DOES GETTING A’S REALLY MATTER? A CONCEPTUALIZATION OF GRADES
AS A MEASURE OF EDUCATIONAL OUTCOMES

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

By
Brandy J. Ellison

Maureen T. Hallinan, Director

Graduate Program in Sociology
Notre Dame, Indiana
October 2008
Does Getting A’s Really Matter? A Conceptualization of Grades as a Measure of Educational Outcomes

Abstract

by

Brandy J. Ellison

Assessing educational outcomes is a matter of great importance for everyone in a society that relies primarily on public schools to shape the next generation of citizens. Given the vital nature of this task, it is imperative that the researchers who undertake it have an informed understanding of the measures used to evaluate educational outcomes. This dissertation contributes to that task by analyzing grades as a potential complement or alternative to standardized test scores as a measure of educational outcomes.

The research is based on a theoretical framework that emphasizes the importance of both instruction and socialization as goals of the education system. Grades are theorized to assess elements of both of these goals because they account for academic achievement and student behavior, while standardized test scores primarily reflect knowledge and instruction. Therefore, grades are expected to provide a more comprehensive evaluation of student achievement than standardized test scores. In addition, grades function as feedback to students that can affect their future outcomes. Given the comprehensive nature of grades and their role in shaping students’ later
outcomes, the theory establishes the necessity of developing a thorough conceptualization of grades.

Both quantitative and qualitative data are utilized to test elements of the theory. The analyses examine several aspects of grades and provide information that contributes to a more complete understanding of grades than has been available in the past. The research confirms that grades account for students’ behaviors and achievement orientations while standardized test scores do not offer a thorough assessment of these characteristics of students. Observations clarify the importance of behavior for mediating opportunities to learn within a classroom setting and support the inclusion of behavioral assessments in grades. The analyses also highlight the role of teachers in interpreting and evaluating students’ behaviors in grades. In addition, teachers are found to mediate the impact of students’ behavior on their own academic achievement and on classmates’ opportunities to learn.

The models estimated in this study demonstrate that grades can serve as an appropriate supplement to standardized test scores or act on their own as measures of educational outcomes. Their viability as outcomes in analyses is determined based on the variables used to predict achievement and the information in which researchers are interested. Because grades act as a proxy for behavior and achievement orientation and also act as feedback to shape students’ academic self-concepts, they are valuable predictors of outcomes that are influenced by those characteristics of students. In many situations, using both grades and standardized test scores to predict outcomes is appropriate because each accounts for different aspects of educational achievement.
Overall, the study establishes the usefulness of grades as measures of educational outcomes. It accomplishes this goal by generating a theoretical basis for understanding the meaning and function of grades. The research also sets a foundation for understanding what grades represent in analyses and how they can be appropriately utilized and interpreted in research.
CONTENTS

Figures ........................................................................................................................................ iv

Tables ......................................................................................................................................... v

Chapter 1: Introduction ............................................................................................................ 7

Chapter 2: Review of the Literature ......................................................................................... 13

Chapter 3: Theoretical Foundation for Understanding Grades as a Measure of 
Educational Outcomes ............................................................................................................ 20
  3.1 Organizational Mandates ................................................................................................. 22
    3.1.1 Instruction ............................................................................................................... 22
    3.1.2 Socialization .......................................................................................................... 24
  3.2 Intersection of Instruction and Socialization ................................................................... 26
  3.3 Transmission of Organizational Mandates ..................................................................... 27
  3.4 Feedback .......................................................................................................................... 28
    3.4.1 Teacher Evaluations of Students ......................................................................... 28
    3.4.2 The Impact of Grades on Students ....................................................................... 30
  3.5 Conclusion ........................................................................................................................ 33

Chapter 4: Relationships Between Teachers’ Reports of Students’ Behaviors and Grades .... 35

Chapter 5: Relationships Between Teachers’ Reports of Students’ Behaviors and Test Scores ........................................................................................................................................ 65

Chapter 6: Relationships Between Students’ Reports of Their Behaviors and Grades ... 77

Chapter 7: Analyses of Classroom Observations .................................................................... 91
  7.1 Data and Methods ............................................................................................................ 92
  7.2 Theoretical Framework for Analyzing Observations ..................................................... 95
  7.3 Explanations for Teachers Not Having Control of Instruction ...................................... 98
  7.4 Determinants of Students’ Classroom Behaviors ........................................................ 103
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5 The Impact of Students’ Behaviors on Instruction</td>
<td>105</td>
</tr>
<tr>
<td>7.6 Management of the Diffusion of the Effects of Students’ Behaviors</td>
<td>108</td>
</tr>
<tr>
<td>7.7 Teachers’ Reactions to Students’ Behaviors</td>
<td>112</td>
</tr>
<tr>
<td>7.8 Conclusion</td>
<td>119</td>
</tr>
<tr>
<td>Chapter 8: Relationships Between Grades and Test Scores and Students’ Educational and Occupational Outcomes</td>
<td>125</td>
</tr>
<tr>
<td>Chapter 9: Conclusion</td>
<td>161</td>
</tr>
<tr>
<td>Bibliography</td>
<td>171</td>
</tr>
</tbody>
</table>
FIGURES

