testing for significance using t tests. Taken together, these results indicate that top dogs perceive a more favorable school environment than bottom and middle dogs.

<Table 3 near here>

The main results remain largely unchanged when accounting for student selection into timing of top dog and bottom dog status (as determined by middle school grade span) using the instrumental variables strategy. In the IV model (Panel B of Table 3), we estimate that top dog status decreases a student's likelihood to report *Fights*, *Gangs* and *StayHome*, and increases a student's likelihood to report *SafeSchool* and *Known* than middle dogs. The estimated effect of bottom dog status is robust to the instrumental variables model for regressions estimating the effect on *Gangs*, *SafeSchool*, and *Known*. Bottom dog status increases a student's likelihood to report *Gangs* and decreases probability to report *SafeSchool* and *Known*.

Our estimated effects of top dog status are stronger when we address the endogeneity of selection into middle school and control for time invariant student characteristics. The results from our preferred models – which include the instrumental variables and student fixed effects – are shown in Panel C of Table 3. We estimate that top dogs compared to middle dogs are 4.6 percentage points less likely to report bullying, 7.7 percentage points less likely to report fights, 6.8 percentage points less likely to report gangs, and 2.1 percentage point less likely to report staying home due to feeling unsafe in school. Furthermore, top dogs are 10.5 percentage points more likely to report feeling safe in hallways, locker rooms, and bathrooms, 11.5 percentage

points more likely to report feeling they are known, and 3.7 percentage points more likely to report feeling welcome.

In summary, OLS, IV, and IV student fixed effects models all indicate that top dogs fare better than middle and bottom dogs. These estimated effects are quite large, showing marked changes in perceptions of the learning environment for students who are top dogs. Conversely, as models better address causality, bottom dogs fare similarly to middle dogs, but worse than top dogs. This suggests that negative coefficients for bottom dogs in an OLS model might be biased by the effects of selection through, for example, endogenous student mobility.

Heterogeneity by Grade and Grade Span Length

Table 4 presents our preferred model estimates of TDBD effects for each middle school grade. We interact TD and BD with student grade to estimate the differential effects of top dog and bottom dog status in 6^{th} and 8^{th} grades, finding that 6^{th} graders have a greater bump from top dog status than do 8^{th} graders. We find 6^{th} grade top dogs face better learning environments than they would as middle and bottom dogs. For example, as a result of top dog status, 6^{th} grade top dogs are 7.6 percentage points less likely to report *Bullying* and 14.7 percentage points more likely to report *SafeSchool* than they would as middle dogs (see Top - Middle: Grade 6 in Table 4). Similarly, as a result of their status, 6^{th} grade top dogs see a 10.6 percentage point decrease in probability of *Gangs* and a 14.6 percentage point increase in probability to feel *Known* as compared to if they were bottom dogs (see Top - Bottom: Grade 6 in Table 4).

<Table 4 near here>

Conversely, there is little difference in the student experiences of 8th grade top dogs and middle dogs. We find a statistically significant impact for only one outcome: feeling known (5.6 percentage point increase compared to middle dogs). There are no other significant differences in bullying, safety, or belonging between top and middle dog 8th grade students.[9]

While the results shown in Table 4 are mean TDBD impact estimates by grade, they omit indicators for whether students serve as top dogs in short or long grade spans. Figure 2 shows estimates of the relationship between the TDBD effect and heap size (number of grades beneath top dogs in the school). As shown in Figure 2, we find that much of the TDBD effect is explained by students at the top of larger heaps (longer grade spans); as the heap size gets larger, so does the TD effect. For example, we estimate that top dog status has a substantially larger effect on the likelihood students report feeling *Known* in schools with a heap size of 8 (12.4 percentage points) than in schools with a heap size of 2 (7.4 percentage points) compared to middle dogs.

<Figure 2 near here>

Conversely, there is no effect of longer heap size on the BD effect (shown in Table S4). That is, for example, 6th graders at the bottom of a long grade span do not report *Bullying* at a higher rate than 6th graders at the bottom of a short grade span. Students in larger heaps may benefit from increased premiums on top dog status without further hurting bottom dogs. This provides evidence that longer grade spans (heaps) for middle school students, such as K-8 schools, may improve student experiences.[10]

In sum, we find that there is a larger top dog effect in 6th than 8th grade and for students serving as top dogs over larger compared to smaller heap sizes. Further, we find that long grade

spans (larger heaps) do not harm bottom dogs as compared to shorter grade spans (smaller heaps). These results suggest, consistent with developmental theory, that timing of top dog status matters and, further, that longer grade spans may help top dogs more than shorter grade spans.

Digging Deeper into the TDBD Effect

Bottom dog effect or new dog effect?

Figure 3 shows estimates of the TDBD effect for returning and new students side by side with our preferred model TDBD estimates (Table S5 shows results in table form). As shown in Figure 3 Panel A, the top dog effect for returning students is about the same as the top dog effect estimates in our preferred model; all results point in the same direction and are of roughly the same magnitude (though a little smaller). That is, top dog status improves the student learning environment, even among returning students.

<Figure 3 near here>

The results in Figure 3 Panel B, are also consistent with the preferred model estimates: the estimated effect of bottom dog status for new students is not statistically different from new middle dogs.[11] Taken together, the results in Figure 3 provide strong evidence that the TDBD effects presented in this paper are due to TDBD status and are not moderated by new student status.

Top dog effect or big dog effect?

The results shown in Table 5 tease apart the extent to which the TDBD effect is explained by student height. Panel A of Table 5 shows the estimates from our preferred model without height controls, but for the subset of students with height measures. As shown in Panel B,

including controls for relative student height (zHeight and the quadratic form) does not change our primary results much. That is, top dogs benefit from their status independent of the role of height.

<Table 5 near here>

As shown in Panel C of Table 5, the main TDBD effect holds even with the inclusion of the interaction terms between TDBD status and student relative height. Top dog status improves student perceptions of the school environment as compared to middle and bottom dog status. For example, top dog status decreases the probability of reporting *Bullying* by 8.2 percentage points as compared to middle dog status. While student height matters, it does not explain the TDBD effect on perceptions of the learning environment.[12]

Effects on Academic Achievement

Last, we turn to empirical tests of the hypothesis that the TDBD effect explains observed losses in academic performance at the time of middle school entry by estimating the effect of TDBD status on student academic achievement. Again, our instrumental variables fixed effects estimates allows a causal interpretation. As shown in Table 6, bottom dog status hurts academic performance and top dog status improves academic performance in 6th grade. These results highlight that declines in academic performance during transitions to middle schools are, in part, a result of transitioning from top dog to bottom (or middle) dog status.

