Reexamining Middle School Effects: A Comparison of Middle Grades Students in Middle Schools and K–8 Schools

CHRISTOPHER C. WEISS and LINDSAY KIPNES
Columbia University

The period of the middle grades has seen numerous reforms to improve education for students in early adolescence. However, although several current reforms seek to overhaul middle schools, only a handful of studies have directly compared the effects of different configurations of grades. Our analysis uses district and student data from one of the few American urban districts that contain both middle schools and K–8 schools. We compare student outcomes in eighth grade, finding few differences by school type. Only self-esteem and perceived threat differ by type of eighth-grade school. We also show that students' self-esteem benefits academic outcomes, a benefit that primarily accrues to students in middle schools.

Introduction

The history of efforts in the United States to develop structures of schooling for the "middle grades"—the span from fifth grade through eighth grade—is one of continual tinkering and persistent dissatisfaction. A division of the 12 years of primary and secondary education into an eight-year elementary school and a four-year high school was the primary model for American public schools throughout the nineteenth century (Manning 2000). Beginning shortly after the emergence of ideas of adolescence as a distinct phase of the life course, one with specific needs, educators strove to adapt or create educational environments and institutions to meet these needs (e.g., Commission on the Reorganization of Secondary Education 1918). Over the past nine decades, schools for educating children in the middle grades have seen numerous revisions and alterations, conducted in an effort to create an educational environment that is suited to the particular academic, social, and emotional needs of students in an often difficult time of life.

FEBRUARY 2006   239

Reexamining Middle School Effects

However, despite numerous modifications, rarely has there been widespread or consistent satisfaction with the forms of middle grades schooling. Few efforts to develop an educational form better suited to middle grades education have satisfied reformers or achieved the desired results. Currently, although there is a diversity of forms for middle grades education, schooling in these grades is now predominantly conducted in middle schools, which typically encompass sixth, seventh, and eighth grades (U.S. Department of Education 2001). Yet despite their emergence as the modal form, few educators and researchers would argue that middle schools represent the solution to the shortcomings of other forms (e.g., Carnegie Council on Adolescent Development 1989; Clark and Clark 1993). Moreover, a number of studies have documented the difficulties students experience in middle school, such as poor grades, behavioral problems, and low self-esteem (e.g., Eccles et al. 1991; Simmons and Blyth 1987). In large part due to these findings and perceptions of middle schools' harmful effects, numerous districts across the United States have begun to eliminate their middle schools, changing their systems of education for middle grades students to other, usually smaller, schooling forms. In one of the largest such efforts to date, officials of the New York City public schools recently announced plans to abolish the majority of the city's 218 middle schools (Herszenhorn 2004).

These efforts to restructure the grades of primary and secondary education into the best schooling forms is part of what Mac Iver and Epstein have called "the longest-running debate in middle level educational research" (1993, 520).[1] Although such reforms purport to be taken in the interest of students, we argue that they are likely to have little effect on students' performance or well-being. Our argument is not that the years of middle school are not difficult for students; instead, we contend that the difficulties experienced in middle school are remarkably similar to those students encounter in other schooling

CHRISTOPHER C. WEISS, PhD, is director of the Quantitative Methods in the Social Sciences Program at Columbia University. His research focuses on how institutions shape children's and adolescents' development, how parenting strategies influence children's developmental trajectories, and structural influences on the transition to adulthood. His current research projects include a study of the effects of grade retention on children's educational trajectories and an investigation of how features of the urban environment influence children's physical activity levels and risk of obesity. LINDSAY KIPNES is currently chief financial officer at Dialectic Capital Management, LLC, in New York, NY. She conducted this research while a student in the Quantitative Methods of the Social Sciences Program at Columbia University, the program from which she recently received her Masters of Arts degree. While attending Columbia, her research focused on how configurations of grades in different schools affect students' outcomes.

forms. These difficulties have come to be seen as a problem with middle schools in large part because of research that has found middle school effects in the absence of comparisons with student outcomes in other types of schooling for the middle grades.

Perhaps what is most notable with respect to the literature on the effects of middle schools—or of any particular school form—is the very limited number of studies that directly compare middle schools with alternative schooling forms. Stated differently, although there are many arguments about the effects of particular schooling forms (such as the middle school) on students' outcomes, surprisingly little research has directly examined the effects of particular schooling forms in a comparative framework. In part, this may reflect a limitation of data and of how reforms are implemented. Because most districts have only one configuration of schooling forms (e.g., elementary school for grades K–5, middle school for grades 6–8, and high school for grades 9–12), comparisons are problematic, since it can be difficult to disentangle district-level differences from school-level differences.

In this analysis, we examine the relationship between student outcomes and schooling form in Philadelphia, a large urban district containing roughly equal numbers of middle schools and K–8 schools. We examine how eighth-grade students' performance in school varies by the form of school they attend, comparing how outcomes differ for students attending middle school and those attending K–8 schools. In short, we find that, although there are substantial differences in the sociodemographic characteristics of students in the two types of schools, there are few differences in student outcomes by the type of school attended. However, one domain in which middle schools differ significantly from other forms of school is self-esteem. We find not only that self-esteem varies significantly but also that students' self-esteem mediates the impact of difficulties experienced in eighth grade, although this protective effect exists only in middle schools.

Background

The History of Middle Grades Education in Separate Institutions

The period of the middle grades in the educational sequence has seen a number of educational reforms that have sought to better tailor instruction and improve student outcomes in these years. As John Lounsbury writes, the development of middle grades education is the "longest-running, most extensive educational reform movement in the United States" (1991, 68). Although it is difficult to establish precisely when and where the first junior high school was established in the United States, the idea for a separate institution devoted to education of

early adolescents emerged in the late nineteenth and early twentieth century (Clark and Clark 1993). The primary forces behind the development of the junior high school were concerns over soaring dropout rates and new ideas about adolescent development, particularly the notion that early adolescents had specific needs that required a unique schooling form (Angus et al. 1988). Reformers pushed for the development of the junior high school, a separate institution providing education for students in grades 7–9 (Cuban 1992). Expansion of the innovation was rapid, and by the middle part of the twentieth century, the junior high school had been accepted as an appropriate institution for the education of early adolescents (Clark and Clark 1993).

However, by the 1950s, there was growing dissatisfaction with junior highs, with increasing concern as to whether the junior high was appropriately responsive to the special needs of early adolescents (Hansen and Hearn 1971). Critics charged that the junior high was falling well short of its goal to effectively educate adolescents and that it merely duplicated high schools in its programs and policies. As Larry Cuban contends, junior high schools mirrored the high schools they were designed to supplant in "curriculum, instruction, organization, teacher attitudes toward subject matter, and extracurricular activities" (1992, 242). In response, beginning in the middle of the twentieth century, some educators began to advocate for the creation of a new school for the middle grades.

Advocates and scholars of the middle school movement of the 1960s sought not only to reorganize the grade sequence of schools but also to make schools more developmentally appropriate for students. Two of the leading middle school advocates, William Alexander and Emmett Williams (1965), called for organizational structures that would create schools-within-schools to foster social ties between teachers and students while utilizing the strengths of teachers with different specialties. Similarly, Eichorn (1966) specified a middle school model that would serve as a bridge to help students make the transition from the classrooms of elementary schools to the departments and class periods of high schools. Eichorn also emphasized a school in which instruction and curricula were tightly integrated (cited in Clark and Clark 1993). These middle schools, reformers argued, were structured both to address the developmental needs of early adolescents and to overcome the limitations of the junior high school.

In the early 1960s, middle schools developed in different pockets across the United States, particularly in the Northeast and the upper Midwest (Eichorn 1984). Educators of the middle school movement promoted practices they believed to be suited to the particular needs of students in the middle grades, such as flexible scheduling, ungraded programs, and team teaching (Alexander and George 1981). Fueled by efforts of educators, scholars, and organizations such as the National Middle School Association, the number of middle schools grew rapidly through the 1970s.

