School Discipline, Student Achievement, and Social Inequality

Richard Arum, Karly Ford, and Melissa Velez

In a series of lectures at the beginning of the 20th century (1902–1903), Emile Durkheim (1973 [1925]) argued for the centrality of school discipline in the process of youth socialization. According to Durkheim, school confronts youth as the first nonfamilial social institution (other than in some cases religion) that teaches students that there are external social norms, values, and rules that structure social interaction. Youth who internalize these norms and values are more likely to demonstrate conventional behaviors associated with productive employment and citizenship as adults, whereas the failure of schools and families to instill these values and norms in children is associated with delinquency, crime, and other outcomes at odds with the goals of these social institutions. While school discipline plays a role in allowing learning to occur and permitting educators to teach in work settings that are safe and professionally conducive to teaching, Durkheim (1973 [1925]) argued that school discipline was "not a simple device for securing superficial peace in the classroom" but, more important, "an instrument—difficult to duplicate—of moral education" (148–49).

Sociologists in particular have been interested in the role of school disciplinary environments in shaping individual student outcomes. Although discipline has long been recognized as a central feature of successful schools, researchers in recent decades have given relatively scant attention to comparative studies on this critical topic. Comparative research on school discipline is especially important at this time because we need to know more about how schools vary systematically in their approaches to discipline and how youth development varies with respect to school context. Such knowledge gives policy makers a sound empirical basis for the formation of

effective educational policies. This book provides the first systematic comparative cross-national study of school disciplinary climates and aspires to contribute to academic knowledge, public understanding, and educational policy formation.

School disciplinary climates are made up of multiple elements and are best conceptualized as joint functions of the actions of students and educators. First, administrators' and teachers' actions to maintain school order—that is, school discipline as *administrative regulation*, or social control—set the parameters within which student attitudes, behaviors, and subcultures in schools are expressed. Second, school discipline manifests itself not just in administrative actions but also in student behaviors, norms, and values—that is, school discipline as *peer environment*. A school's *disciplinary climate* thus can be conceptualized as a product of the actions of teachers and administrators, the cultural beliefs and behaviors of students, and the interactions between students and educators that shape the school's organizational culture. These school-level processes, of course, are embedded in and structured by larger institutional, social, and cultural contexts that constitute the organizational field in which schools are situated.

School disciplinary climates not only are potentially associated with academic achievement but also provide the institutional context in which student moral development occurs. Sociologists have focused on the potentially critical role school plays in shaping the attitudes and dispositions of youth. Following Durkheim's thinking, sociologists have understood the moral dimensions of schooling as those aspects primarily related to the capacity to shape youth attitudes and dispositions in a manner aligned with normative expectations.

This contrasts with other developmental approaches that have understood moral education in terms of growth in the capacity of individuals to apply abstract systems of moral reasoning to problems they confront in the world (e.g., Kohlberg 1981). For sociologists, moral education is considerably simpler; it can be understood as schooling—and particularly school discipline—that plays a potentially critical role in children's and adolescents' internalization of conventional social expectations and norms.

Although social scientists regard school disciplinary climate as central to school effectiveness and individual development, surprisingly little systematic empirical research has been focused on examining the causes and

consequences of variation in these climates. An early exception to this lack of attention to school discipline is the work of James Coleman and his colleagues. In 1959 Coleman postulates in *Adolescent Society* that students' educational orientations and behaviors are the product of specific school organizational environments and that these behavioral orientations can shape educational outcomes. The 1966 "Coleman Report," *Equality of Educational Opportunity*, further develops this theme, demonstrating that, after controlling for students' social background, "differences between schools account for only a small fraction of differences in pupil achievement" (22). To the extent that schools do affect student outcomes, however, the most important characteristic of schools is not facilities, curriculum, or teacher quality but peer environment. Coleman notes that "a pupil's achievement is strongly related to the educational backgrounds of the other students in the school" (22) and maintains that

> a child's fellow-students provide challenges to achievement and distractions from achievement; they provide the opportunities to learn outside the classroom, through association and casual discussions. Indeed, when parents and educators think of a "good school" in a community, they most often measure it by the kind of student body it contains: college-bound and high achieving. (183)

Coleman and his colleagues extend this focus on peer environments in the United States with research using the High School and Beyond study, a national probability survey of students in the 1980s. Coleman's colleague Tom DiPrete (1981) and his associates find that students in school climates with stricter discipline in 10th grade have lower rates of 12th grade misbehavior. In subsequent work exploring differences between U.S. public and private school student outcomes, Coleman and Thomas Hoffer (1987) link student behavioral climate to growth in cognitive performance between 10th and 12th grades and to differences across school sectors. Coleman, Hoffer, and Sally Kilgore (1982, 171) find that high school sophomores' reports of other student misbehavior (i.e., peer absenteeism, cutting class, students fighting each other, and students threatening teachers) were less frequent in Catholic schools and account for 33 percent of the higher performance of Catholic students relative to public school students on standardized reading tests and 46 percent of the difference between them on standardized mathematics tests.