Figure 4.1: The graph presents the distribution of math test scores by grades. The data source is the National Education Longitudinal Study of 1988. The graph was created using Stata 10. ................................................................. 41

Figure 4.2: The graph presents the distribution of English test scores by grades. The data source is the National Education Longitudinal Study of 1988. The graph was created using Stata 10. ................................................................. 42
TABLES

Table 4.1 Descriptive Statistics For Variables In The Analyses ........................................ 37

Table 4.2 Regression of Math and English Grades On Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, Behaviors, and Educational Futures .................................................................................................................. 45

Table 5.1 Regression of Math and English Test Scores On Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, Behaviors, and Educational Futures .................................................................................................................. 67

Table 5.2 Regression of Math and English Test Scores On Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, and Behaviors ................................................. 72

Table 6.1 Descriptive Statistics for Student Variables in the Analyses................................. 78

Table 6.2 Regression of Math Grades On Students’ and Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, Behaviors, and Educational Futures ........................................................................................................ 79

Table 6.2 (Contd.) .................................................................................................................. 80

Table 6.3 Regression of English Grades On Students’ and Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, Behaviors, and Educational Futures ........................................................................................................ 80

Table 6.3 (Contd.) .................................................................................................................. 81

Table 8.1 Descriptive Statistics for Dependent Variables in the Analyses......................... 126

Table 8.2 Multinomial Logistic Regression of Grades and Test Scores On Enrollment in Postsecondary Education ......................................................................................................................... 131

Table 8.3 Multinomial Logistic Regression of Grades, Test Scores, and Control Variables On Enrollment in Postsecondary Education ................................................................................. 132
Table 8.4 Predicted Probabilities for Postsecondary Enrollment Based On Grades and Test Scores

Table 8.5 Predicted Probabilities for Postsecondary Enrollment Based On Grades, Test Scores, and Control Variables

Table 8.6 Logistic Regression of Grades, Test Scores, and Control Variables On Graduation from a Four-year College

Table 8.7 Predicted Probabilities for Graduation from a Four-year College Based On Grades, Test Scores, and Control Variables

Table 8.8 Predicted Probabilities for Graduation from a Four-year College Based On Grades, Test Scores, and Control Variables for Students Who Enrolled in a Four-year Program

Table 8.9 Regression of Grades and Test Scores On Self-reported Income in 1999

Table 8.10 Regression of Grades, Test Scores, and Educational Attainment On Self-reported Income in 1999

Table 8.11 Predicted Probabilities for Income in 1999 Based On Grades, Test Scores, and Control Variables

Table 8.12 Predicted Probabilities for Income in 1999 Based On Grades, Test Scores, Control Variables, and Educational Attainment
CHAPTER 1:
INTRODUCTION

Legislators in Texas recently implemented a controversial plan to grant admission to a state university to all students who graduate in the top 10% of their high school class based on their grade point average (Glater 2004). Critics of the plan state that students from low-performing schools who graduate in the top 10% of their class will not be as well-prepared for postsecondary education as students who graduate below the top 10% in high-performing schools. They believe that high grades do not accurately indicate preparedness for college and standardized test scores are a better measure for granting college admission because they rank students across schools and provide an objective measure of college preparedness.

An examination of educational research and policy indicates that the belief in the appropriateness of standardized test scores as a measure of educational success is widespread. Researchers, educators, and policymakers frequently utilize standardized test scores to assess students, teachers, schools, and instructional programs. Following the passage of the No Child Left Behind Act, this practice has become even more common. Under the system of accountability established by the act, standardized test scores are used to determine if students meet adequate yearly progress in academic subjects (U.S. Department of Education 2001). Decisions regarding funding, instruction, and school leadership are made based on these results.
The debate surrounding the Texas plan illustrates an unresolved conflict regarding measures of educational assessment and their use in research and policy decisions. The prevalence of test scores as indicators of students’ educational outcomes for research and policy purposes indicates that standardized test scores have become the default winner in the debate. This is troublesome because the improper use of educational assessment measures in research may result in misleading conclusions or futile reform efforts.

The value of assessing educational outcomes is not in doubt. Education is publicly-funded and it contributes to the goal of a knowledgeable and socialized citizenry. For these reasons, individuals and organizations have a well-founded interest in evaluating educational successes and failures. What is missing from the evaluation efforts is an exploration of possible supplements or alternatives to test scores, such as grades.

Two basic conflicts exist regarding whether or not standardized test scores or grades are the best measure to use to examine educational outcomes. The first is whether a norm- or criterion-referenced measure is most valid for assessing educational outcomes. Standardized test scores are criterion-referenced in that they offer a baseline for performance standards and rank students independently of each other. Grades are likely to be norm-referenced because they reflect what individual teachers expect of students and implicitly or explicitly compare students to other students within the classroom or school, thereby reflecting organizational variations and classroom processes that vary across schools.

The second conflict revolves around what aspects of education will benefit students most as they transition through school and into occupations. The ostensible
rationale behind the use of test scores is that they are impersonal, unbiased, and only measure knowledge gained in the classroom. Although these characteristics of standardized test scores have been debated, the scores are still thought to provide the closest available approximation of the qualities listed above. Therefore, researchers and policymakers believe that they provide an objective evaluation of knowledge and students’ progress. Indeed, if knowledge gained is the outcome in which evaluators are most interested, standardized test scores may be the most appropriate measure. However, the amount of knowledge that students gain during their time in school may not be the best predictor of outcomes of interest to researchers and policymakers.