While quickly adopted, however, middle schools proved disappointing in

practice, with studies on the effects of middle schools showing that "the new middle schools and old high schools were surprisingly alike in actual practice" (Lounsbury 1991, 64). Reviews of research conducted during the first 20 years of middle schools found that a change in a school's grade span had little effect on educational practices in the middle grades (Calhoun 1983). Middle schools proved to possess many of the unfavorable characteristics of junior high schools, and the reform had fallen well short of its goals, such as integrating the departmentalization of subjects to promote students' critical thinking (Cuban 1992). As one review of middle school education phrased it, "changes were restricted largely to the names of schools and the grades they contained" (Mac Iver and Epstein 1992, 835).

However, up until very recently, the middle school movement has remained strong. In the past decade, the number of middle schools has increased by 41 percent, with a corresponding decline in the number of junior high schools (U.S. Department of Education 2001). Like the junior high school, whose wrongs it was intended to right, according to numerous researchers and educators, the middle school has not lived up to its potential. Based on these perceptions, several districts have recently initiated reforms to dismantle their middle schools and have sought to educate students in the middle grades through other schooling forms. In efforts to assess whether the middle school as an educational form is indeed failing its students, researchers from a variety of disciplines have examined a host of academic, psychological, and social outcomes. Almost all have concluded that middle schools are not good for early adolescents (e.g., Anderman and Maehr 1994; Eccles et al. 1991; see also Juvonen et al. 2004).

However, as we argue below, seldom have these conclusions been drawn from direct comparisons between middle schools and other schooling forms, a limitation of research design that has led to a distorted picture of the impact of middle schools. As a result, current initiatives to reform or eliminate middle schools are being undertaken with an inadequate understanding of the middle school's effects relative to those of alternative schooling forms.

Research on the Effects of Middle Schools

Although few studies have actually compared student outcomes in different forms of middle grades education, numerous scholars and practitioners have argued that middle schools influence students' behaviors and outcomes in negative ways. Perhaps the single largest domain of middle school effects research is self-esteem. Although the specific mechanisms responsible for the decline vary, a number of researchers have offered data to argue that middle schools are detrimental to students' self-esteem, especially for girls. A recent study of middle schools documented a significant decline in students' self-

esteem between sixth grade and eighth grade (Reddy et al. 2003). Some scholars, particularly Simmons and Blyth (1987), argue that the decline in self-esteem is due to the coincidence of the difficult transition to middle school with pubertal development. Research in this arena has suggested that the transition to middle school intensifies a developmental decline in adolescents' self-esteem, especially for females (Simmons et al. 1979).

Yet it must be stressed that, although the majority of these studies employ arguments that suggest a comparative framework, few actually compare students in middle schools to those in other forms. In one of the few studies to undertake such a comparison, Blyth et al. (1978) found that seventh-grade males in junior high school do not experience the growth in self-esteem of the K–8 seventh graders, nor do they experience the decline in self-esteem of the females in junior high school.

Self-esteem is not the only domain examined in this literature. Several studies have shown that the middle school is an alienating environment, one that negatively influences students' sense of school belonging or connectedness to their school (Eccles et al. 1991; Seidman et al. 1994). In another of the few studies that employ a direct comparison of different forms, Anderman (2002) shows that students who attended K–8 or K–12 schools in the middle grades reported a slightly greater sense of belonging as compared to students in middle schools. These feelings, in turn, are positively related to optimism and GPA and negatively related to depression, social rejection, and school problems (Anderman 2002). Similarly, Blyth et al. (1978) find that the K–8 school structure supports students' involvement with their peers and with extracurricular activities, while the junior high school dampens students' participation, despite the larger number of extracurricular activities offered.

Interpersonal relationships are also demonstrably less positive in middle schools than in other schooling forms, and some researchers have argued that the middle school provides a structure to facilitate negative behaviors such as cruelty or meanness among their students (Merten 1997). From a developmental perspective, the middle grades are generally a time of growing concern for popularity, with students placing increasing importance on interpersonal relationships. This shift in emphasis often results in increasingly nonconforming peer values, social competition, and mean behavior (Eder 1985; Seidman et al. 1994). The issue is more problematic in middle schools, some argue, because adults in the school do not have as much of an opportunity to know what goes on among students, as instruction is structured such that students move from classroom to classroom, limiting student-teacher interaction (Wilson and Herriot 1989). Similarly, students spend more time outside the classroom, which means that adult intervention in the social arena is scarce (Merten 1997).

Some researchers have extended this argument about the detrimental effects of limited contact between teachers and students to make a more general

point that the middle school structure does not match students' academic and intellectual needs. Research by Eccles and her colleagues (Eccles et al. 1991; Eccles and Midgely 1989) argues that specific features of the middle school environment are responsible for students' difficulties in middle school. Junior highs and middle schools feature curricula that are focused more on impersonal tasks and less on personal relationships between students and teachers, in contrast to elementary schools' greater focus on personal aspects of schooling (e.g., Midgley and Edelin 1998). Students' difficulties arise from the "mismatch" between their needs and the structure of the junior high environment.

One primary reason for this mismatch may be attributed to the classroom teacher. Most middle schools feature a school day that is segmented into several subject-specific periods, with teachers specializing in academic subject areas. As a result, middle school teachers are subject specific and see many students in one day, each for a short period of time. This structure does not allow for much individualized student attention or personal student-teacher relationships. As a consequence, motivational or academic problems are more likely to go unnoticed (Eccles et al. 1993). Such limited contact with teachers is not unnoticed by students. Seidman and his colleagues (1994) find that students perceive less support from their teachers and greater hassles in daily school life in junior high. Comparisons of middle school teachers and elementary school teachers show that middle school teachers typically grade more stringently than elementary school teachers (Eccles and Midgley 1989). The tougher grading standards and teacher expectations are related not only to students' grades but also to their academic self-perceptions (Murdock et al. 2000).

Finally, research has documented differences in school safety. Anderman found that students who attended schools with a K–8 or K–12 grade structure were less likely to report feeling victimized or that their school was unsafe, and they were less likely to get into trouble for bad behavior than adolescents in the traditional middle school grade structure (2002). Similarly, in a study using data from the National Education Longitudinal Survey, Anderman and Kimweli (1997) found that students in K–8 schools reported lower levels of crime victimization and held safer perceptions of their schools than students in the same grade attending middle schools. Additionally, Astor et al. (2001) found that sixth graders in middle schools were much more likely than sixth graders in elementary schools to perceive multiple and specific threats in their school environments. Blyth et al. identified similar differences, particularly along gender lines, arguing that the transition to junior high was one in which students moved from a relatively safe environment to a "hostile one in which males were much more likely to experience an act of victimization" (1978).

However, not all comparative research has found the middle school lacking relative to other forms of middle grade education, though any benefits found have not been consistently seen across studies. Mac Iver and Epstein (1993)

found that middle schools emphasize active learning techniques and other beneficial instructional approaches more than other schooling forms for the same grades. More recently, Weiss and Bearman (2004) found that transitions had little discernible impact for students moving from eighth grade to ninth grade (to a higher form of instituition—the high school), as compared to students who moved from eighth grade to ninth grade within the same institution. What few significant differences they did find were in a positive direction, suggesting that students who are socially or academically stigmatized in eighth grade are less likely to carry their stigma across a school transition.

In addition to these findings, other research sheds further light on the effects of middle schools, helping to contextualize the middle school effect. A study by Larson and Richards (1991) showed that, although boredom in school is typical for students in middle school, students in the same grades in other schooling forms exhibit identical levels of boredom. These findings suggest that, for some outcomes, the particular school form may be less important than factors such as stage of development. Moreover, there is evidence that some outcomes that may be affected by attendance at a middle school are transient in their effect. Kinney (1993), for example, shows that, for many students who have low social status in middle school, quality of social life improves significantly as they make the transition to high school.