Additional research on U.S. schools highlights how misbehaving students have lower levels of educational achievement as measured by changes in grades and test scores (Myers et al. 1987). Consistent with this position, researchers such as Paul Barton, Richard Coley, and Harold Wenglinsky (1998) argue that stricter school disciplinary practices are associated with improved student behavior. Richard Arum (2003) highlights how students' perception of the fairness (or legitimacy) of school discipline has a greater impact on student outcomes than perceived strictness of school regulation and sanctions against students. Research on school behavioral climates often also draws on student-level victimization data. This research suggests that certain school factors, such as school size, student composition, and school location, are associated with variation in victimization rates (Gottfredson and Gottfredson 1985).

In recent decades, educational systems have increasingly monitored and assessed reports of victimization in efforts to reduce school violence and increase school safety. Social scientists have been integral to this development, organizing themselves through such collaborative efforts as the International Conference on School Violence and the *International Journal of School Violence* (Benbenishty and Astor 2008). While these efforts have typically focused on developing programmatic interventions to reduce violence and improve school climate in settings with high levels of victimization, they have only recently begun contributing to the advance of social scientific knowledge on the structural and organizational differences across countries that are associated with student disciplinary climates (Benbenishty and Astor 2008).

Only a few systematic efforts have been launched to identify and understand how school discipline is structured differently across modern industrial societies. Two strands of comparative research are exceptions to this lack of empirical work. In the first, a few comparative studies highlight cultural explanations for differences in school discipline between the United States and Asian countries (e.g., Stigler, Lee, and Stevenson 1987). While these studies are fascinating and informative, accounting for cross-national differences largely in terms of national cultures, they do not systematically map out or account for differences in school disciplinary climates in either structural or institutional terms. A second line of research takes advantage of cross-national datasets to identify variation in student victimization, but it fails to advance compelling structural and institutional accounts of these

differences (e.g., Craig et al. 2009; Akiba et al. 2002). Wendy Craig and colleagues (2009) analyze the Health Behavior in School-Aged Children (HBSC) 2005–2006 survey of more than 200,000 students in 40 countries and conclude that adolescents in Baltic nations have higher rates of victimization. Motoko Akiba and colleagues (2002) find that student victimization is lower in countries with higher levels of economic development (measured by gross domestic product per capita) but is not influenced by the income inequality within a country or by the percentage of linguistic minorities as a share of the population.

In this book a group of leading international social science researchers addresses this dearth of studies with systematic comparative research on a set of nine strategically chosen national case studies. The goal of the project is to identify the institutional determinants of variation in school discipline and the association of these school contexts with student achievement. How do these countries vary in terms of population heterogeneity, organizational structure of the educational system (e.g., centralization, privatization, and tracking), legal rights of students, administration of school discipline, and other factors that could potentially account for differences in school disciplinary climates? Which structural and institutional factors at the country and school level are related to school disciplinary climates? And finally, how are associations between social background and academic achievement related to school discipline?

RESEARCH FRAMEWORK AND COMPARATIVE METHODOLOGY

To address these questions, we employ a collaborative comparative methodology similar to that used by Yossi Shavit and Hans-Peter Blossfeld (1993), Shavit and Walter Müller (1998), Arum and Müller (2004), and Shavit, Arum, and Adam Gamoran (2007). Research teams in nine countries address the overarching research questions using a common theoretical and methodological framework, which allows comparison across countries. Figure I.1 highlights the conceptual framework adopted for the analysis. The authors of the country chapters begin their exploration of school discipline by providing a detailed description of four institutional dimensions of school disciplinary context theorized to have important implications for structuring student behavioral climates. These dimensions are defined as
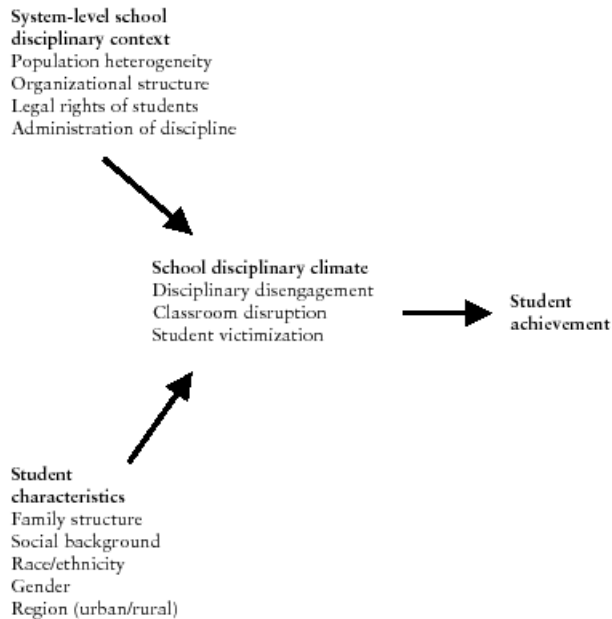
System-level school disciplinary context
Population heterogeneity
Organizational structure
Legal rights of students
Administration of discipline