<Table 6 near here>

Might TDBD affect student academic outcomes through student experience of the learning environment? While we cannot provide causal estimates to test this, we provide correlational evidence in Table 7, which shows results from models linking academic performance to bullying, safety and belonging. To be clear, the relationship could be bi-directional and we are unable to distinguish that here. In Panels A and B of Table 7, we show that students reporting *Bullying*, *Fights*, *Gangs*, and *StayHome* (*SafeSchool, Welcome, and Known*) in the middle grades, have lower (higher) levels of math and ELA achievement, respectively. Further, as shown in Table 7 column (8), many of these relationships are robust to controlling for all measures of student perceptions of the school environment simultaneously and all survey responses are correlated with achievement in the intuitively appealing direction even in cases in which the estimates are not significant. Consistent with the TDBD hypothesis, we find the measures of student experiences jointly significant in predicting math and ELA achievement.

<Table 7 near here>

In summary we find that top dog status improves academic achievement, while bottom dog status is somewhat deleterious to achievement. These estimated effects are quite large, showing marked changes in academic achievement, especially for students who are top dogs. We provide some correlational evidence that this might operate through changes in perceptions of the school learning environment, though other explanations are plausible.

Discussion

This paper offers the first credibly causal evidence on the TDBD hypothesis. Using data on 6th, 7th, and 8th graders in NYC public schools, we find that top dogs are less likely to report problems with bullying or safety, and are more likely to report feeling welcome and belonging in school, compared to bottom dogs. These effects are robust to controls for a variety of student characteristics, student fixed effects, and corrections for potential selection into grade spans. Conversely, bottom dogs are more likely to report bullying, feeling unsafe, and like they do not belong in school than they do as middle or top dogs. Unlike the top dog effect, our results suggest that the bottom dog effect results, at least in part, from student selection. We find moving from elementary to middle school hurts bottom dogs because they lose the top dog status they previously held in their old school. Put differently, the TDBD effect is significant, both substantively and statistically.

Our results also suggest that students may benefit from longer grade spans. We find the top dog effect is larger in schools with longer grade spans (larger heap sizes), while the effect on bottom dogs in longer grade spans (at the bottom of larger heap sizes) is no worse than in schools with shorter grade spans. Moreover, in longer grade spans, the closer students are to the top the better they do; that is, promotion through school improves learning environments.

We also find that there is a larger top dog effect in 6th grade than in 8th grade, which is partially a result of heap size (6th grade is most often a terminal grade for long grade spans, while 8th grade is most often a terminal grade for short grade spans).[13] In part, this result may reflect the greater ability of 8th graders to adapt to environments in which they do not have top dog status than 6th graders due to developmental differences across early adolescents.

We explore possible mechanisms for the TDBD effect, including (1) student height and (2) whether a student is returning or new to a school. First, we find that the jump in perceptions of learning environment for top dogs comes from the status afforded to them by their grade and not their height. Second, we find the top dog effect holds even when controlling for continued enrollment in the same school. While being a new student and student height affect student experiences, they do not drive or explain the TDBD phenomenon.

We suggest that other plausible explanations for the negative consequences of the middle school environment on the whole are unlikely. First, we estimate the TDBD effect in models that include student fixed effects and, therefore, estimate the impact of TD and BD status within students over time. In alternative model specifications, we find, for example, that 8th graders in a 6-8 middle school environment have better perceptions of the learning environment than they did as 6th graders in 6-8 schools, suggesting that there is an independent top dog effect in addition to any plausible negative consequences of the middle school environment on the whole. Further, in the context of NYC, the difference in school characteristics between middle schools and elementary schools is not pronounced for students in 6th, 7th, and 8th grades. For example, elementary schools are not better resourced than middle schools in NYC (Rockoff and Lockwood, 2010). Moreover, while there are differences across grade spans in terms of class size and whether students have subject-specific teachers, these differences are negligible for students within the same grade (Rockoff and Lockwood, 2010). For example, 6th, 7th and 8th grade students attending K-8 schools are typically assigned subject-specific teachers. Thus, we also minimize the possibility of these otherwise plausible explanations for top dog effect identified in this paper.

We also provide evidence on academic achievement. While previous research offers many explanations for drops in student performance at middle school entry, our results suggest that the TDBD phenomenon is an important one of them. We find that top dog status improves academic performance, using the same IV we use for the bullying, safety, and belonging outcomes. We further provide descriptive evidence that the top dog effect on academic performance could plausibly operate through perceptions of the learning environment, though these results simply model the correlations between perceptions of the learning environment and academic achievement. Further work is needed to determine if the TDBD effect on academic achievement operates through perceptions of the school learning environment.

While every school has both top and bottom dogs, grade organization defines when and how frequently students serve as top and bottom dogs. Thus, our results can inform policy decisions on school organization. We find, for example, that the top dog premium increases in the length of the grade span. Moreover, our results offer insight into how school administrators may want to target their resources. We find, for example, that even returning students in the middle of a grade span feel less like they belong (less likely to report they are *Known* or *Welcome*) than top dogs. This suggests benefits to targeting resources to foster more welcoming environments for middle dogs, and not just new students. Further, our evidence links TDBD status to academic outcomes, suggesting that fostering safer environments for bottom dogs may ease their transition to middle school and improve academic performance as well.

Our results provide empirical support for the TDBD hypothesis, even after addressing endogeneity of school grade span choice and time-invariant student characteristics. This suggests that the effect of TDBD status on student experiences ought to be considered to make optimal decisions on grade span length. In particular, the evidence in this paper suggests that longer

grade spans that enable middle grade students to serve as relative top dogs would improve student experiences in school and academic achievement. Attending a K-8 school as opposed to a 6-8 school, for example, would benefit 6th graders because they would no longer be "new dogs" in the school, would benefit 7th grade students because they would hold a higher relative position than had they attended a 6-8 school, and would benefit 8th graders because they would hold top dog status over a longer heap size. While wholesale school reorganization nationwide would be costly, there may be more opportunity to make such changes in urban areas, especially if such school districts are growing or declining and K-8 schools provide more efficient building use. Moreover, in places that do not reorganize elementary and middle school grade spans, this paper provides strong evidence that resources should be committed to fostering safe learning environments for students who are not top dogs. While we can only speculate on how these policy recommendations would impact elementary-aged students, we are more certain of the positive effects they can have on the experiences and academic achievement of students in the middle grades.

Notes

[1] Students who are top dogs or bottom dogs in 7th grade (K-7, 6-7, 7-8, 7-10, 7-11, and 7-12 schools) are excluded, because these grade spans most typically reflect schools that are phasing in or out, respectively. 4,301 7th grade students were excluded in all. While some of these configurations are standard in other school districts (such as 7-8), they are not in NYC. Similarly, students who are bottoms dogs in 8th grade (schools serving 8th graders only and 8-12 schools) are excluded, because all schools with these grade spans are being phased out. 612 8th grade students were excluded in all.