The primary aspect of education that test scores do not fully reflect is that education is an inherently social process. Because tests are standardized and impersonal they do not account for the social framework in which learning takes place. And, because they measure academic performance at only one point in time, they do not reflect the fact that learning is developmental and takes place across time and in changing social contexts. Finally, tests underestimate the importance of meeting behavioral standards for achieving successful outcomes in school and work settings.

Grades are more likely than standardized tests to account for social aspects of education. However, many researchers are reluctant to use grades in analyses for this reason. They believe that grades offer a subjective interpretation of achievement which is clouded by social factors that are unrelated to academics. It is possible, however, that the social facet of education, as captured in grades, is an important reflection of students’ opportunities for success in the future. Because test scores are not based on information
regarding students’ social experiences in school, they may lack predictive power when analyzing outcomes that have a social component.

In order to address the conflicts described above, the research presented here will begin an examination of grades as an alternative or complement to standardized test scores for assessing educational outcomes. It will contribute to this goal by exploring the relationships between classroom factors and students' outcomes as measured by their grades. The factors to be analyzed include instructional experiences and students’ achievement orientation as measured by social interactions, behavior, and aspirations.

The research will also examine the effects of academic feedback and the predictive power of outcome measures by analyzing how well grades predict students' likelihood of completing future educational transitions and occupational outcomes. The research will focus on grades assigned to secondary school students because the students are at a point in their educational careers when they have gained an awareness of the requirements – both academic and social – for earning good grades and can choose to act on that information. In addition, both students and teachers may be considering the role of grades in determining later educational and occupational outcomes.

Students’ grades are not the only measure that could be used to analyze educational outcomes. However, grades were chosen as the most feasible complement or alternative to standardized test scores for several reasons. First, like standardized test scores, they are nearly universal as an educational assessment tool in the United States and they are widely available in educational datasets. Second, grades may capture social inputs that are not reflected in standardized test scores and, therefore, may explain a broader domain of outcomes. Finally, grades are likely to be more salient to students,
teachers, and parents as indicators of academic progress than are standardized test scores and therefore may have a greater impact on students and their educational trajectories.

The first reason for examining grades noted above - their widespread usage and availability - makes them both an important and viable option for evaluating educational assessment measures. The second two reasons - the fact that grades capture more social inputs and provide more salient feedback - suggest that grades are a different and more comprehensive measure of educational processes than test scores. Because of these differences, grades offer a compelling alternative to standardized test scores as a measure for educational evaluation purposes. The ultimate goal of the research is to develop a conceptual understanding of grades and determine how their use in research and policy can contribute to the continuing goal of improving the educational system.

It is hypothesized that grades will reflect aspects of students’ instructional experiences, social interactions, behavior, and acceptance of educational norms. Specifically, students’ grades will improve as the quality and quantity of instruction increases, if they have positive social interactions in school, adopt classroom norms, and value educational pursuits. Because grades are hypothesized to be a more socially inclusive measure than standardized test scores and to act as more salient academic feedback, it is also hypothesized that high grades will be strong predictors of students’ outcomes such as college graduation or post-high school income, even after controlling for test scores.

Research will be conducted in three phases. First, quantitative analyses will be performed to determine which educational processes – both academic and social – contribute to grades. This is done because understanding the determinants of grades will
enable researchers to better understand how to use them in research both as indicators of educational progress and as predictors of outcomes. Second, classroom observations will be used to gain a deeper understanding of the day-to-day production of grades within a classroom setting. This portion of the research is undertaken because grades represent an interactive classroom process that takes place over a significant period of time. Finally, quantitative analyses will be performed to determine how well high school grades function as predictors of college enrollment, college graduation, and income. This portion of the research will establish the validity of grades as predictors of students’ future outcomes.
CHAPTER 2:
REVIEW OF THE LITERATURE

Sociologists utilize standardized test scores and grades in research, as both dependent and independent variables, on a regular basis. When analyzing achievement levels and achievement gains, standardized test scores are the most commonly employed outcome within the sociology of education literature. Several prominent findings are based on standardized test scores as a dependent variable. These include examinations of the effects of school sector (Hoffer, Greeley, and Coleman 1985; Gamoran 1996), ability grouping (Gamoran 1992), parental involvement (Sui-Chu and Willms 1996), sibship size (Downey 1995), social networks (Morgan and Sorensen 1999), school organization (Chubb and Moe 1990), and race (Smerdon 1999) on students’ academic achievement. An implicit underpinning of this research is either that standardized test scores reflect educational achievements beyond how much knowledge a student gained or that knowledge gained is a sufficient indicator of educational outcomes. A corollary assumption is that whatever is measured by standardized test scores can accurately predict students’ opportunities for success in future educational locations, employment, and other life outcomes.

Grades are less common as a dependent variable in sociological literature. When they are used, the study is frequently investigating grading practices and the role of noncognitive behaviors in determining grades rather than assessing students’ academic
achievement outcomes (Farkas et al. 1990; Rosenbaum 2001; Kelly 2008). Those who
do use grades as a dependent variable in models of students’ academic achievement often
do not provide explicit reasons for their choice but it may be inferred that they chose
grades for theoretical reasons (Ainsworth-Darnell and Downey 1998) or because the
dataset did not contain test scores (DiMaggio 1982; Ferguson 2001).