School disciplinary climate
Disciplinary disengagement
Classroom disruption
Student victimization

Student achievement

Student characteristics
Family structure
Social background
Race/ethnicity
Gender
Region (urban/rural)

Figure I.1.   Research design

(1) population heterogeneity and (2) organizational structure of educational systems, two components of national contexts that vary cross-nationally and that we quantitatively identify for this study, and (3) the legal rights of students and (4) the administration of discipline, two additional elements of national differences that we describe qualitatively.

Population heterogeneity is explored in terms of patterns of immigration and socioeconomic inequality that create pedagogical challenges and, often, public anxiety regarding assimilation, integration, or youth socialization. Previous social science scholarship suggests that student misbehavior can emerge from oppositional subcultures associated with these social differences (e.g., Ogbu 1978; Fordham and Ogbu 1986).

Organizational structure of educational systems is identified primarily in terms of school stratification (e.g., Müller and Shavit 1998), with particular focus on curricular tracking early in a student's academic career. School systems vary cross-nationally by whether they are stratified by curricular tracking, which separates youth at early ages on the basis of ability, academic performance, occupational orientations, and other factors,

including such ascriptive characteristics as race, social class, and gender. In addition, where applicable for the educational systems of particular countries, we consider variation in school centralization (i.e., organization and administration from a central national bureaucratic authority) and privatization (i.e., private sector alternatives to publicly organized schooling). These institutional configurations are expected to affect student behavior through the sorting and segregation of students prone to misbehavior, individual assessments of the likelihood of success related to school persistence, and within-country variance of institutional characteristics.

Legal rights of students are examined and described qualitatively in terms of students' legal entitlements related to the administration of school discipline. We also explore court adjudication in student and parent disputes with school administrators over discipline issues. What formal legal rights are afforded to students in the country? Have formal definitions of the legal rights of students diffused internationally? Is there any evidence (e.g., media reports) that students and parents use the legal system to challenge school disciplinary practices (e.g., are educators ever sued for disciplining students)? Earlier research (Arum 2003) demonstrates how laws influence disciplinary practices and the perceived legitimacy of administrative efforts to sanction disruptive student behavior. Labor economist Philip Babcock (2009) further extends this line of research to highlight associations between legal environments, school disciplinary climate, and student behavior in the National Study of Adolescent Health[1] (Add Health), a study of adolescents in 132 schools, grades 7–12, across the United States. Babcock demonstrates that legal environments are associated with school disciplinary policies and that students attending schools with stricter discipline have lower truancy, greater likelihood of high school graduation, and improved employment outcomes.

Administration of school discipline is identified and described qualitatively in terms of the particular disciplinary forms and procedures employed by an educational system. We explore how schools impose discipline on students, to what extent school discipline is formalized, and how schools deal with minor and serious disciplinary problems. In the case of serious disciplinary issues, we pay particular attention to the use of mandatory school transfers, expulsion, and corporal punishment. Theoretically, these institutional practices should have direct effects on student behavioral climate.

Following descriptions of institutional variation in each country, authors assess within-country variation in school discipline and its association with individual student background characteristics and student achievement using 2003 Trends in International Mathematics and Science Study (TIMSS) data (Mullis et al. 2004). Analyses in each country explore the association between student background, school-level characteristics (i.e., student body composition in terms of student social background and heterogeneity), and school disciplinary climate. Researchers model the association between school disciplinary climate and student achievement on math and science tests and model how variation in school disciplinary climate mediates the effects of social background on student achievement. In many chapters, authors use other national data to supplement the comparative TIMSS analysis. This chapter provides details on the 2003 TIMSS data, measures, results from pooled analysis of all countries' data, and a summary of each country's results.