[2] We test the sensitivity of the results to additional binary configurations (for example, bullying takes a value of 0 for students answering "never" only and a 1 for students answering "some of the time," "most of the time," or "all of the time"). The results reported in this paper are not sensitive to alternative variable construction and are available upon request.

[3] We assess the uniqueness of these measures using a factor analysis, finding the seven school environment outcomes fall into three main factors, but further finding that the uniqueness of these measures are relatively high, ranging from 0.59 to 0.92. We, therefore, see value in reporting the results for the seven individual outcomes and do so throughout this paper.

[4] The percent change better captures how additional grades served affect student relative position than linear measures, but results are similar using both linear and logarithmic heap size.

[5] We use a GMM distance test to assess the endogeneity of the TD and BD variables, comparing OLS results to IV results. We reject the null hypothesis that TD and BD are exogenous for all but two of the nine outcomes (the two test score outcomes and the seven survey outcomes) we use in this paper, suggesting students sort even conditional on observed student characteristics.

[6] We use a vector of binary variables to instrument for TDBD status, which include indicators for 3rd grade enrollment in a K-5, K-6, K-8, K-4, 3-6, 3-5. 1-5, or K-12 school, among others, and an interaction of each of these indicators with current grade.

[7] As a reference, when Fights is the outcome of interest, the fully specified equation for model (1) takes the form: $Fights_{ist} = \beta_0 + \beta_1 TD_{ist} + \beta_2 BD_{ist} + \beta_3 G_{ist} + \beta_4 X_{ist} + \alpha_i + \mu_{ist}$.

[8] In addition to the heap size models shown in this paper, we also estimate individual TDBD effects for the most common middle school grade spans (that is 6-8, K-6, K-8, 6-12, and 5-8). We estimate the effects of student grades within a given grade span as they are promoted through grades 6, 7, and 8 as an "intention-to-treat" model, fixing students to grade span in 6th grade: $BSB_{ist} = \beta_0 + \beta_1{}^{GS} TD_{ist} + \beta_2{}^{GS} BD_{ist} + \beta_3 G_{ist} + \beta_4 X_{ist} + \alpha_i + \mu_{ist}$, where all variables are as previously defined and superscript "GS" indicates that student i's grade span in 6th grade is interacted with the variable (for example, top dog status). These results are consistent with the findings in shown in Figure 2 and are available upon request.

[9] We also estimate the impact of 8th grade top dog status by tracking individual students as they move from 7th to 8th grade (rather than comparing different students in these grades). We find 8th grade top dog status reduces Bullying and Gangs over 7th grade middle dog status (a 4.0 and 4.2 percentage point decline, respectively), suggesting learning environments improve as students age through school.

[10] As noted previously, we also estimate TDBD effects as students move from 6th to 8th grade in common grade spans. Estimates are shown in Table S3. Consistent with the heap size estimates, these results show that schools with longer grade spans (K-8, 6-12, and 5-8) have larger TDBD effects than those with shorter grade spans (6-8). Moreover student promotion within a grade

span improves perceptions of the learning environment as students move from bottom to middle and eventually to top dog status.

[11] Returning bottom dogs and new top dogs are exceedingly rare and may reflect time varying unobserved characteristics of students or their schools. The standard errors for the impact estimates of these two statuses are very large, and the estimates themselves should be interpreted with caution. The purpose of this analysis is to test the robustness of the TDBD effect for returning top dogs (as compared to returning middle dogs) and new bottom dogs (as compared to new middle dogs). Not to estimate, for example, if returning bottom dogs (who almost exclusively attend schools in the process of phase out) differ from new bottom dogs.

[12] Given differences in male and female growth patterns and social norms around height, we estimate the effects of TDBD status and height separately by gender. These results are largely consistent with the results in Table 5 and across both genders. The results of TDBD effect by gender and height are available upon request of the authors.

[13] K-6 is a much more common grade span than 3-6. 6-8 is much more common than K-8 or 5-8.

References

Akos, P. (2006). Extracurricular participation and the transition to middle school. *RMLE Online: Research in Middle Level Education*, *29(9),* 1-9.

Alspaugh, J. W. (1998). Achievement loss associated with the transition to middle school and high school. *Journal of Educational Research*, *92(1),* 20-25.

Borg, M.G. (1999). The extent and nature of bullying among primary and secondary schoolchildren. *Educational Research*, *41(2),* 137-153.

Bloom, H. S., & Unterman R. (2012). Sustained Positive Effects on Graduation Rates Produced by New York City's Small Public High Schools of Choice. New York: MDRC.

Blyth, D.A., Hill, J. P., & Smith, C. K. (1981). The influence of older adolescents on younger adolescents: Do grade level arrangements make a difference in behaviors, attitudes, and experiences? *Journal of Early Adolescence*, *1,* 85-110.

Blyth, D.A., Simmons, R.G., & Bush, D. (1978). The transition into early adolescence: A longitudinal comparison of youth in two educational contexts. *Sociology of Education, 51(3),* 149-162.

Byrnes, V., & Ruby, A. (2007). Comparing achievement between K–8 and middle schools: A large-scale empirical study. *American Journal of Education*, *114,* 101–135.

Carolan, B.V., Weiss, C.C., & Matthews, J.S. (2015). Which middle school model works best? Evidence from the early childhood longitudinal study. *Youth & Society*, *46(5),* 591-614.

Cook, P. J., MacCoun, R., Muschkin, C., & Vigdor, J. (2008). The negative impacts of starting middle school in sixth grade. *Journal of Policy Analysis and Management*, *27(1),* 104-121.

Cordes, S., Schwartz, A.E., & Stiefel, L. (2014). Does residential mobility harm school performance? *Working Paper, Institute for Education and Social Policy*. Retrieved from author.

Eccles, J.S. (1999). The development of children ages 6 to 14. *The Future of Children*, *9(2),* 30-44.

Eccles, J.S., Lord, S., & Midgley, C. (1991). What are we doing to adolescents? The impact of educational contexts on early adolescents. *American Journal of Education, 99(4),* 521-542.

Eccles, J.S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education*, *3,* 139-186.

Eccles, J.S., Wigfield, A., Midgley, C., Reuman, D., Iver, D.M., & Feldlaufer, H. (1993). Negative effects of traditional middle schools on students' motivation. *The Elementary School Journal*, *93(5),* 553-574.

Fuster, J.M. (2002). Frontal lobe and cognitive development. *Journal of Neurocytology*, *31(3-5),* 373-385.

Graham, S., Bellmore, A., & Juvonen, J. (2003). Peer victimization in middle school: When self and peer views diverge. *Journal of Applied School Psychology*, *19,* 117–137.

Hanushek, E. A., Kain, J.F., & Rivkin, S.G. (2004). Disruption versus Tiebout improvement: The costs and benefits of switching schools. *Journal of Public Economics*, *88(9),* 1721-1746.

Howley, C. (2002). Grade-span configurations. *School Administrator, 59(3)*, 24-29.