In sum, although the research literature contains numerous papers on how particular forms of schooling influence students' outcomes and behaviors, research on the effects of schooling form generally has been limited to studies that contrast students in a particular form against a hypothetical group in other forms. The majority of studies that report on the effects of school form do not directly compare student outcomes in different school forms; rather, in many cases the counterfactual is assumed but not tested. Somewhat surprisingly, this weakness in the educational literature has persisted despite periodic calls for such research from scholars. Three decades ago, in the introduction to the *Report of the National Panel on High Schools and Adolescent Education*, J. H. Martin (1974) wrote, "Surprisingly, we found no research with significant findings to substantiate one organizational pattern over another . . . all [patterns] lack a validating research base" (cited in Blyth et al. 1978). Such a comparison is essential to understanding how school form influences student outcomes. In the analysis that follows, we examine how students' experience varies by school form, comparing students in Philadelphia in the middle grades of the education system.

Research Questions

For the most part, research on student outcomes and performance in the middle grades focuses on the academic performance of adolescents. However,

many of the same theoretical bases for this research indicate that we might expect similar differences in nonacademic outcomes. We expand the set of outcomes that are typically examined to incorporate several nonschool outcomes as well. Our investigation is organized around three related research questions shown below.

Question 1. Do outcomes in eighth grade vary based on the type of school form a student attends?   Although students in middle schools are presumed by many to have more negative outcomes in eighth grade relative to their peers attending K–8 schools, differences in student outcomes between school forms remain largely untested. Here we focus on a set of academic and nonacademic outcomes, examining, first, whether there is any variance in a set of outcomes. We expect to find, in this component of the analysis, that students attending middle schools have worse outcomes relative to those in K–8 schools.

Question 2. If we find differences in outcomes between school types, are these differences due to composition differences in their students?   Less clear, however, is whether these differences persist when examined in a multivariate framework. It may be that what appear as differences between schooling forms can be accounted for by compositional differences between the two types of schools. In the second stage of analysis, we examine whether the differences observed in bivariate comparison persist after controlling for a set of school-level and individual-level predictors.

Question 3. Do the effects of self-esteem on school-based outcomes vary by schooling form?   Since much of the literature argues that school form operates through the proximate mechanism of psychological well-being, particularly self-esteem, we examine the extent to which psychological outcomes vary by school type. We then extend these findings through models examining whether the effects of this form of well-being are different in different types of schools.

In pursuing this analysis, we use data from one of the few settings to have significant numbers of both K–8 schools and middle schools for middle grades education: the School District of Philadelphia. We seek to expand the existing findings of the effects of middle grades schooling by directly comparing student outcomes for those attending middle schools with outcomes for students in K–8 schools.

Data

To conduct these analyses, we use data from the Philadelphia Education Longitudinal Study (PELS), an ongoing longitudinal study of public high school students in Philadelphia. The study was designed to collect rich, detailed data on numerous dimensions of students' lives from eighth grade onward, beginning at the point shortly before they make the transition to high school,

and then monitor their experiences throughout the high school years. The study now has seven waves of data, supplemented with administrative records from various sources. For these analyses, we use only data collected in the first wave, which were collected through telephone interviews of parents and students in the summer before students began high school.

The sample for PELS was chosen through a stratified random sampling procedure using data from the School District of Philadelphia. From a list of all 93 schools attended by eighth graders in the district, a sample of 45 eighth-grade schools was chosen through random selection. These schools were then ranked according to the size of the eighth-grade class, and a random sample of students was drawn from each school using a data set containing district records for all eighth graders. For schools with 250 or fewer eighth graders, a random sample of approximately 26 percent of the eighth-grade class was drawn.[2] For larger schools, the sample size was 16 percent.

The first wave of the study began during the summer of 1996, when telephone interviews were conducted with 1,483 students and parents of those who had recently completed the eighth grade. This post-eighth-grade wave of the study had two primary goals: to obtain baseline information about parents' and students' educational attitudes, plans, and experiences and to investigate the factors that make the transition to ninth grade difficult for some students.

In addition to the information collected through the telephone interviews, the School District of Philadelphia has provided extensive data for the study. Data for each adolescent interviewed in PELS are linked with district information on grades, standardized test scores, attendance, and disciplinary problems. Each interview record is also linked with a district file containing information about which courses students took in eighth grade and ninth grade, what grades they received in these courses, how often they were absent, whether and what disciplinary problems they had, and what curricular track, if any, they were enrolled in during eighth grade.

Philadelphia as a Research Setting

The School District of Philadelphia is the nation's ninth largest, enrolling 204,851 students, which include those in early childhood programs (Sable and Young 2003). The racial composition of public school students in June 2001 was 65.3 percent African American, 4.9 percent Asian, 13.1 percent Hispanic, 0.2 percent Native American, and 16.4 percent white (School District of Philadelphia 2003b). The district also has one of the nation's highest percentages of Title I–eligible schools, with 98.9 percent of the district's schools eligible in the 2001–2 school year (Sable and Young 2003).

Philadelphia provides a unique setting in which to examine differences

between middle grades students in middle school and K–8 settings. The School District of Philadelphia has roughly equal numbers of middle schools and K–8 schools, with slightly more K–8 schools in the district than middle schools. Philadelphia represents a particularly interesting site in which to examine the effects of middle schools, because the district recently embarked upon an ambitious middle school reform the rationale of which closely resembles the research described in the previous section. As noted earlier, the issue of what effects middle schools have on their students has particular policy salience in Philadelphia. The district's Secondary Education Movement Strategic Plan, which commenced in 2003 and is expected to be completed in 2007, states that "the middle school organization has a history of failure in urban schools" and specifies a series of steps to reduce the number of middle schools in the district (School District of Philadelphia 2003a). Under the plan, the school district will convert nine middle schools into "smaller" high schools of from 800 to 1,000 students. Simultaneously, the elementary schools that send students to those middle schools will increase by one grade level each year in order to obtain a K–8 configuration (School District of Philadelphia 2003b).

In an effort to gain support for this transition, the reform document states several ways that K–8 schools are believed to serve adolescents better than the traditional middle school. For example, the district states that middle schools tend to be very large and to lack the "family atmosphere" of the K–8 schools that results from groups of students being together over a long period of time. Similarly, the K–8 organization enables the formation of close relationships between parents and teachers, a type of relationship that is less likely to develop in the middle school (School District of Philadelphia 2003b).

Table 1 presents the number of middle schools and K–8 schools in the School District of Philadelphia, the number of students served by each type, and some very basic information on the students of the two forms. These data are taken from the School District of Philadelphia and were current for the 1995–96 school year, the year that the first wave of PELS was deployed in the field.

As the figures in table 1 show, the number of middle schools and K–8 schools are roughly equal, with slightly more K–8 schools in the district than middle schools. However, because of their larger average size, the middle schools serve substantially more students than do the K–8 schools. The average K–8 school has 72 eighth-grade students, while the average middle school has nearly four times as many (268). As a result of the middle school's greater average size, nearly three-quarters (73 percent) of the district's eighth graders attend a middle school. The middle portion of table 1 documents differences in teacher characteristics between the two school types. Middle schools have a lower percentage of certified teachers than do K–8 schools. Middle school

Reexamining Middle School Effects

TABLE 1

Descriptive Characteristics of Philadelphia Eighth-Grade Schools

	Middle Schools	K–8 Schools
Number of schools	38	41
Total number of eighth-grade students	10,335	3,671
Median number of eighth-grade students per school	268	72
Teacher characteristics:		
Percent certified, 1999–2000	87.1	96.2
Retention rate (three-year), beginning 1999–2000	59.1	67.7
Average years of experience	11.4	14.3
Student characteristics:		
Median percent African American	90.7	41.9
Median percent receiving AFDC	60.1	42.4

SOURCES.—School data and student characteristics: School District of Philadelphia. Teacher data: Neild et al. (2003).

NOTE.—AFDC = Aid to Families with Dependent Children.

teachers are also less experienced and are more likely to leave their position within three years than their counterparts in K–8 schools.