TIMSS 2003 DATA

The TIMSS project, conducted by the International Association for the Evaluation of Educational Achievement (IEA), gathered data from fourth and eighth grade students,[2] teachers, and administrators in 49 countries, resulting in a dataset containing over 360,000 students, 25,000 teachers, and 12,000 principals. All participating countries employed a stratified multistage sampling design in which at least 150 schools were chosen at the first stage using probability-proportional-to-size sampling, and then one or two classes within each school were randomly selected at the second stage. This book includes TIMSS eighth grade data collected from Canada, Chile, Israel, Italy, Japan, South Korea, the Netherlands, Russia, and the United States.[3] Finally, two qualifications are worth noting. First, data on the United States and Canada were limited to specific states or provinces and thus are not nationally representative. Second, the United States did not ask students about individual victimization.

The primary goal of the TIMSS project was to measure trends in math and science achievement across countries. To that end TIMSS researchers in each country administered comprehensive math and science assessments near the end of the school year in each country. Countries whose school

year ends in November or December administered the tests in October or November of 2002, while countries whose school year ends in June administered the tests in April, May, or June of 2003. In addition to the assessments, students filled out questionnaires on their attitude toward school, their home environment, and the school climate. Math and science teachers of sampled students provided information about their professional training and background, instructional practices, and curriculum. Principals (or headmasters) were also administered a questionnaire on school demographics, teacher support, school staffing and resources, and math and science course offerings. Student, teacher, and administrator surveys contained information pertinent to this book about school disciplinary climates, including questions about school violence, school safety, and student victimization.

MEASUREMENT OF VARIABLES

Cognitive Performance: Math and Science Assessments

Math and science assessments contained two dimensions, a content domain and a cognitive domain, administered in multiple-choice and open-ended formats. In math there were five content domains (numbers, algebra, measurement, geometry, and data) and four cognitive domains (knowing facts and procedures, using concepts, solving routine problems, and using reasoning). The science assessment had five content domains (life science, chemistry, physics, earth science, and environmental science) and three cognitive domains (factual knowledge, conceptual understanding, and reasoning and analysis). Because the extensive number of questions would have overburdened students, they were administered a subsection of items chosen using a matrix-sampling technique in which questions were chosen to create balanced designs containing math and science items. To create individual test scores, responses were scaled to derive estimates of what students would have scored if they had completed the entire test. Five imputed scores were provided for each student in the math and science subsections, respectively, each with an average of 500 and a standard deviation of 100. For this book, the math score is the average of the five imputed math scores, the science score is the average of the five imputed science scores, and the combined score is the sum of the averaged math and science scores.

Student Characteristics

Student background characteristics were drawn from student responses to the student questionnaire. Gender was measured using a dummy variable indicating a male student, a continuous age variable was calculated by TIMSS researchers using student reports of birth month and year, and immigrant status was measured using a dummy variable indicating that a student was not born in the country where the test was taken. Highest parental education was measured as the highest level of education completed by a student's mother (or stepmother or female guardian) or father (or stepfather or male guardian) and was standardized across countries into the following categories: did not go to school or did not finish primary school, finished primary school, finished lower secondary school, finished upper secondary school, finished postsecondary nontertiary education (e.g., vocational training), finished tertiary education designed to provide direct access to the labor market, finished tertiary education designed to prepare students for skilled work, and received beyond tertiary education. A categorical variable indicated the number of books students reported having in their household (none or very few, enough to fill one shelf, enough to fill one bookcase, enough to fill two bookcases, and enough to fill three or more bookcases), and a continuous variable indicated the number of people in the students' homes (two, three, four, five, six, seven, eight, or more). Finally, student expectations about educational attainment were measured using a dummy variable indicating that students expected to finish college.

School and Community Characteristics

School and community variables included principal reports of the school size (adjusted for skewness in the distribution with a logarithmic transformation), the school's highest grade level, and a categorical measure of the size of the community the school was located in (fewer than 3,000 people, 3,001 to 15,000 people, 15,001 to 50,000 people, 50,001 to 100,000 people, 100,001 to 500,000 people, and more than 500,000 people).

Because principal responses were not available for several school-level measures of interest, they were derived by taking the average of individual student measures. Specifically, the number of male students,[4] the number of immigrant students, and the highest level of parental education were identified by averaging the characteristics of sampled students by school.

The average numbers of males and immigrants in the school were then divided into categories based on an evaluation of the naturally occurring breaks in the distribution across countries. As a result, the average number of male students in a school was divided into categories of 0–45 percent male, 46–60 percent male, and more than 60 percent male. The average number of immigrant students in the school was divided into categories of 0 immigrants, 1–10 percent immigrants, and more than 10 percent immigrants. For our project, we calculate the school-level average of the highest level of parental education, and we create a measure of variation in parental education by dividing the school-level standard deviation of parental education by the mean of the highest parental education of the school.