Keating, D.P. (1990). Adolescent thinking. In S.S. Feldman & G.R. Elliott (Eds.) *At the threshold: The developing adolescent* (54-89). Cambridge, MA: Harvard University Press.

Kim, Y.S. & Leventhal, B. (2008). Bullying and suicide. A review. *International Journal of Adolescent Medicine and Health*, *20(2),* 133–154.

Lacoe, J.R. (2013). Too scared to learn? The academic consequences of feeling unsafe at school. *Working Paper #02-13, Institute for Education and Social Policy*.

Lee, V. E., & Smith, J. B., (1997). "High school size: Which works best and for whom?" *Educational Evaluation and Policy Analysis, 19(3),* 205-227.

Nakamoto, J. & Schwartz, D. (2010). Is Peer Victimization Associated with Academic Achievement? A Meta-analytic Review. *Social Development*, *19(2),* 221-242.

Napier, P. (2008). *Negotiating the complexities of elementary grade span reconfiguration: A review of literature*. Boston, MA: Boston College.

Offenberg, R. M. (2001). The efficacy of Philadelphia's K-to-8 schools compared to middle grades schools. *Middle School Journal, 32(4),* 23-29.

Pellegrini, A.D., Long, J.D., Solberg, D., Roseth, C., Dupuis, D., Bohn, C., & Hickey, M. (2010). Bullying and social status during school transitions. In S.R. Jimerson, S.M. Swearer, D.L. Espelage, E.W. Gutgsell & J.M. Gutgsell (Eds.), *International Handbook of School Bullying: An International Perspective* (199-210). New York: Routledge.

Poiner, J. (2015, April 10). Should districts get rid of middle schools? *Ohio Gadfly Daily*. Retrieved from http://edexcellence.net/articles/should-districts-get-rid-of-middle-schools?utm_source=Fordham+Updates&utm_campaign=19bd655b51-20150419_LateLateBell4_17_2015&utm_medium=email&utm_term=0_d9e8246adf-19bd655b51-71539577

Powell, A.G., Farrar E., & Cohen D.K., (1985). *The shopping mall high school: Winners and losers in the educational marketplace.* Boston: Houghton Mifflin Company.

Rockoff, J. E., & Lockwood, B. B. (2010). Stuck in the middle: Impacts of grade configuration in public schools. *Journal of Public Economics, 94(11-12),* 1051-1061.

Roeser, R.W., Eccles, J.S., & Sameroff, A.J. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology, 10(2),* 321-362.

Rubenstein, R., Schwartz, A.E., Stiefel, L., & Zabel, J. (2009). Spending, Size, and Grade Span in K-8 Schools**.** *Education Finance and Policy*, *4(1),* 60-88.

Schwartz, A.E., Stiefel, L., & Cordes, S. (2016). Moving Matters: The Causal Effect of School Mobility on School Performance. *Education Finance and Policy.* 10.1162/edfp_a_00198

Schwartz A. E., Stiefel, L., Rubenstein, R., & Zabel, J. (2011), "The Path Not Taken: How Does School Organization Affect 8th Grade Achievement?" *Educational Evaluation and Policy Analysis*, *93(3),* 293-317.

Schwartz, A. E., Stiefel, L., & Wiswall, M. (2016). Are all schools created equal? Learning environments in small and large public high schools in New York City. *Economics of Education Review, 52,* 272-290.

Schwerdt, G. and West, M. R. (2013). The impact of alternative grade configurations on student outcomes through middle and high school. *Journal of Public Economics, 97(1-2),* 308-326.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. New York, NY: Gruyter.

Voss, L.D. & Mulligan, J. (2000). Bullying in school: are short pupils at risk? Questionnaire study in a cohort. *BMJ*, *320,* 612-613.

Wang, M.T. & Holcombe, R. (2010). Adolescents' perceptions of school environment, Engagement, and academic achievement in middle school. *American Educational Research Journal, 47(3),* 633-662.

Weiss, C. C., & Kipnes, L. (2006). Reexamining middle school effects: A comparison of middle grades students in middle schools and K-8 schools. *American Journal of Education*, *112(2),* 239-272.

Yurgelun-Todd, Deborah. (2007). Emotional and cognitive changes during adolescence. *Current Opinion in Neurobiology*, *17(2),* 251-257

Table 1. Measures of Bullying, Safety, and Belonging

Category	NYC School Survey Question	Variable Name	=1 if respond
Bullying	"Students threaten or bully other students at school"	Bullying	"Most" or "All of the Time"
Safety	"Students get into physical fights at my school"	Fights	
	"There is gang activity in my school"	Gangs	
	"I stay home because I don't feel safe at school"	StayHome	
	"I am safe in the hallways, bathrooms, and locker rooms at my school"	SafeSchool	"Agree" or "Strongly Agree"
Belonging	"Most of the teachers, counselors, school leaders, and other adults I see at school every day know my name or who I am"	Known	
	"I feel welcome in my school"	Welcome	

Table 2. Descriptive Statistics, 6th Grade Students, by School Grade Span, 2008

	Bottom Dogs		Middle Dogs		Top Dogs
	6-8	**6-12**	**5-8**	**K-8**	**K-6**
% in Grade Span	64.8	3.0	4.2	8.7	6.9
Age	11.2	11.3	11.2	11.3	11.3
Percent:					
Female	50.9	53.7	52.3	51.4	51.5
Ethnicity					
Black	24.5	36.5	28.2	39.0	30.3
Hispanic	38.0	45.4	54.4	36.8	39.4
Asian	20.1	6.7	10.1	10.8	20.4
White	17.5	11.3	7.3	13.4	9.9
English as a home language	52.4	67.2	46.2	63.4	52.8
LEP (Limited English Proficiency)	9.5	6.1	13.0	7.6	8.6
Part-Time Special Education	11.1	13.8	6.5	12.2	11.6
Free Lunch Eligible	70.0	78.4	74.8	75.0	75.8
Reduced Price Lunch Eligible	10.1	7.7	10.6	8.1	8.8
Residential Borough					
Manhattan	10.6	20.0	10.2	16.1	17.1
Bronx	17.2	34.8	33.4	26.7	14.0
Brooklyn	34.4	39.4	26.0	33.4	4.3
Queens	28.1	5.6	30.2	23.4	64.3
Staten Island	9.7	0.2	0.0	0.0	0.0
N	33,624	1,545	2,154	13,086	3,588

Sample includes 2008 6th graders who answer all seven survey questions included in this paper and who attend schools with grade spans enrolling at least 2% of the 6th grade student population. For other grade spans see Table S1.