Carbonaro (1998) uses both grades and standardized test scores as dependent
variables in separate models. He acknowledges that the two dependent variables may
represent different academic outcomes and his results indicate that this is true because
intergenerational social closure is significantly positively related to test scores but not
significantly related to cumulative grade point average. Downey and Vogt Yuan (2005)
also use both test scores and grades as dependent variables for their research examining
gender differences in academic outcomes. They reason that the two measures represent
distinct academic outcomes because female students receive better grades than male
students in all subjects even though male students perform better on standardized tests of
math ability. They theorize that, because of gender socialization, female students receive
better grades because they exhibit more appropriate classroom behavior than male
students. They find that teachers do express more approval of female students’ classroom
behavior and this helps to explain why female students receive better grades than their
male peers, indicating that behavior impacts grades more than it impacts standardized test
scores.

Overall, sociological studies that use grades as a dependent variable are an
exception to the rule. The preference for using standardized test scores as a dependent
variable when examining the quality of instruction and effects of social factors on
educational outcomes indicates that many sociologists do not consider grades to be a reliable measure of academic achievement. Those studies which do exist indicate that grades and standardized test scores represent different aspects of students’ educational outcomes. However, they do not provide a thorough basis for understanding what each measure represents and how social and academic factors may affect them differently.

When used as control or explanatory variables, standardized test scores and grades are often considered interchangeable or highly similar. Researchers may employ one or the other depending on which is available or most easily utilized. Some studies use both measures. These studies, although not intended as analyses of assessment measures, may reveal differences in the validity of grades and standardized test scores as predictors of educational outcomes. For instance, Carbonaro (2004) and Ellison (2005) each found that standardized test scores predict college enrollment but fail to predict college graduation while grades are significantly associated with each outcome. These findings are incidental to the larger purpose of each study, but they are enlightening in that they indicate differences in the predictive validity of the two measures.

The evidence provided above signifies a gap in the sociological literature. While sociologists are interested in examining academic outcomes, they have not thoroughly considered the best indicators for undertaking the analyses. The research does indicate that grades and standardized test scores are different measures of academic outcomes because they are both predicted by disparate educational factors and function differently as predictors of later outcomes. Therefore, indiscriminately choosing one or the other for use in research may be inappropriate if it is not truly representative of the outcome in which the researcher is interested. The preference for using standardized test scores
when the option is available also indicates that grades may be undervalued as a measure of educational outcomes. An assessment of grades is needed in order to determine how they may most appropriately be used in research.

With a few exceptions that are discussed below, sociologists have left the matter of assessing these measures of educational outcomes to educational psychologists. Popular journals in this field, such as the Journal of Educational Measurement, are devoted mainly to examining standardized tests and determining their validity as indicators of students’ knowledge gains. While this pursuit is important, it does not fully acknowledge the social context of education and the potential usefulness of more subjective, socially-embedded measures such as grades.

Research in educational psychology that has compared grades and standardized test scores has uncovered important differences. In research on the effects of prosocial and antisocial behavior on students’ grades and standardized test scores, Wentzel (1993) notes that social behavior is a stronger predictor of students’ grades than their standardized test scores. This result is enlightening because it indicates important distinctions between the two measures. Although this was not a main focus of her research, the finding leads her to comment that the link between social behavior and different educational assessment measures is not fully understood and should be explored.

Willingham et al. (2002) perform a more thorough analysis of differences in the predictors of grades and test scores from an educational measurement perspective. The article is primarily atheoretical in that it does not present a strong conceptual basis for understanding differences in grades and standardized test scores for individual students.
After analyzing the two measures, the authors conclude that grades and standardized tests do account for different aspects of students’ educational careers and that corrective factors can be used to increase the correlation between the two measures. However, some of the elements that the authors view as error in grades may, from a sociological viewpoint, be seen as important indicators of students’ opportunities for success in other educational settings and occupations. Ultimately, they do recommend using both measures if possible when making high-stakes decisions regarding students because the two measures complement each other in important ways. They do not, however, fully consider the potential benefits of utilizing a measure such as grades that may account for social, behavioral, and academic factors.

A number of educational psychology studies comparing grades and standardized test scores have also looked at the predictive validity of the two measures. The findings consistently indicate that high school grades are a valuable measure for predicting students’ college outcomes. Munro (1981) and Hoffman (2002) reported that high school grades are a better predictor of academic performance in college than standardized test scores from high school. In a comparison of majority and non-majority groups on a college campus, Hoffman and Lowitzki (2005) also found that the predictive validity of each measure varies by group. Non-majority students’ academic achievement was better predicted by high school grades than by standardized test scores while the predictive validity of test scores remained high for majority students. These studies predominantly focus on the implications of these results for college admissions policies and attempts to create campus diversity. They do not offer a strong theoretical basis for understanding the links between the high school assessment measures and future educational outcomes.
Again, this research suggests important differences between grades and standardized test scores that deserve further exploration.