School Disciplinary Climate

School disciplinary climate was measured using principal, teacher, and student information collected from their respective questionnaires. Disciplinary disengagement was measured using principal responses to questions regarding how often arriving late at school, absenteeism, and skipping class occurred among eighth graders at their school (never, rarely, monthly, weekly, or daily).[5] If principals answered at least two of these questions, the responses were averaged to create the discipline disengagement index (Cronbach alpha is .76 for all schools; individual country alphas ranged from .67 to .87). Higher numbers on the index indicate more disciplinary disengagement in the school.

Frequency of classroom disruption was measured by a question administered to math and science teachers asking how often disruptive students limit how they teach their math or science class (not at all, a little, some, a lot). The school average of these responses was calculated separately for math and science teachers; those averages were then averaged to create one teacher average per school. Higher averages indicate more classroom disruption.

Student perceptions of disciplinary climate were captured using a student victimization index created by summing variables indicating whether students had reported that something of theirs was stolen, they had been hit or hurt by other students, or they had been made fun of or called names in school during the past month (α = .53 for all students; individual country α's ranged from .47 to .56).[6] A school-level student victimization measure

was calculated by taking the school-level average of individual student measures.[7] Higher numbers on these indices indicate more victimization.

All chapters include descriptive statistics, school-level models predicting disciplinary climate, and student-level models predicting student outcomes. Some present additional TIMSS analyses or supplementary data from other sources. School-level models use ordinary least squares regression to examine the relationship between school-level characteristics and school-level disciplinary climate measures. Student-level analyses can take several forms but at a minimum include hierarchical linear models (HLMs) that examine the relationship between student-level measures, school-level characteristics, disciplinary climate, and math and science test scores. Some countries supplement the test score models with additional models examining students' college expectations or student victimization as outcomes.[8]

Using HLMs for the student-level outcomes is appropriate for these data because they adjust for the clustering of students within schools by calculating separate student- and school-level equations. In the student-level equation, an outcome Y for student i in school j is predicted by k student-level variables:

$$Y_{ij} = \beta_{0j} + \sum_{1}^{k} \beta_{kj} X_{kij} + r_{ij},$$

where β_{0j} represents the average outcome of school j adjusted for student characteristics (X) included in the model and r_{ij} is the student-specific error.

In school-level equations, the school average of the student outcome is predicted by

$$\beta_{0j} = \gamma_{00} + \sum_{1}^{k} \gamma_{0k} W_{kj} + u_{0j},$$

where school-level outcomes (β_{0j}) are predicted by the sum of the average intercept across groups (γ_{00}); a vector of school-level variables (W_{kj}), including school discipline measures; and a school-specific error (u_{0j}).

A second school-level equation indicates that the effects of the student-level variables are assumed not to vary across schools:

$$\beta_{kj} = \gamma_{k0}.$$

While our multilevel models control extensively for covariates that might affect estimates of relationships of interest, it is important for readers to recognize that our comparative project relies on cross-sectional observational data that limit the extent to which causal inferences can be drawn. In particular, disciplinary problems and student achievement are endogenous, and thus from cross-sectional observational data one would not want to assume causality. Nevertheless, descriptive modeling of how school discipline and its association with school inputs and outcomes vary within and across countries can still usefully contribute to sociological understandings and inform educational policy and practice by identifying how these factors covary with respect to other variables.

WEIGHTS

For analyses to accurately reflect the populations of participating countries, we add sampling weights provided by TIMSS to the student- and school-level analyses. School-level analyses include a school-level weight that adjusts for a school's probability of selection and nonparticipation. Student-level analyses for individuals are weighted using a student-level weight, HOUWGT (house weight), that adjusts for the selection probability of schools, classrooms, and students and for school, classroom, and student nonparticipation. Student-level analyses conducted on the pooled sample of all countries in the project are weighted using a variation of the student-level weight, SENWGT (senate weight), that is adjusted so that each country contributes equally to the analyses regardless of population.

MISSING DATA

Unless otherwise noted, authors use mean substitution to handle missing data.[9] All missing student- and school-level covariates are mean substituted with the exception of the disciplinary climate measures and student- and school-level gender measures. Missing data from continuous variables are replaced with the means from the sample, and missing-data dummy variables are created to flag observations with mean substitution. Missing data from categorical measures are set to 0 and then flagged using a missing-data dummy variable. The missing-data dummy variables are included in all multivariate analyses but not reported in the tables.