Table 3. Regression Results, Baseline, Instrumental Variables and IV Student FE Models

	Bullying	Fights	Gangs	StayHome	SafeSchool	Known	Welcome
A. OLS							
Top Dog	-0.003	0.008	-0.018**	-0.001	0.028***	0.015*	0.006
	(0.010)	(0.014)	(0.008)	(0.003)	(0.010)	(0.009)	(0.006)
Middle Dog	—	—	—	—	—	—	—
Bottom Dog	0.023**	0.020	0.039***	0.010***	-0.040***	-0.037***	-0.012**
	(0.012)	(0.016)	(0.009)	(0.003)	(0.012)	(0.011)	(0.006)
Top – Bottom	-0.026**	-0.012	-0.057***	-0.011***	0.068***	0.052***	0.018***
Student FE	NO	NO	NO	NO	NO	NO	NO
IV	NO	NO	NO	NO	NO	NO	NO
B. IV							
Top Dog	-0.003	-0.047*	-0.044**	-0.014**	0.073***	0.050**	0.016
	(0.021)	(0.029)	(0.017)	(0.007)	(0.020)	(0.022)	(0.013)
Middle Dog	—	—	—	—	—	—	—
Bottom Dog	0.013	-0.012	0.038***	0.006	-0.026*	-0.075***	0.003
	(0.015)	(0.020)	(0.012)	(0.005)	(0.015)	(0.016)	(0.009)
Top – Bottom	-0.016	-0.035	-0.082***	-0.020***	0.099***	0.125***	0.013
Student FE	NO	NO	NO	NO	NO	NO	NO
IV	YES	YES	YES	YES	YES	YES	YES
C. IV Stud. FE							
Top Dog	-0.046**	-0.077**	-0.068***	-0.021**	0.105***	0.115***	0.037***
	(0.021)	(0.031)	(0.020)	(0.008)	(0.021)	(0.017)	(0.013)
Middle Dog	—	—	—	—	—	—	—
Bottom Dog	-0.020	-0.071***	-0.001	-0.006	0.013	-0.001	0.020**
	(0.014)	(0.017)	(0.011)	(0.006)	(0.013)	(0.012)	(0.009)
Top - Bottom	-0.026	-0.006	-0.067***	-0.015**	0.092***	0.116***	0.017
Student FE	YES	YES	YES	YES	YES	YES	YES
IV	YES	YES	YES	YES	YES	YES	YES
Observations	252,687	252,887	252,298	250,513	251,688	259,120	259,892
Students	92,840	92,861	92,751	92,379	92,560	94,273	94,418

Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Model controls include: indicators for female, black, Hispanic, Asian, other non-White, free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade and student age. Cohort fixed effects are also included in the OLS and IV specifications. Reference Group = Middle Dogs

Table 4. Regression Results, Same Grade, Different Schools, IV and Student FE

	Bullying	Fights	Gangs	StayHome	SafeSchool	Known	Welcome
Top Dog:							
Grade 6	-0.054**	-0.052*	-0.102***	-0.022**	0.116***	0.113***	0.045***
	(0.022)	(0.031)	(0.022)	(0.009)	(0.022)	(0.019)	(0.013)
Grade 8	-0.024***	-0.043***	-0.010**	0.000	0.048***	0.051***	0.008**
	(0.005)	(0.008)	(0.004)	(0.002)	(0.006)	(0.004)	(0.003)
Middle Dog:							
Grade 6	0.022*	0.060***	0.013	0.007	-0.031***	-0.038***	-0.001
	(0.012)	(0.014)	(0.009)	(0.005)	(0.011)	(0.009)	(0.008)
Grade 7	—	—	—	—	—	—	—
Grade 8	-0.040*	-0.041	-0.042*	0.003	0.041	-0.006	-0.012
	(0.022)	(0.054)	(0.024)	(0.018)	(0.033)	(0.017)	(0.014)
Bottom Dog:							
Grade 6	-0.003	-0.018***	0.004	-0.001	-0.011**	-0.033***	0.020***
	(0.005)	(0.006)	(0.004)	(0.002)	(0.005)	(0.005)	(0.003)
$Top^6 – Middle^6$	-0.076***	-0.112***	-0.115***	-0.029***	0.147***	0.151***	0.046***
$Top^6 – Bottom^6$	-0.051**	-0.034	-0.106***	-0.021**	0.127***	0.146***	0.025*
$Mid^6 – Bottom^6$	0.025*	0.078***	0.009	0.008	-0.020*	-0.005	-0.021**
$Top^8 – Middle^8$	0.016	-0.002	0.032	-0.003	0.007	0.057***	0.020
Observations	252,687	252,298	250,513	252,887	251,688	259,120	259,892
Students	92,840	92,751	92,379	92,861	92,560	94,273	94,418

Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Controls in all models include indicators for free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade, student age, and student fixed effects. Reference Group = Grade 7 Middle Dogs

Table 5. Digging Deeper: Regression Results, Big Dogs and TDBD Effect

	Bullying	Fights	Gangs	StayHome	SafeSchool	Known	Welcome
A. Preferred[†]							
Top Dog	-0.068***	-0.123***	-0.085***	-0.028***	0.120***	0.123***	0.053***
	(0.024)	(0.036)	(0.024)	(0.010)	(0.025)	(0.020)	(0.015)
Middle Dog	---	---	---	---	---	---	---
Bottom Dog	-0.023	-0.071***	0.002	-0.009	0.012	0.000	0.026**
	(0.016)	(0.019)	(0.012)	(0.007)	(0.014)	(0.013)	(0.011)
B. Height							
Top Dog	-0.058**	-0.115***	-0.087***	-0.029**	0.112***	0.123***	0.048***
	(0.028)	(0.043)	(0.029)	(0.012)	(0.029)	(0.023)	(0.016)
Middle Dog	—	—	—	—	—	—	—
Bottom Dog	-0.027*	-0.075***	-0.000	-0.010	0.015	-0.001	0.028**
	(0.016)	(0.019)	(0.012)	(0.007)	(0.014)	(0.013)	(0.011)
zHeight	-0.007	-0.007	0.002	0.001	0.005	-0.004	0.004
	(0.005)	(0.007)	(0.005)	(0.002)	(0.005)	(0.004)	(0.003)
$zHeight^2$	-0.001	-0.001	0.001	0.001**	0.002**	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
C. Interaction							
Top Dog	-0.082***	-0.105***	-0.040***	-0.147***	0.140***	0.101***	0.075***
	(0.0305)	(0.026)	(0.013)	(0.0387)	(0.028)	(0.026)	(0.019)
Middle Dog	—	—	—	—	—	—	—
Bottom Dog	-0.062***	-0.024	-0.015	-0.116***	0.069***	-0.007	0.037**
	(0.021)	(0.017)	(0.010)	(0.024)	(0.021)	(0.020)	(0.015)
TD * zHeight	-0.007*	-0.001	0.001	-0.009**	0.010**	0.001	0.000
	(0.004)	(0.003)	(0.002)	(0.005)	(0.004)	(0.004)	(0.003)
MD * zHeight	-0.010*	0.005	0.004	-0.006	0.007	-0.001	0.000
	(0.005)	(0.004)	(0.003)	(0.006)	(0.005)	(0.005)	(0.004)
BD * zHeight	-0.011*	-0.003	-0.002	-0.014*	0.011*	-0.007	0.003
	(0.007)	(0.006)	(0.003)	(0.008)	(0.007)	(0.006)	(0.005)
Observations	218,416	218,090	216,512	218,536	217,630	223,809	224,514
Students	82,748	82,651	82,277	82,755	82,531	84,167	84,298

Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. Sample includes students with measured height only (over 30,000 observations include no measured height and are included from this analysis). zHeight measures student relative height as compared to other students in the school. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Controls in all models include indicators for free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade, student age, and student fixed effects. Reference Group = Middle Dogs

[†]Sample restricted to students with height data.