The populations served by these types of school differ significantly as well. The lower part of table 1 shows data on characteristics of the student population in both types. While only 42 percent of eighth-grade students enrolled in K–8 schools are African American, over 90 percent of the students in middle schools are. Similarly, a greater percentage of middle school students come from families receiving public assistance. Taken together, the figures of table 1 show significant differences in student composition between middle schools and K–8 schools. The student population of the middle schools is much larger and poorer than that of the K–8 schools, and it has a substantially greater portion of African American students. Therefore, we control for these independent variables, as well as other potentially confounding observable differences, in our analysis.

The sociodemographic differences shown in table 1 also highlight the difficulties of comparing these school forms, since their populations are quite different. Given that middle schools were not randomly assigned to different neighborhoods in the city, any comparison of school form has problems of endogeneity. Since middle schools have greater proportions of students with characteristics associated with lower performance in school, it might be difficult to disentangle the middle school effects from differences due to the different compositions of both school types. As a practical matter, in these analyses the issue of endogeneity is particularly problematic since we find few differences in outcomes by schooling form. However, the matter is important to note as a difficulty in specifying effects of nonrandomly assigned treatments such as schooling form.

TABLE 2

Means and Standard Deviations for Variables Used

	Mean	SD
Grade average for eighth grade	77.86	8.19
Failed subjects for eighth-grade final grades (received an F = 1)	.325	.469
Excessive missed school (20 or more absences = 1)	.182	.386
Suspension (suspended = 1)	.434	.496
Self-esteem	.000	1
Safety	.000	1
Threat (threat = 1)	.401	.49
Like school	.000	1
African American	.601	.49
Held back	.33	.47
Female	.496	.5
Primary parent high school or less	.671	.471
Welfare	.128	.334
Number of eighth graders	1,013	584
Percent students African American	.598	.313

Measures

Predictors

Our analysis controls for various factors that might be related to the outcomes and therefore could be confounding variables. These variables are also included in an attempt to better focus on our main predictor of interest, middle school. We control for two school-level predictors in our models to account for the extent to which a middle school effect may be due to differences in school contexts: school size and racial composition.[3] Our measure of school size is the number of eighth-grade students in the school in the 1995–96 school year, while the racial composition measure is the percentage of the student population that is African American. Data for both of these school-level measures are taken from reports of the school district. Means and standard deviations for all variables used in this analysis are presented in table 2.

Among the individual-level predictors, we include a dichotomous variable related to students' race, African American, drawn from students' self-reports. This measure is equal to one if the student listed African American as a primary racial/ethnic identity.[4] Similarly, we include a dichotomous variable for student's gender. Rather than include a predictor for student's age, we use a measure of whether the student had been retained at least once during the

schooling career prior to the end of eighth grade. Because the sample was chosen based on eighth-grade attendance in the 1995–96 school year, age and previous retention are highly correlated, preventing inclusion of both measures in our models. Since the retention variable has been shown to be a more powerful predictor of academic and behavioral difficulties, we include it in these analyses. The measure we use is a dichotomous one, with data taken from the parent interview, and is equal to one if the student has been held back a grade during their schooling.

We also include two variables related to parents' socioeconomic status. We dichotomize parents' education level, with a variable equal to one if the parent has a high school degree or less and zero if the parent has more education than this. In addition, an income-related measure is included, one that equals one if the parent reported that the household received supplementary income from public assistance or welfare and/or food stamps. Finally, our primary predictor of interest is a dummy variable for middle school attendance. This measure is equal to one if the student attends/attended a middle school as opposed to a K–8 school.

Outcomes

We examine several academic and nonacademic outcomes, the data for which are taken from students' survey responses and their official records from the School District of Philadelphia. We examine two outcomes related to students' grades in school. The first is a measure of the students' average final grades for all courses, with the exception of their physical education courses. This variable ranges from 0 to 100. The second measure is dichotomous, equal to one if the student received an "F" as the final grade for any course and zero if he or she passed all courses. Two other measures are taken from school district records. The first, taken from students' attendance records, is equal to one if the student missed 20 percent or more of the school days in an academic year and zero otherwise. Students' disciplinary data were used to create a dichotomous measure equal to one if the student was suspended during the school year and zero if not.

We also use data from the survey to examine students' feelings about school and self-esteem. Two variables capture students' self-reported risk of physical harm in and around the school. Our measure of threat is a dichotomous one indicating whether or not the student had been threatened by another student at school. A second measure was created through factor analysis and captures how safe students feel in and around their school.[5] Students' feelings toward their school were created through factor analysis based on their opinions about school and school subjects. Finally, our variable for self-esteem was calculated

by factor analysis based on answers to questions measuring the student's general feeling of self-worth and satisfaction.

Analytical Approach

In our attempt to better discern any effects that attending a middle school has versus attending a kindergarten through eighth-grade school, we control for socioeconomic and other key characteristics that could affect the academic and nonacademic outcomes examined. Because PELS data are nested, with a group of students clustered within a group of schools, a multilevel analysis strategy is required (Bryk and Raudenbush 1992). Traditional linear regression models used on such data would result in biased standard errors. Thus, we also utilize multilevel regression of the various outcomes on the control factors that allows us to see if the type of school attended is significantly related. Moreover, use of this model controls for unobserved factors at the school level that may account for differences in students' outcomes.

We begin our analysis by comparing student characteristics and outcomes in eighth grade, using this bivariate comparison to get a sense of just how different these two schooling forms are in this district. We then estimate a series of multilevel models using MLWin to determine whether type of school attended has a significant effect on the outcome and how much of the variance in outcomes is explained by this factor. MLWin is a software package for fitting multilevel models; it is similar to HLM (hierarchical linear modeling). MLWin was developed by the Centre for Multilevel Modelling team at the Institute of Education, University of London. Please see Goldstein (1995) and Rabash et al. (2000) for further details.

Results

The first step in this analysis is to examine how student and family characteristics differ by school type, a comparison initiated in table 1. However, the figures presented in table 1 are for the district overall. How does this analysis sample compare? The distribution of student characteristics for the entire sample as well as the breakdown by type of school are shown in table 3. The school-type comparisons also include *t*-tests to assess whether there is a statistically significant difference in the characteristics of middle schoolers and those attending K–8 schools. As the data in table 3 document, the population attending middle schools has parents with lower education levels and is more likely to have received public assistance. Both of these differences are highly significant ($p < .001$). In addition, the percentage of students who have been

TABLE 3

Socioeconomic Characteristics of Students by School Type

	Middle School (%)	K–8 School (%)	Total (%)
Race:			
African American $=1$	58.8	61.8	60.2
Hispanic $= 1$***	17.2	9.0	13.5
Gender (female $=1$)	49.5	49.7	49.6
Parent education (high school or less $= 1$)***	71.3	61.9	67.1
Poverty (welfare recipient $= 1$)***	36.9	26.5	32.2
Student held back in school $= 1$*	35.4	30.2	33.0
Number of observations	755	620	1,375

* $p = .05$.
*** $p = .001$.

held back a grade is greater in middle schools than in K–8 schools ($p < .05$). Taken together, the figures of table 3 show significant differences in the characteristics of students who attend these two kinds of schools in Philadelphia.

In terms of academic outcomes, the distribution of values for the outcomes we examine by type of school attended is shown in table 4. There is a significant difference in grade average, with those in middle school having worse grades. In addition, the percentage of middle school students who failed a class in eighth grade is statistically significantly higher than that of K–8 students. Similarly, students in middle school were significantly more likely to have had poor attendance records than their K–8 counterparts. Some of the most striking differences between the two school types are those outcomes related to safety and threat. For both outcomes, adolescents in middle school feel significantly less safe and more threatened than those in K–8 schools. Finally, consistent with other research, students in middle school have lower levels of self-esteem than do those in K–8 schools.

Taken together, the results presented in tables 3 and 4 document significant differences between these two forms of eighth-grade school in the characteristics of students and in their well-being. In the next section of analysis, we examine whether these differences persist in a multivariate framework.