Some sociologists have undertaken more in depth examinations of either standardized tests scores or grades. Research on standardized tests has often focused on disparities in test scores across groups of students. For instance, The Black-White Test Score Gap (Jencks and Phillips 1998) examines standardized tests and IQ tests from a variety of angles to assess what affects students’ scores and how the scores matter for later educational and employment outcomes in an effort to identify the factors that contribute to the persistent racial gap in scores. Jencks (1998; 1972) also offers discussions of the meaning of standardized tests both in terms of what they actually measure and what they are perceived to measure by the public and by researchers.

Rosenbaum (2001) performed one of the few sociological analyses of high school grades as indicators of achievement and presents evidence suggesting that both cognitive and noncognitive factors shape grades and that these factors are important in determining future educational outcomes and earnings. \(^1\) Specifically, he reports that grades account for noncognitive factors including school attendance, discipline problems, and extracurricular leadership. He also finds that noncognitive factors predict educational attainment and earnings nine years after high school and grades mediate some of these effects.

Rosenbaum notes that many employers do not consider grades to be accurate measures of skills that are relevant for the work environment. However, the noncognitive

\(^1\) Farkas et al. (1990) report similar findings regarding the effects of noncognitive factors on grades for middle school students.
factors measured in grades reflect behaviors and abilities that are important in many work settings. The established relationship between grades and later earnings provides evidence for the connection between the determinants of grades and occupational success. These findings provide an initial basis for developing a sociological understanding of what grades measure and how they can most effectively be used in research.

All of the research described above is important in that it helps to set a foundation for developing a theory regarding grades as a form of educational assessment. The research indicates that grades and standardized test scores are two separate and distinct indicators of academic outcomes and their use may lead to different results and perhaps invalidate comparisons of research findings that employ different measures. However, none of it fully considers the meanings of grades and the potential ramifications of their use in research. This research will address that gap by undertaking an examination of grades from a sociological viewpoint that can further inform the understanding of grades, particularly in comparison to standardized test scores. The application of a uniquely sociological viewpoint to the issue may result in better research and policy decisions that account for the social context of education.
CHAPTER 3:
THEORETICAL FOUNDATION FOR UNDERSTANDING GRADES
AS A MEASURE OF EDUCATIONAL OUTCOMES

Schools are significant societal institutions that most residents of the United States will have contact with throughout their lives. Like most organizations, schools are characterized by a variety of divisions tasked with different functions, a hierarchy of power, and the ultimate goal of producing a specific product. In the case of schools, classrooms represent divisions of the organization and each class is designed to meet a specific purpose. That purpose may be educating students of a certain age or ability level or providing instruction on a specific academic topic. Within the classroom, teachers have authority over students in that they organize and plan instruction and students are expected to meet academic and behavioral demands. The organization is designed to facilitate the production of well-educated and socially-capable citizens. These two mandates are consistent across schools despite the fact that, because of organizational variations, schools may differ in how they achieve the goals or define the final product.

Given the ubiquity of educational organizations and the fundamental nature of their goals, assessing the success of schools is a primary concern. Currently, the predominant mode of evaluation is standardized test scores. Test scores gauge progress toward achieving the first goal of schools outlined above: creating well-educated students. They do not, however, offer a full accounting of the mandate to produce
socialized students. In a nominal way, they measure behaviors that are indicative of socialization in that most students need to pay attention in class to gain knowledge and be willing to put effort into a test in order to earn a decent score. This indicates an orientation toward academic achievement. Beyond that, standardized tests do not directly measure students’ success or failure at adopting the norms of the school and classroom.

In addition, standardized test scores are not a high-impact form of academic feedback for most students. Students, teachers, and parents may consider test scores invalid because they do not measure students’ progress or effort, because the student had a bad day when he or she took the test, or because they do not believe the test score accurately assessed knowledge transmitted within the classroom. Opposition to standardized testing and its use in high-stakes decisions such as graduation and funding is illustrated by a teacher who received a two week suspension after refusing to administer a standardized test to his students (Shaw 2008). The teacher claimed that he did what he felt was best for his students because the tests do not offer appropriate evaluations of students’ abilities and efforts. He called for efforts to develop more appropriate methods of assessing students’ academic outcomes. While the extreme nature of his refusal is rare, it illustrates a growing disillusionment among educators with the continued reliance on standardized test scores to evaluate students and schools.

In addition, the value of standardized test scores as feedback may be limited because students, teachers, and parents are not informed of the results of tests – particularly those administered for research purposes. If they are made aware of students’ results, they may not be able to interpret the score because they are not familiar with the metric or standards for the test. These factors limit the ability of standardized
tests to act as meaningful academic feedback that can shape students’ educational trajectories.

Grades offer a more holistic evaluation than standardized test scores of students’ successes and failures at meeting the organizational mandates of the educational process. They also provide an intuitive method for students and significant others to understand students’ academic progress and outcomes. In order to better understand these characteristics of grades, the theory outlined below describes the production and negotiation of grades within a high school classroom setting. While factors outside of the classroom, such as family and social activities, play a role in determining academic outcomes, the theory does not directly address those elements. Any effects they may have are considered inputs to the classroom setting in terms of how they affect students’ academic performance and behaviors. The theory highlights both the intellectual and social aspects of grades and illustrates the possible rewards of using grades as a measure of educational assessment.