Table 6. Regression Results, Impact of TDBD status on Academic Achievement by Grade, IV Student FE

	Reading	Math
Grade 6:		
Top Dog	0.100***	0.160***
	(0.033)	(0.033)
Middle Dog	-0.005	0.017
	(0.021)	(0.026)
Bottom Dog	-0.024***	-0.014
	(0.009)	(0.011)
Grade 7:		
Middle Dog	—	—
Grade 8:		
Top Dog	0.003	-0.003
	(0.010)	(0.012)
Middle Dog	-0.002	-0.116*
	(0.041)	(0.066)
Student FE	YES	YES
IV	YES	YES
Observations	251,566	290,213
Students	86,284	99,591

Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. ELA and Math are student z-scores on achievement exams. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Model controls include: indicators for female, black, Hispanic, Asian, other non-White, free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade and student age, and student fixed effects. Reference Group = Grade 7 Middle Dogs

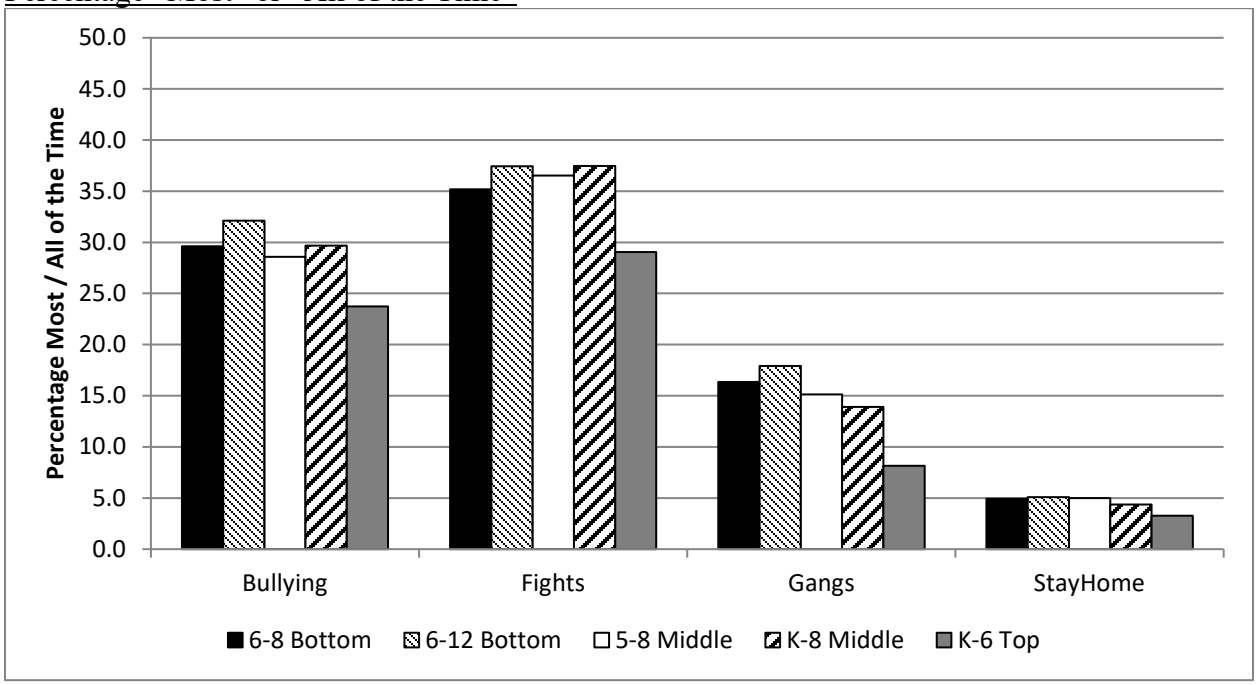
Table 7. Bullying, Safety, and Belonging and Academic Achievement, 6th, 7th & 8th Grades

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Math								
Bullying	-0.013***							0.001
	(0.004)							(0.004)
Fights		-0.018***						-0.009**
		(0.005)						(0.005)
Gangs			-0.026***					-0.012**
			(0.005)					(0.005)
StayHome				-0.054***				-0.036***
				(0.008)				(0.008)
SafeSchool					0.020***			0.010**
					(0.005)			(0.005)
Known						0.012***		0.007
						(0.004)		(0.005)
Welcome							0.035***	0.030***
							(0.005)	(0.006)
Constant	-0.119	-0.146	-0.134	-0.129	-0.171*	-0.184**	-0.196***	-0.171*
	(0.097)	(0.097)	(0.097)	(0.097)	(0.097)	(0.078)	(0.075)	(0.099)
N	303,757	303,445	301,570	304,081	302,846	310,189	311,243	283,457
Students	123,499	123,507	123,310	123,545	123,392	124,137	124,329	121,388
R-squared	0.861	0.861	0.862	0.861	0.861	0.860	0.860	0.865
B. ELA								
Bullying	-0.013***							-0.006
	(0.005)							(0.005)
Fights		-0.011**						0.003
		(0.004)						(0.006)
Gangs			-0.015***					0.002
			(0.006)					(0.006)
StayHome				-0.033***				-0.025***
				(0.007)				(0.008)
SafeSchool					0.011**			-0.006
					(0.005)			(0.005)
Known						0.006		-0.007
						(0.005)		(0.006)
Welcome							0.018***	0.013*
							(0.006)	(0.007)
Constant	-0.119	-0.128	-0.117	-0.119	-0.143	-0.131	-0.138	-0.153
	(0.111)	(0.107)	(0.106)	(0.107)	(0.106)	(0.102)	(0.101)	(0.132)
N	302,281	301,959	300,124	302,578	301,378	308,674	309,645	282,169
Students	123,259	123,265	123,059	123,297	123,149	123,893	124,077	121,150
R-squared	0.780	0.780	0.781	0.780	0.780	0.779	0.779	0.783