Multilevel Models: Outcomes from School Data

Results of our multilevel models are presented in tables 5 and 6. Table 5 shows estimates of the models predicting four school-related outcomes, while

TABLE 4

Comparison of Mean Outcomes by School Type

	Middle School (%)	K–8 School (%)
Grade average for eighth grade***	77.2	78.7
Failed subjects for eighth-grade final grades (received an F = 1)***	37.5	26.5
Excessive missed school (20 or more absences = 1)*	20.0	16.0
Suspension (suspended = 1)	45.4	40.9
Self-esteem**	−.077	.094
Safety**	−.076	.092
Likes school	−.013	.016
Threat (threat = 1)***	44.2	35.2

* $p = .05$.
** $p = .01$.
*** $p = .001$.

table 6 presents results for nonacademic outcomes. For all models of dichot-omous outcomes, we use second-order penalized quasi-likelihood (PQL) es-timation procedures to minimize downward bias in between-group variances (Guo and Zhao 2000; Rodríguez and Goldman 1995). For each outcome examined, three models are shown: the first displays results of a model esti-mated with only the middle school variable included as a predictor, the second adds the two school-level predictors, and the third adds the set of individual-level sociodemographic predictors.[6]

Looking across the different models in table 5, the most striking feature of the table is that middle school attendance is unrelated to all four outcomes in any of the models examined. Middle school attendance is not significant when it is the only predictor in the model, nor is it significant when school- and individual-level variables are controlled. Neither grade average nor the likelihood of failing at least one course is significantly different for those who attend middle schools relative to those attending K–8 schools. In the first column of the table, the results reveal that those who attend a middle school have a slightly lower grade average than those attending a K–8, although the difference is far from significant. When the school-level predictors are added in the second column, the coefficient for the middle school variable is reduced. Both school-level predictors are significant, as discussed below. When the individual predictors are added, as seen in the third data column, the effect is further reduced. The results of the models for predicting whether a student had failed a course are similar.

TABLE 5

Hierarchical Linear Wave 1 Models: Predicting Academic Outcomes

												ACADEMIC OUTCOMES						
	Grade Average			Received an F[a]			Many Absences[b]			Suspension[c]		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Middle school	−.470 (.679)	−.229 (.640)	−.274 (.594)	.281 (.216)	.218 (.214)	.212 (.222)	.192 (.223)	.174 (.214)	.117 (.191)	.107 (.176)	.072 (.161)	.053 (.160)
Number of eighth graders		−.001** (.000)	−.001** (.000)		.000** (.000)	.000** (.000)		.000** (.000)	.000** (.000)		.000** (.000)	.000** (.000)
Percent students African American		−5.179** (1.001)	−3.293** (1.009)		1.139** (.338)	.898** (.385)		1.300** (.347)	1.537** (.417)		.917** (.258)	.520 (.297)
Black			−1.526** (.538)			.146 (.211)			−.663 (.261)			.407*** (.177)

256

Held back	−3.979***	.827***	.766***	.455***
	(.440)	(.158)	(.188)	(.145)
Female	3.739***	−.975***	−.010	−.841***
	(.392)	(.153)	(.185)	(.133)
Primary parent high school or less	−.638	.266	.270	.180
	(.439)	(.173)	(.218)	(.147)
Welfare	−1.290**	.339*	.622***	.317*
	(.455)	(.166)	(.193)	(.148)

[a] Dichotomous variable reflecting whether the student received an F on the eighth-grade final report card or not.
[b] Dichotomous variable where many absences is defined as being absent more than 20% of the year.
[c] Dichotomous variable reflecting whether or not the student had been suspended.

* $p = .05$.
** $p = .01$.
*** $p = .001$.

257

TABLE 6

Hierarchical Linear Wave 1 Models: Predicting Nonacademic Outcomes

NONACADEMIC OUTCOMES

	Self-Esteem[a]			Safety[a]			Threat[b]			Like School[b]		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Middle school	−.167* (.069)	−.170** (.065)	−.165* (.066)	−.143 (.090)	−.116 (.086)	−.114 (.087)	.430*** (.128)	.439*** (.128)	.453*** (.130)	−.093 (.069)	−.079 (.065)	−.088 (.067)
Number of eighth graders		.000 (.000)	.000 (.000)		.000 (.000)	.000 (.000)		.000 (.000)	.000 (.000)		.000 (.000)	.000 (.000)
Percent students African American		.184 (.105)	.071 (.127)		−.490*** (.135)	−.448** (.148)		−.045 (.208)	.137 (.265)		.103 (.105)	.082 (.128)
Black			.154* (.075)			−.028 (.079)			−.146 (.166)			.007 (.076)

258

Held back	−.255*** (.063)	−.091 (.064)	.259 (.141)
			−.141* (.064)
Female	−.134* (.057)	−.039 (.057)	−.232 (.127)
			.112* (.057)
Primary parent high school or less	.075 (.063)	−.055 (.064)	−.076 (.140)
			.019 (.063)
Welfare	.025 (.065)	.052 (.067)	−.247 (.145)
			.074 (.065)

[a] Variable calculated by factor analysis. Survey questions used to compute the variable are presented in the appendix.
[b] Dichotomous variable reflecting whether or not the student has been threatened by another student at school.
* $p = .05$.
** $p = .01$.
*** $p = .001$.

259

Reexamining Middle School Effects

In contrast to the lack of a significant relationship with middle school attendance, the two contextual variables are significantly related to both grade average and the odds of receiving an F in a course. School size is negatively related to grade average and positively related to the odds of course failure. Likewise, the greater the percentage of students who are African American, the lower the grade average and the higher the odds of failing at least one course.

Among the individual-level factors, students' race, sex, previous grade retention, and family welfare status all exert significant influences. African American students have average grades that are more than two points (out of 100) lower than their those of their white classmates; however, there is no significant difference in the odds of failing a class by race. Females have significantly better grades and are much less likely to fail at least one course than their male counterparts. Students who have been held back at least one year in school and those whose families receive welfare have significantly lower grades and greater odds of failing a course than do their counterparts.

A similar pattern is evident in the two school-behavior outcomes we examine in table 5. Attendance at a middle school is not significantly related to having a large number of absences or the likelihood of being suspended during the school year. This absence of a significant relationship is present in the model with middle school predictor only and in the other two models as well. For absences and suspensions, both contextual variables are significant. Students who attend larger schools and schools with a higher percentage of African American students are significantly more likely to have a large number of absences and to be suspended from school.

Of the other predictors included in the full model, having been retained a grade and family welfare status are both positively related to having a high number of absences. Both of these factors are significantly and positively related to the likelihood of being suspended during eighth grade. Moreover, race and gender are also related to the odds of suspension, with African Americans having greater odds and females having lower odds relative to white and male students, respectively. Taken together, the results of the models examining these four academic outcomes show no discernible effect of middle school attendance.

Multilevel Models: Outcomes from Self-Reported Data

Table 6 presents results of four models examining nonacademic outcomes. Looking at the values of the predictors for middle school across the table, there are two outcomes in which middle school is a significant predictor in the expected direction. In the first, the model predicting students' self-esteem,

the results show that students who attend a middle school have significantly lower levels of self-esteem than their counterparts at K–8 schools. In this model, race, gender, and previous retention status are also significantly related.

The other outcome for which middle school attendance is a significant predictor is students' perception of threat in the school environment. Students who attend a middle school perceive their school environment as significantly more threatening than do those who attend a K–8 school. The control variables are also significant predictors of perceived threat, with females reporting lower levels of threat than males and children whose families live in poverty experiencing less threat than those who are better off. The other two outcomes examined—whether students like their school and their perceived level of safety at school—are both unrelated to middle school attendance.