3.1 Organizational Mandates

3.1.1 Instruction

Classroom instruction is the main method for achieving one of the primary goals of schools – educated students. Teachers’ methods of instruction and expectations for student achievement can vary across classrooms. The primary, organizationally-sanctioned, reason for this is ability grouping (Hallinan and Ellison 2008). Ability grouping is designed to place students in classes according to their academic capabilities,
to make instruction easier for teachers, and to expose students to an appropriate level of instruction. This type of variation is more common in high schools where students will be grouped in separate classrooms rather than in groups within a classroom or pull-out classes, as is common in elementary schools. The stated function of this grouping is to provide appropriate instruction to students with varying abilities and learning styles. Because the students in different groups ostensibly have different abilities, teachers will alter their instructional content and pedagogy accordingly and have differing expectations for students’ academic performance.

Instruction may also vary across classrooms as an outcome of student-teacher interactions. Within classrooms, the organization and its authority relations are clear. The teacher has control or desires control and the students are expected to be relatively passive (Jackson 1968). However, teachers are outnumbered by students and this situation provides the students some degree of power within the classroom. This type of asymmetrical power is also found in the average workplace. Bridges (1970) posits several ways in which workers who are not in a position of authority can affect the actions of their superiors. Some of the tactics Bridges lists for achieving these ends include forming coalitions, deceit, flattery, and threats. In these situations, the person or group in power may knowingly, unknowingly, or unwillingly submit to the will of subordinates. McFarland (2001) has found that some of these same forces are in effect in the classroom, allowing students to manipulate and negotiate with their teachers in a variety of ways. This indicates that student behavior can affect instruction by facilitating or impeding instruction or influencing teachers to change instructional plans.
The effects of these organizational realities of schools on instruction shape results on grades. The type of knowledge and level of academic performance necessary to earn a specific grade will vary across classrooms because individual teachers hold different expectations for students and set different standards for grades. In addition, students’ academic accomplishments will be compared within the classroom and their outcomes may be determined in part by this comparison. Farkas et al. (1990) found that teachers tend to assign a high grade to a student if the student performs better academically than his or her classroom peers. However, the same student might not receive as high of a grade for the same level of achievement if classmates equal his or her academic performance. Therefore, classroom context can affect grades through academic expectations and comparisons of students.

Grades also reflect power relations in schools and classrooms by registering variation in how students and teachers interact based on the social dynamics described above. They are a negotiated outcome of day-to-day interactions over several months between the teacher and individual students and/or groups of students. Students may accept the grades they receive or alter their academic performance or behavior to change their grades. They may also find other methods; such as protest, negotiation, or threat; to convince teachers to change instructional methods or grades.

3.1.2 Socialization

The other primary organizational mandate of schools is to socialize students. Dreeben (1968) highlights the important role that schools play in teaching children to function outside of their families and in preparing them for adulthood and work. The
author outlines educational goals that are essential for students to succeed but not directly related to the transmission of subject knowledge.

Dreeben (1968) focuses on the similarities in socialization processes across schools, indicating that there is a common goal in the United States in terms of how students should be socialized. Other authors, though, have highlighted how the process may differ across schools (Bowles and Gintis 1976; Anyon 1980). Their research indicates that students are socialized differently based on their social class positions, with lower-income students taught to follow instructions and obey authority and higher-income students taught to think critically and make decisions independently. This indicates that differences in how students are socialized may be a function of organizational variations within schools. More recently, though, research by Lareau (2000) found that these differences in students’ socialization outcomes result more from students’ home life than from their experiences in classrooms. From this perspective, differences in behavior based on social class are a type of input to the classroom rather than an outcome of classroom processes.

Despite the fact that students enter schools with varying types of preparation and social skills because of their different backgrounds, the majority of teachers desire all students to achieve a similar baseline level of socialization (Dreeben 1968; Jackson 1968). This minimum level will allow their classes to proceed smoothly and prepare students for life outside of the family and, eventually, outside of classrooms. Because these aspects of socialization are necessary to ensure that classrooms function and that instruction can occur, they will be relatively constant across all schools in the United
States and not subject to the organizational variations, such as ability grouping, that create differing instructional and academic expectations across classrooms and schools.

Dreeben (1968) outlines four norms that students need to adopt in order to differentiate from their family unit and prepare for future occupations and civic life: independence, achievement, universalism, and specificity. The first two – independence and achievement – are manifest and directly related both to educational outcomes and future occupational success. The second two – universalism and specificity – are latent and related more to self-concept and civic engagement. This research will focus primarily on the norms that are directly addressed within the classroom and that are most closely related to academic outcomes – independence and achievement.

Students’ acceptance of the norms of independence and achievement contributes to behaviors that are commonly considered to be educationally-appropriate and conducive to classroom learning. These include valuing academic success, accepting and taking account of evaluations, following instructions, and seeking to improve academic assessments (see Jackson’s work on “crowds, praise, and power” (1968) for a similar discussion). Well-socialized students, as defined by these criteria, are likely to receive positive feedback from teachers – both informally through classroom interactions and formally through grades.