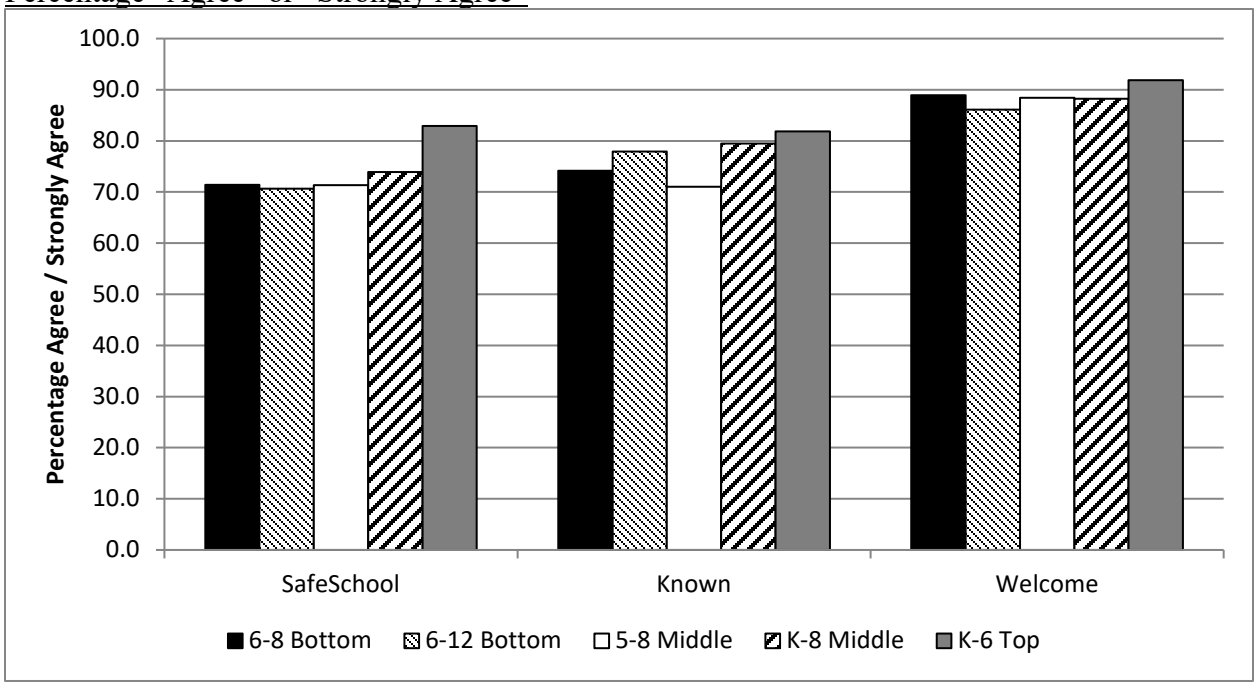
Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. ELA and Math are student z-scores on achievement exams. Controls in all models include: indicators for female, black, Hispanic, Asian, other non-White, free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade, student age, and student fixed effects.

Figure 1. Bullying, Safety, and Belonging by Relative Position and Grade Span, 6th Grade Students, 2008

Percentage "Most" or "All of the Time"

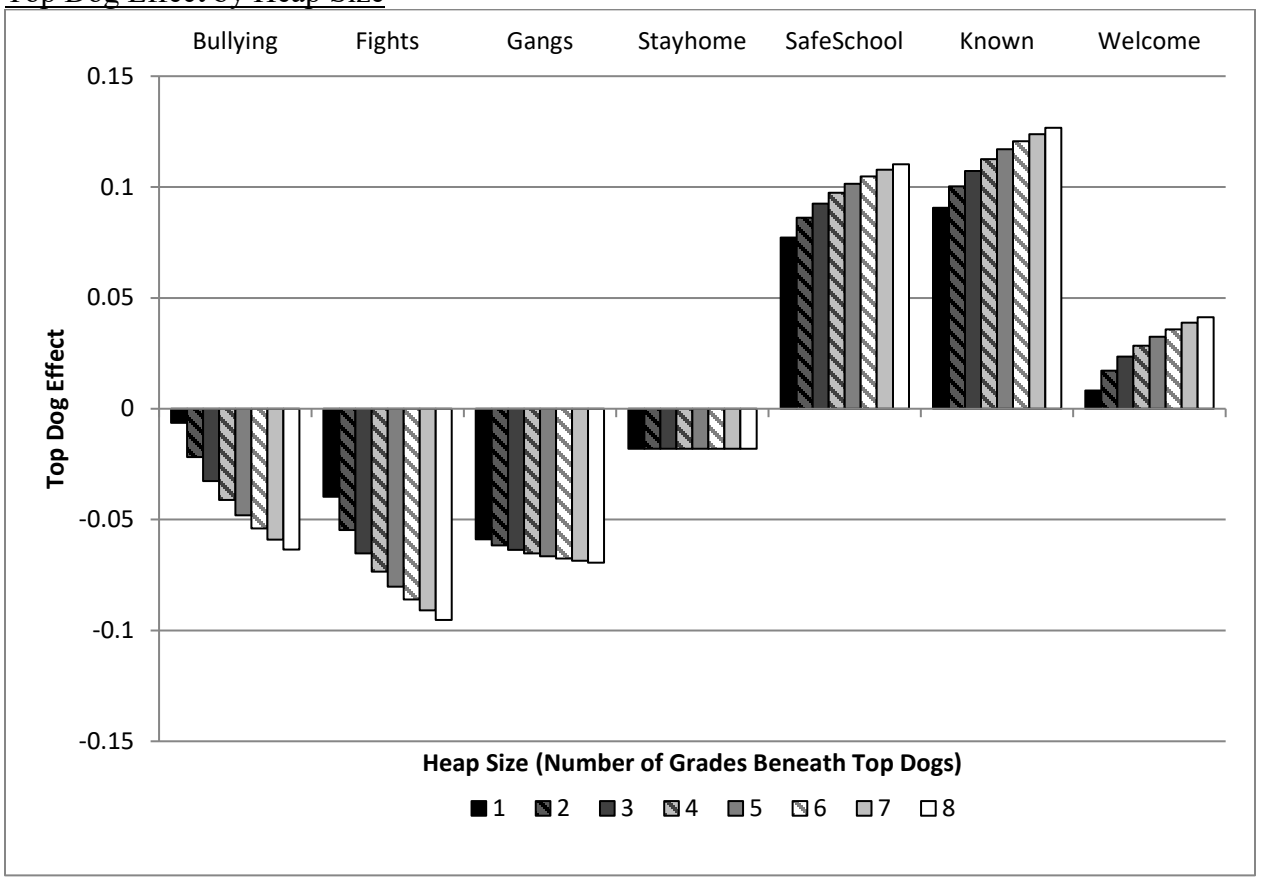


Percentage "Agree" or "Strongly Agree"



Sample includes 2008 6th graders who answer all seven survey questions included in this paper and who attend schools with grade spans enrolling at least 2% of the 6th grade student population. There are 65,852 students in 6-8 schools, 3,453 students in 6-12 schools, 4,301 students in 5-8 schools, 13,086 students in K-8 schools, and 7,079 students in K-6 schools.