Taken in sum, the results of tables 5 and 6 reveal far fewer differences in student outcomes by school type than previous research would suggest. None of the school-based outcomes show any significant differences by school type. Only perceptions of threat and self-esteem significantly vary by school type. These results call into question whether districts' efforts to eliminate middle schools are well founded. We also find that the characteristics of schools—number of students and the percentage of students who are African American—are significantly related to most of these outcomes. In models not shown, we examine whether these variables might be suppressing a middle school effect. The middle school coefficient remained insignificant in these models. In the final section of this analysis, we look at whether and to what extent self-esteem varies in its effect by type of school that the student attends.

Examining the Effect of Self-Esteem on Other Outcomes

In this final section of the analysis, we examine whether the relationship between self-esteem and student outcomes varies by school type. It may be that self-esteem does not have a uniformly positive (or negative) relationship with performance in school; rather, its effect might vary by type of school. To examine whether this is the case, we estimate two sets of models: in the first, we test whether the effect of self-esteem varies by school type through a series of interaction models, while in the second, we run separate regression models by school type.

Table 7 presents the results of the interaction models. For each outcome, two models are estimated. In the first, the variable for self-esteem is added to the full models of table 5. Then, in the second model, an interaction term is added to test whether effects of self-esteem vary by type of school attended.

Looking across models, the coefficient for self-esteem is significantly related to all four of the school-based outcomes. Students with higher self-esteem have

TABLE 7

Hierarchical Linear Wave 1 Models with Self-Esteem–Middle School Interaction

| | ACADEMIC OUTCOMES | | | | | | | |
| | Grade Average | | Received an F[a] | | Many Absences[b] | | Suspension[c] | |
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
Middle school	−.047	−.061	.151	.138	.095	.093	.028	.029
	(.591)	(.590)	(.224)	(.225)	(.191)	(.194)	(.159)	(.160)
Number of eighth graders	−.001*	−.001*	.000*	.000*	.000	.000	.000	.000
	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)	(.000)
Percent students African American	−3.270**	−3.285**	.917*	.930*	1.565***	1.587***	.533	.534
	(.999)	(.999)	(.389)	(.389)	(.419)	(.422)	(.296)	(0.296)
Black	−1.707**	−1.719**	.203	.209	−.644*	−.646*	.435*	.436*
	(.532)	(.532)	(.214)	(.214)	(.263)	(.263)	(.178)	(.178)
Held back	−3.688***	−3.693***	.750***	.756***	.725***	.729***	.417**	.417**
	(.437)	(.438)	(.160)	(.160)	(.190)	(.190)	(.146)	(.146)

262

Female	3.888***	3.882***	−1.040***	−1.041***	−.032	−.032	−.866***	−.865***
	(.387)	(.388)	(.156)	(.156)	(.186)	(.186)	(.134)	(.134)
Primary parent high school or less	−.713	−.718	.284	.290	.288	.289	.192	.192
	(.433)	(.433)	(.174)	(.175)	(.219)	(.220)	(.148)	(.148)
Welfare	−1.310**	−1.305**	.364*	.362*	.625**	.622**	.327*	.327*
	(.449)	(.449)	(.168)	(.168)	(.193)	(.194)	(.148)	(.148)
Self-esteem	1.144***	.906**	−.341***	−.174	−.163	.013	−.157*	−.145
	(.210)	(.318)	(.083)	(.132)	(.097)	(.162)	(.072)	(.109)
Self-esteem × middle school		.417		−.274		−.276		−.021
		(.420)		(.169)		(.202)		(.114)

[a] Dichotomous variable reflecting whether the student received an F on the eighth-grade final report card or not.
[b] Dichotomous variable where many absences is defined as being absent more than 20% of the year.
[c] Dichotomous variable reflecting whether or not the student had been suspended.
* $p = .05$.
** $p = .01$.
*** $p = .001$.

263

Reexamining Middle School Effects

better grades, are less likely to fail a course, are less likely to be absent a great deal, and have lower odds of being suspended in eighth grade. However, there appear to be few significant differences in the effect of self-esteem between the two schooling forms, as attested by the fact that in none of the models is the interaction term significant. Students who have higher self-esteem have significantly better grades, as shown in the coefficient for the main effect of self-esteem; however, the benefit of self-esteem is not significantly different for students in middle schools as compared to those in K–8 schools. Two of the other three outcomes also show benefits of self-esteem, but this does not vary by school form. Students with higher self-esteem are less likely to receive an F as a final course grade on their report card; however, this effect does not vary by school type. As with the model estimates shown in tables 5 and 6, the coefficient for middle school is not significant in any of the models of table 7.

Finally, we estimated separate models for the two schooling forms. Although the interaction models provide the best test of whether the effects of self-esteem vary significantly between K–8 and middle schools, we believe it useful to present these separate models as well. The results are presented in table 8.

For three of the four outcomes examined here, the benefit of self-esteem is greater for students in middle schools than for those attending K–8 schools. In the first two columns of table 8, we see that the coefficient for self-esteem is almost twice as large for middle school students as for those in K–8 schools. Similarly, the benefit of self-esteem on the odds of failing a class is greater for those in middle schools. Again, these results are presented largely for descriptive purposes, although they do help to contextualize the kinds of factors that shape eighth-grade students' performance in different types of school, and, too, they point to a potentially important direction for future research.

Conclusions

This investigation of the effects of attendance at a middle school yields several unanticipated findings. Perhaps most surprising is that our models showed no middle school effect at all for six of the eight outcomes we examined in our analysis. In contrast to previous research findings and widely held beliefs about the effects of middle schools, our findings offer little support for reformers seeking to improve students' performance in the middle grades by eliminating middle schools. Across a variety of outcomes, our results show that there is little difference in student performance based on the type of school that they attend. This is not to say that middle schools are not difficult environments for eighth graders; however, as our results make clear, the environment of the middle school is no more detrimental to students' performance than that of

264   American Journal of Education

TABLE 8

Hierarchical Linear Wave 1 Models with Self-Esteem: Separate Models by School Form

	SCHOOL-BASED OUTCOMES							
	Grade Average		Received an F[a]		Many Absences[b]		Suspension[c]	
	Middle School	K–8 School	Middle School	K–8 School	Middle School	K–8 School	Middle School	K–8 School
Black	−1.601*	−1.476*	.391	−.004	.102	−.759**	.899***	.374*
	(.682)	(.602)	(.235)	(.246)	(.209)	(.274)	(.186)	(.191)
Held back	−4.621***	−3.863***	.921***	.742***	.748***	.875***	.439**	.329
	(.588)	(.616)	(.180)	(.217)	(.194)	(.239)	(.172)	(.188)
Female	3.708***	4.328***	−.918***	−1.175***	−.014	−.075	−1.138***	−.740***
	(.542)	(.557)	(.178)	(.218)	(.193)	(.238)	(.166)	(.173)
Primary parent high school or less	−.551	−.335	.252	.103	.206	.293	.488**	−.037
	(.625)	(.598)	(.204)	(.232)	(.237)	(.275)	(.187)	(.188)
Welfare	−.782	−2.366***	.244	.598*	.586**	.793**	.400*	.194
	(.603)	(.657)	(.190)	(.237)	(.202)	(.256)	(.178)	(.205)
Self-esteem	1.619***	.883**	−.449***	−.235*	.222	−.259	−.242**	−.098
	(.279)	(.665)	(.093)	(.296)	(.122)	(.200)	(.086)	(.086)

[a] Dichotomous variable reflecting whether the student received an F on the eighth-grade final report card or not.
[b] Dichotomous variable where many absences is defined as being absent more than 20% of the year.
[c] Dichotomous variable reflecting whether or not the student had been suspended.

* $p = .05$.
** $p = .01$.
*** $p = .001$.

the K–8. Although much has been made about the negative consequences that middle schools have on students' performance in the middle grades, we find little effect.

Generally, for the outcomes we examine in this analysis, we find little evidence that the form of the school that a student attends in eighth grade is of great concern. More interesting is the fact that we do not find differences where a body of previous research indicates we should: in grades, likelihood of failure, and so forth. Where we do find significant differences, they are in the realms of self-esteem and threat. The statistical significance of self-esteem in our results replicates the findings of previous research. Previous research has attributed this decline to the discontinuity of the transition and the resulting increase in students' sense of anonymity (Simmons et al. 1979).