3.2 Intersection of Instruction and Socialization

Implicit in the norms of independence and achievement is a desire to achieve the other main goal of schools – creating educated students. Students who are well-socialized should value education and pursue academic success. Their concepts of
academic success will vary, though, because these concepts are communicated through classroom experiences which are subject to the organizational variations – ability grouping and classroom interactions – described in the section on instruction. Well-socialized students who are in different schools or classrooms will develop discrepant ideas of what constitutes academic success, yet they will all want to achieve the broad goal of academic success. By pursuing this goal, students facilitate instruction and create positive student-teacher interactions because they are working in concert with the teacher to achieve academic success as defined within their classrooms. Students who fail to achieve the goal of socialization will likely have negative educational experiences both socially and academically because they do not behave in ways that are conducive to classroom learning.

3.3 Transmission of Organizational Mandates

Within schools, the organizational mandates of instruction and socialization are transmitted to students in a variety of ways. One medium of transmission is the organizational structure of the school itself. Students are divided by age, grade, and subject; clear authority relations are established; a schedule is in place that shapes daily life in the school and classroom. These and other aspects of school structure implicitly communicate the goals and methods of education to students.

Another primary conduit of information regarding organizational mandates is the classroom teacher. Teachers explicitly communicate the goals of instruction and socialization to students within the classroom. They demand certain behaviors and levels
of performance from students and distribute rewards and punishments based on their perceptions of students’ progress toward the goals.

Based on the information they receive from the organization and teachers, students come to understand the goals of the education system. They can then choose, to some degree, whether to accept or reject these goals. Within a high school setting, most of the students will already be familiar with the goals and methods to achieve them based on earlier academic experiences. Although academic goals will vary from classroom to classroom, students will be aware that, in order to succeed, they must understand the teachers’ academic and behavioral standards and work to reach them. Students can then choose, to a degree, if they want to commit the effort to working to achieve the goals.

Students’ levels of dedication to achieving the goals of education are determined by numerous variables that shape their academic lives – including, but not limited to, their innate abilities, the tenor of classroom interactions, and their long-term plans. Whatever the determinants of students’ commitment to the goals of instruction and socialization, their ability and commitment to meeting the demands of the education system will be reflected in their academic performance and classroom behavior.

3.4 Feedback

3.4.1 Teacher Evaluations of Students

Teachers are the primary judges of students’ success or failure in achieving the goals of education. A large part of their jobs is to evaluate academic performance and report on it. They have the responsibility of assigning students’ grades based on
requirements they have mandated for the course and their understandings of students’ academic performance and behavior.

Teachers provide feedback on students’ progress toward classroom goals in several ways other than through the final grade. Throughout daily life in the classroom, they praise and admonish students for their work and behavior. Teachers also provide formal feedback to students on assignments and tests. Frequent feedback from teachers reminds students of requirements and offers them an opportunity to maintain or change their performance and behavior based on the criteria presented by the teacher. This process culminates in a final grade for the quarter or semester. The grade represents the end product of negotiations regarding work and behavior between the teacher and student.

Classroom behavior may not be perceived of as a primary component of high school grades. However, there are several behaviors related to the norms of achievement and independence that are regularly factored into grades. These include whether or not students complete homework or attend class. Also, many teachers include a participation element in grades. Participation can encompass a number of behaviors including student willingness to talk during discussion, attentiveness during lectures, and a variety of other requirements imposed by the teacher. In addition, students who do not value achievement and independence or students who behave inappropriately in the classroom are less likely than well-socialized peers to benefit from the instruction that is provided. They may fail to pay attention to instruction, neglect to do work assigned for class, or engage in inappropriate work behavior (such as cheating or plagiarism) that results in grade reductions.
As was mentioned earlier, the academic requirements for grades may vary across classrooms. These variations are designed to meet the academic needs of students. However, behavioral requirements imposed by teachers, such as those described in the previous paragraph, are expected to be relatively similar. This is both because they are necessary to ensure that academic work can be accomplished, as described in the section on socialization, and because of the nature of the teaching profession. Lortie (1975) found that teaching is an inherently conservative occupation in which teachers tend to share uniform practices and expectations in the classroom. These viewpoints are entrenched because of the methods for training and professionalizing teachers. The outcome is that teachers are likely to value similar behaviors and indications of academic commitment from students.

For the reasons noted above, classroom grades represent teachers’ evaluations of both students’ academic performance and their classroom behavior. The standards for academic performance will vary across classrooms because of age-grading, ability grouping, subject matter, and context. Therefore, a high grade in one class may represent a level of academic performance that would earn a mediocre grade in another classroom. However, student orientation toward academic success (as a result of socialization) will factor into grades across all classrooms in comparable ways because it is a universal element of the classroom environment and teachers are likely to value the behaviors similarly across courses.

3.4.2 The Impact of Grades on Students

Grades are not just an outcome of, but also an input to, educational processes in the form of feedback on educational progress. Feedback plays an important role in
education. Because of socialization, most students come to see the institution of education as legitimate and adopt the goal of achieving academic success (Meyer 1977). The same is true of teachers and parents, many of whom have been educated and socialized in a similar system. Because they accept the broad goal of becoming educated, students, parents, and teachers will value and take into consideration feedback that indicates whether or not students are achieving those goals. Students who are not socialized to their particular educational settings are less likely to consider the feedback because they will doubt its validity or simply not be interested in whether or not they are making progress. However, the feedback may still be important to parents and teachers and their reactions to the feedback will have effects on the students.