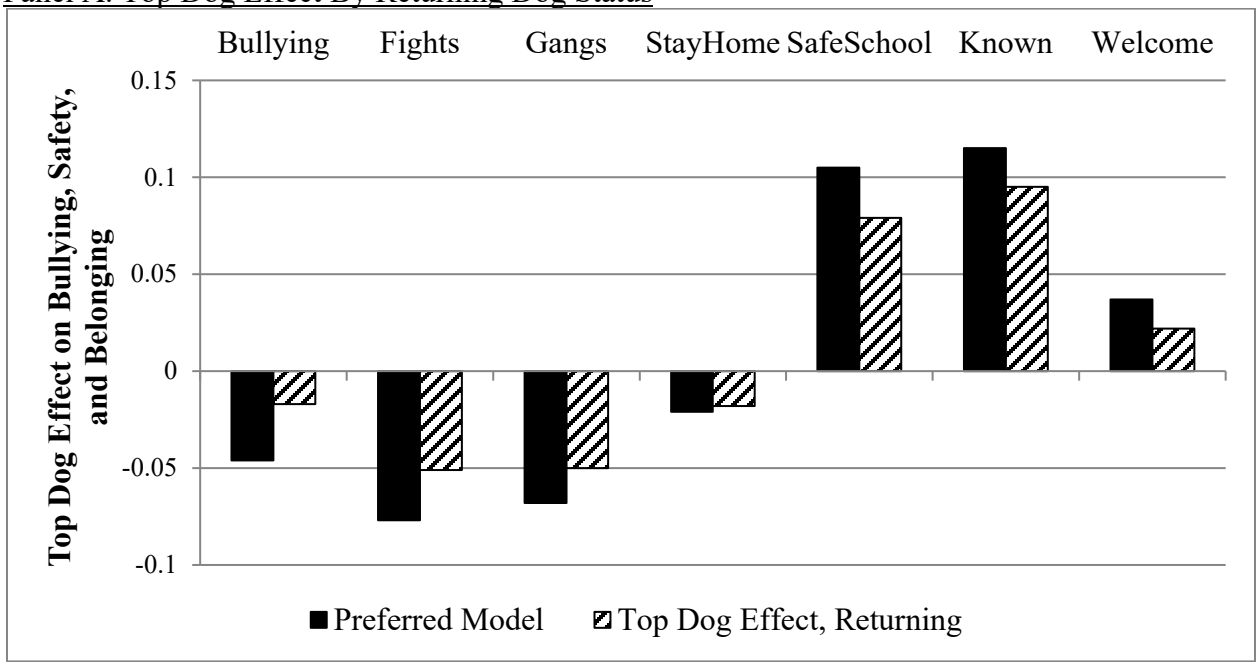
Figure 2. Regression Results by Heap Size (Number of Grades Underneath Top Dogs)



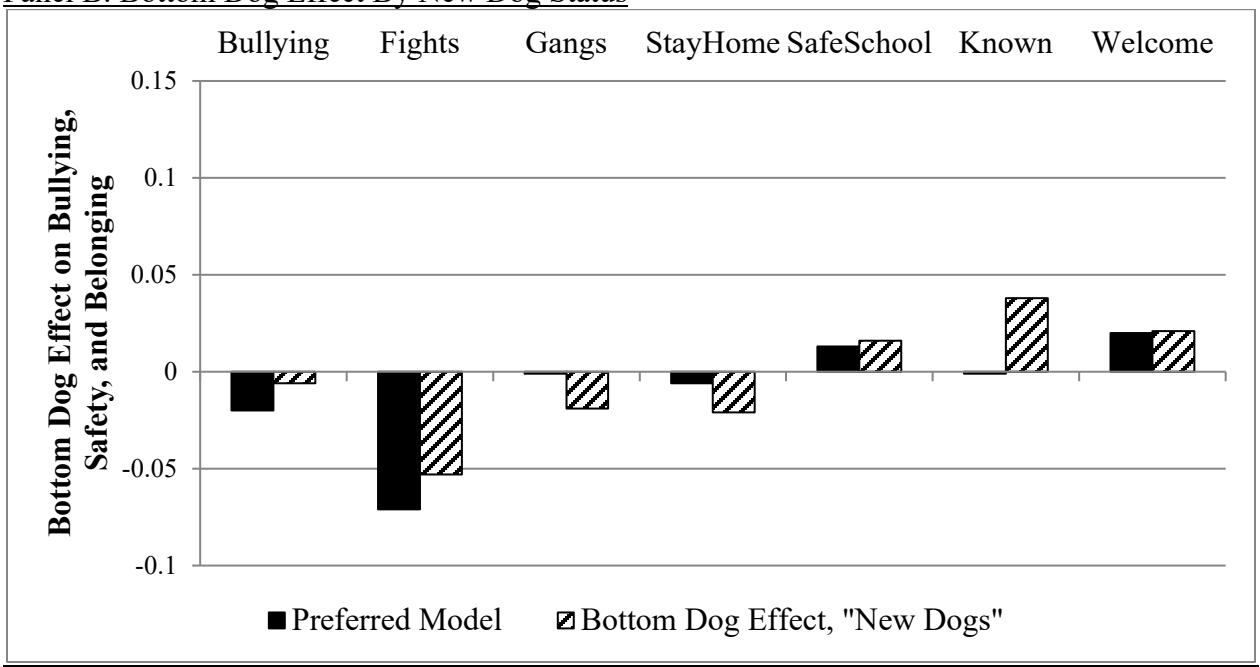
Robust standard errors in parentheses (***p<0.01, **p<0.05, *p<0.1) clustered by middle school. Heap is the natural logarithm of the number of grades underneath the top dogs in a school. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Controls in all models include indicators for free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade, student age, and student fixed effects. Reference Group = Middle Dogs. See Table S4 for coefficient estimates and standard errors.

Figure 3. Digging Deeper: Regression Results, Returning & New Dogs and TDBD Effect

Panel A. Top Dog Effect By Returning Dog Status



Panel B. Bottom Dog Effect By New Dog Status



New students are in their first year in a school. Returning students are those attending the same school in the previous year. None of the reported estimates are statistically different across models. IV estimates instrument for TDBD status using the grade span of student's 3rd grade school. Controls in all models include indicators for free lunch eligibility, reduced lunch eligibility, home language is English, limited English proficiency, part-time special education, borough of residence, grade, student age, and student fixed effects. Reference Group = Middle Dogs. Returning bottom dogs and new top dogs are exceedingly rare and results are not shown here. See Table S5 for coefficient estimates and standard errors.