We are unclear how to attribute the differences between our results for outcomes other than self-esteem and those of previous research in this area. Our reading of the literature finds that previous studies have suggested comparative benefits from other schooling forms but have seldom examined whether such differences hold. The absence of such comparisons is largely due to the implementation strategies of middle school reform, through which a district will adopt a single district-wide schooling form. There are very few districts like Philadelphia, in which middle school reform left some of the previous schools intact. Having multiple schooling forms in a single district provides a distinct comparative advantage for understanding how middle schools shape students' outcomes—and how different school forms for eighth grade yield remarkably similar effects on students' well-being. It is outside the scope of our analysis to examine the numerous ways that self-esteem has been operationalized in other research; however, it bears mentioning that self-esteem is a complex concept, one that is difficult to measure. Self-esteem also varies over time and may be influenced by a student's academic performance. Disentangling the temporal ordering and conceptual robustness of our findings on self-esteem is an important direction for future research.

One might argue that the results of this analysis represent either exceptionally high-performing middle schools and/or exceptionally low-performing K–8 schools. Research on middle schools has documented substantial variation in the specific features of the middle school model as they are actually incorporated in middle schools, suggesting that our focus on grade span may miss instructional elements that could affect student outcomes (e.g., Lipsitz et al., 1997; Mac Iver and Epstein 1991). Although we lack data to examine directly the operation or environment of the schools in our sample, evidence from other sources suggests that these schools are not atypical. Indeed, it is difficult to imagine that the district could undertake a reform to change middle schools to K–8 schools if it were indeed the case that its middle schools were exceptionally high performing or its K–8 schools exceptionally low perform-

ing. However, with our data, we are unable to do more than note that it is highly unlikely that the excellence of the middle schools or the poor quality of K–8 schools account for our results.

The significance of the two school-level measures in several of our models does potentially point to an advantage of the current policy reforms to eliminate middle schools. The finding that larger school size is associated with worse grades, higher odds of failure, and other outcomes is of potential importance and is consistent with other research on school size. However, our models indicate that such policies may not be adequately directed, since they are not focused specifically on school size. Because we are unable to examine procedures and instructional practices in these schools, we can only point to these findings as an important direction for future research.

One limitation of our data should be noted. Since outcome and predictor data were collected at the same time, it is impossible to precisely determine whether the effects operate in the direction that we argue. It could be that poorer grades result in lower self-esteem for this sample of students. For eighth grade, we do not have the panel data required to examine this possibility; however, in future research, we plan to look at whether and to what extent self-esteem helps or hinders outcomes as students move between eighth grade and ninth grade.

Our analysis is intended to inform efforts to improve education in the middle grades, such as the large-scale overhauls of middle schools currently under way in Philadelphia, New York, and other districts across the country. There would seem to be little question that the period of the middle grades is not an easy time, with many difficulties emerging at this stage of the developmental life course while other problems intensify. Our findings, unfortunately, cannot speak to the question of what interventions are most likely to improve the educational outcomes or general well-being of students in the middle grades. We know nothing of the particular curricular materials, instructional strategies, or grouping patterns used in these schools.

However, our findings can speak to the value of the current reforms to eliminate middle schools. Our results suggest that such reform efforts are likely to have little effect on students' performance or well-being in the middle grades. Such reforms are largely based on understandings of the differences in student outcomes between middle schools and other schooling forms. While they cannot address how eighth-grade performance can be enhanced, what these results clearly indicate is that efforts to improve these outcomes through changes in schooling form are unlikely to succeed.

Reexamining Middle School Effects

Appendix

Description of the Construction of Variables for All Measures

Self-Esteem

Standardized factor-weighted composite ($M = 0$, SD $= 1$) of four items. Composite created using principal components factor analysis. Answer categories are strongly agree (1), agree (2), disagree (3), and strongly disagree (4). For these measures, Cronbach's alpha $= .68$ and the factor eigenvalue $= 2.07$. Scores on the measure range from -4.49 to 1.25.

- You feel that you are very good at your school work.
- You have a lot of friends.
- You are happy with yourself most of the time.
- You like the kind of person you are.

Safety

Standardized factor-weighted composite ($M = 0$, SD $= 1$) of four items. Composite created using principal components factor analysis. Answer categories are not safe (1), somewhat safe (2), mostly safe (3), and very safe (4). For these measures, Cronbach's alpha $= .77$ and the factor eigenvalue $= 2.37$. Scores on the measure range from -2.65 to 1.19.

- In your eighth-grade school, how safe did you feel in your classes?
- In your eighth-grade school, how safe did you feel in the hallways and bathrooms?
- In your eighth-grade school, how safe did you feel outside around the school?
- In your eighth-grade school, how safe did you feel traveling between home and school?

Like School

Standardized factor-weighted composite ($M = 0$, SD $= 1$) of three items. Composite created using principal components factor analysis. Answer categories strongly agree (1), agree (2), disagree (3), and strongly disagree (4). For these measures, Cronbach's alpha $= .77$ and the factor eigenvalue $= 2.37$. Scores on the measure range from -2.65 to 1.19.

- In your eighth-grade year, you learned a lot in school.
- The topics you studied were usually interesting.
- You usually looked forward to school.

Notes

1. See also Jenkins and McEwin (1992).

2. This rule held except in the case of some very small schools. In order to ensure a sufficient number of cases from each school to detect school effects, a minimum of 12 cases was drawn from each school.

3. We would like to have modeled the effects of the measures of teacher characteristics shown in table 1, but we did not have those data for each school in our sample. Therefore, rather than reduce the number of cases, we decided not to use these variables.

4. The question used in PELS asks students to choose all racial identities that apply, a format similar to the question used in the 2000 U.S. Census.

5. More detail about the items used to create scores through factor analysis and their factors loadings is contained in the appendix.

6. We have not included figures about the explanatory power of the models, since the most commonly used statistical measures of goodness of fit (Wald test and intraclass correlation coefficient) have had their appropriateness for binary outcomes questioned in the statistical and public health literature. For further details see Bingenheimer and Raudenbush (2004), Bryk and Raudenbush (1992), and Snijders and Bosker (1994).

References

Alexander, William M., and P. S. George. 1981. *The Exemplary Middle School*. New York: Holt, Rinehart & Winston.

Alexander, William M., and Emmett L. Williams. 1965. "Schools for the Middle Years." *Educational Leadership* 23 (3): 217–23.

Anderman, Eric M. 2002. "School Effects on Psychological Outcomes during Adolescence." *Journal of Educational Psychology* 94 (4): 795–809.

Anderman, Eric M., and David M. S. Kimweli. 1997. "Victimization and Safety in Schools Serving Early Adolescents." *Journal of Early Adolescence* 17 (4): 408–38.

Anderman, Eric M., and Martin L. Maehr. 1994. "Motivation and Schooling in the Middle Grades." *Review of Educational Research* 64 (2): 287–309.

Angus, David L., Jeffery E. Mirel, and Maris A. Vinovskis. 1988. "Historical Development of Age-Stratification in Schooling." *Teachers College Record* 90 (2): 211–36.

Astor, Ronald A., Heather A. Meyer, and Ronald O. Pitner. 2001. "Elementary and Middle School Students' Perceptions of Violence-Prone School Subcontexts." *Elementary School Journal* 101 (5): 511–28.

Bingenheimer, Jeffrey B., and Stephen W. Raudenbush. 2004. "Statistical and Substantive Inferences in Public Health: Issues in the Application of Multilevel Models." *Annual Review of Public Health* 25:53–77.

Blyth, Dale A., Roberta G. Simmons, and Diane Bush. 1978. "The Transition into

Early Adolescence: A Longitudinal Comparison of Youth in Two Educational Contexts." *Sociology of Education* 51 (3):149–62.

Bryk, Anthony S., and Stephen W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage.

Calhoun, Fredrick S. 1983. *Organization of the Middle Grades: A Summary of Research*. Arlington, VA: Educational Research Service.

Carnegie Council on Adolescent Development. 1989. *Turning Points: Preparing American Youth for the Twenty-First Century*. New York: Carnegie Corporation.

Clark, Sally N., and Donald C. Clark. 1993. "Middle Level School Reform: The Rhetoric and the Reality." *Elementary School Journal* 93 (5): 447–60.

Commission on the Reorganization of Secondary Education. 1918. *Cardinal Principles of Education*. Washington, DC: U.S. Bureau of Education.

Cuban, Larry. 1992. "What Happens to Reforms That Last? The Case of the Junior High School." *American Educational Research Journal* 29 (2): 227–51.

Eccles, Jacquelynne, Sarah Lord, and Carol Midgely. 1991. "What Are We Doing to Early Adolescents? The Impact of Educational Contexts on Early Adolescents." *American Journal of Education* 89 (4): 521–42.

Eccles, Jacquelynne, and Carol Midgely. 1989. "Stage/Environment Fit: Developmentally Appropriate Classrooms for Early Adolescents." In *Research on Motivation in Education*, vol. 3, ed. R. E. Ames and C. Ames. New York: Academic Press.

Eccles, Jacquelynne S., Allan Wigfield, Carol Midgley, David Reuman, Douglas Mac Iver, and Harriet Feldlaufer. 1993. "Negative Effects of Traditional Middle Schools on Students' Motivation." *Elementary School Journal* 93 (5): 553–74.

Eder, Donna. 1985. "The Cycle of Popularity: Interpersonal Relations among Female Adolescents." *Sociology of Education* 58 (3): 154–65.

Eichorn, D. H. 1966. *The Middle School*. New York: Center for Applied Research in Education.

Eichorn, D. H. 1984. "The Nature of Transescents." In *Perspectives: Middle School Education*, ed. J. H. Lounsbury. Columbus, OH: National Middle School Association.

Goldstein, Harvey. 1995. *Multilevel Statistical Models*. New York: Wiley.

Guo, Guang, and Hongxin Zhao. 2000. "Multilevel Modeling for Binary Data." *Annual Review of Sociology* 26:441–62.

Hansen, John H., and Arthur C. Hearn. 1971. *The Middle School Program*. Chicago: Rand McNally.

Herszenhorn, David M. 2004. "City Plans to Eliminate Most Middle Schools." *New York Times*, March 4.

Jenkins, D. M., and C. K. McEwin. 1992. "Which School for the Fifth Grade? Programs and Practices in Three Grade Organizations." *Middle School Journal* 23 (4): 8–12.

Juvonen, Jaana, Vi-Nhuan Le, Tessa Kaganoff, Catherine Augustine, and Louay Constant. 2004. *Focus on the Wonder Years: Challenges Facing the American Middle School*. Santa Monica, CA: RAND.

Kinney, David A. 1993. "From Nerds to Normals: The Recovery of Identity among Adolescents from Middle to High School." *Sociology of Education* 66 (1): 21–40.

Larson, Reed W., and Maryse H. Richards. 1991. "Boredom in the Middle School Years: Blaming Schools versus Blaming Students." *American Journal of Education* 99 (4): 418–43.

Lipsitz, Joan, Anthony W. Jackson, and Leah Meyer Austin. 1997. "What Works in Middle Grades Reform?" *Phi Delta Kappan* 78 (7): 517–19.

Lounsbury, John H. 1991. *As I See It*. Columbus, OH: National Middle School Association.

Mac Iver, Douglas J., and Joyce L. Epstein. 1991. "Responsive Practices in the Middle Grades: Teacher Teams, Advisory Groups, Remedial Instruction, and School Transition Programs." *American Journal of Education* 99 (4): 587–622.

Mac Iver, Douglas J., and Joyce L. Epstein. 1992. "Middle Grades Education." In *Encyclopedia of Educational Research*, 6th ed., ed. Marvin C. Alkin. New York: Macmillan.

Mac Iver, Douglas J., and Joyce L. Epstein. 1993. "Middle Grades Research: Not Yet Mature, but No Longer a Child." *Elementary School Journal* 93 (5): 519–33.

Manning, M. Lee. 2000. "A Brief History of the Middle School." *Clearing House* 73 (4): 192.

Martin, J. H. 1974. *Report of the National Panel on High Schools and Adolescent Education*. Washington, DC: U.S. Office of Education.

Merten, Don E. 1997. "The Meaning of Meanness: Popularity, Competition, and Conflict among Junior High School Girls." *Sociology of Education* 70 (3): 175–91.

Midgley, Carol, and K. C. Edelin. 1998. "Middle School Reform and Early Adolescent Well-Being: The Good News and the Bad." *Educational Psychologist* 33 (4): 195–206.

Murdock, Tamara B., Lynley H. Anderman, and Sheryl A. Hodge. 2000. "Middle-Grade Predictors of Students' Motivation and Behavior in High School." *Journal of Adolescent Research* 15 (3): 327–51.

Neild, Ruth Curran, Elizabeth Useem, Eva F. Travers, and Joy Lesnick. 2003. *Once and for All: Placing a Highly Qualified Teacher in Every Philadelphia Classroom*. Philadelphia: Research for Action.

Rabash, Jon, William Browne, Harvey Goldstein, Min Yang, Ian Plewis, Michael Healy, Geoff Woodhouse, David Draper, Ian Langford, and Toby Lewis. 2000. *A User's Guide to MLWin, Version 2.1a*. London: Institute of Education, University of London.

Reddy, R., J. E. Rhodes, and P. Mulhall. 2003. "The Influence of Teacher Support on Student Adjustment in the Middle School Years: A Latent Growth Curve Study." *Development and Psychopathology* 15 (1): 119–38.

Rodríguez, Germán, and Noreen Goldman. 1995. "An Assessment of Estimation Procedures for Multilevel Models with Binary Responses." *Journal of the Royal Statistical Society*, Series A, 58 (1): 73–89.

Sable, Jennifer, and Beth Aronstamm Young. 2003. *Characteristics of the 100 Largest Public Elementary and Secondary School Districts in the United States: 2001–02*. NCES 2003–353. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

School District of Philadelphia. 2003a. "Secondary Education Movement Strategic Plan," March 24, http://www.philsch.k12.pa.us/offices/communications/press_releases/032403/brief4.html.

School District of Philadelphia. 2003b. "Snapshot of the School District at Philadelphia," http://www.philsch.k12.pa.us/aboutus/.

Seidman, Edward, Larue Allen, J. Lawrence Aber, C. Mitchell, and J. Feinman. 1994. "The Impact of School Transitions in Early Adolescence on the Self-System and Social Context of Poor Urban Youth." *Child Development* 65 (4): 507–22.

Simmons, Roberta G., and Dale A. Blyth. 1987. *Moving into Adolescence: The Impact of Pubertal Change and School Context*. New York: De Gruyter.

Simmons, Roberta G., Dale A. Blyth, Edward F. Van Cleave, and Diane Mitsch Bush. 1979. "Entry into Early Adolescence: The Impact of School Structure, Puberty, and Early Dating on Self-Esteem." *American Sociological Review* 44 (6): 948–67.

Snijders, T. A., and R. J. Bosker. 1994. "Modeled Variance in Two-Level Models." *Sociological Methods and Research* 22 (3): 342–63.

U.S. Department of Education. 2001. *Digest of Educational Statistics*. Washington, DC: U.S. Department of Education.

Weiss, Christopher C., and Peter S. Bearman. 2004. "Fresh Starts: School Form and

Reexamining Middle School Effects

Student Outcomes. ISERP Working Paper 04-05, Institute for Social and Economic Research and Policy, Columbia University, New York.

Wilson, Bruce L., and R. E. Herriot. 1989. *Formal Subunits within American Public Schools: Their Structure, Power, and Consequences*. Philadelphia: Research for Better Schools.