Grades are a particularly salient form of feedback. Students and their significant others receive grades frequently in the form of assignments and test results and they receive report cards at regular intervals. Some schools have instituted programs requiring parents to review assignments or pick up report cards to ensure that parents are aware of students’ performance. Therefore, from a social perspective, grades will offer a readily-available and easy to understand form of feedback to those who are interested in students’ progress (Entwisle, et al. 2000).

Given their visibility and frequency, grades will influence those who receive them through reactions to the grades. And, despite the organizational variations in expectations and grading practices across schools and classrooms that were described above, it is likely that students, teachers, and parents will respond to grades in similar ways. The main concern will be that students meet the goals of their particular
organizational setting which represents the wider goal of becoming well-educated, despite the fact that the definition of success may vary from classroom to classroom.

In this process, students begin to view themselves as the types of students indicated by the grades received. Teachers rank students through grades and react to students according to these categorizations. Parents view grades and react to them at face value and treat their child accordingly. Peers, whether or not they are explicitly aware of other students’ grades, also react to the hierarchy created within the classroom. Through these perceptions, grades and the meanings attached to them are reified and lead to social consequences regarding students’ perceptions of their academic abilities and status as a students.

Responses to educational assessments impact students’ educational futures (Rosenthal and Jacobson 2003). Students are labeled as a certain kind of learner based on the assessments. This can perpetuate students’ current behavior or create a self-fulfilling prophecy wherein they meet the expectations reflected by their grades. This response serves as a further input to the educational process and has independent impacts on academic outcomes through the reinforcements of or changes in behavior of the students and significant others. The trajectory that this establishes for students may play a role in determining future academic and occupational outcomes because the feedback influences levels of dedication and commitment to education and academic achievements. Students’ perceptions of their academic abilities, as based on their grades, can also have impacts beyond the high school classroom because they may influence their decisions regarding whether or not to pursue more education.
3.5 Conclusion

The preceding discussion highlights the variability of grading practices across educational locations and the influence of factors other than knowledge gained on grades. This indicates that grades are a more subjective assessment than standardized test scores. For this reason, in situations where there is a discrepancy between grades and standardized test scores, many people believe that the results on standardized tests will be more valid than grades in terms of predicting future outcomes. This belief is illustrated in the conflict regarding Texas’ education policy described in Chapter One. In addition, Rosenbaum (2001) found that employers frequently report that they do not trust information regarding potential employees that is furnished by schools. They dismiss grades as indicators of relatively meaningless knowledge and do not believe grades measure skills that would be applicable in the workplace.

This may not be the case, however, because knowledge levels, as represented by standardized test results, are not the only requirement for success in education and occupations. People must have a willingness and ability to adapt to organizational norms and have a minimum level of social acumen in order to succeed in most organizations. Grades reflect these types of soft skills whereas standardized test scores cannot measure them. Standardized test scores primarily reflect hard skills, such as subject-specific knowledge and the ability to complete related tasks.

The differences noted above indicate that grades may have certain advantages compared to test scores when predicting outcomes. Grades are determined in a situation that approximates settings in which students will later need to be successful – such as college classrooms and workplaces. Therefore, grades account for factors that will
impact outcomes in those similar settings. Test scores are created in more artificial settings that are designed to make soft skills irrelevant or to not measure them.

By considering the social context in which education takes place and the social effects of educational assessments, it becomes clear that an assessment measure such as standardized test scores may not offer a full summary of educational experiences or function as the best predictor of future educational transitions. However, grades, because they account for social processes, may be weak indicators of knowledge gains. This does not mean, though, that neither measure should be used in research. What it indicates is that the value of each measure should be carefully considered in research designs in order to ensure that the most appropriate measure is utilized. The following analyses will contribute to this goal by offering an evaluation of grades and examining their content and meaning.
This section presents the results of analyses of data reported by teachers as it relates to students’ grades. Teachers’ reports of students’ activities and behavior are used as predictors of grades instead of students’ reports of their own activities in order to maintain consistency with the theoretical basis of the research, as outlined above. The theory indicates that students’ instructional placements, social interactions, and behaviors will have impacts on their grades. These impacts will operate primarily through teachers’ perceptions of students because teachers witness and interpret students’ actions and behaviors and ultimately assign grades. While teachers’ understandings of students may be incomplete or inaccurate, those understandings will be part of what teachers use to determine grades – whether implicitly or explicitly.

The analyses use data from the National Education Longitudinal Study of 1988 (NELS:88). The data were gathered by the National Center for Education Statistics (NCES) beginning in 1988. The sample includes over 12,000 students who were sampled from schools during their eighth-grade year. Schools were chosen randomly from a file of the 40,000 public and private schools enrolling eighth graders in 1988. Respondents were re-surveyed in 1990, 1992, 1994, and 2000. Respondents’ teachers were surveyed in 1988, 1990, and 1992.
The NELS:88 database is well-suited for the planned quantitative analyses because of the variety of information that it contains. It provides important demographic data on student respondents, such as gender, race, and parents’ socioeconomic status. The teacher surveys provide information regarding teachers’ perceptions of the specific students who are participants in the survey. The dataset also contains students’ grades as taken directly from high school transcripts.

Students are included in the analyses if their transcript-indicated math or English grades are available for their sophomore year (1990). The sample is further narrowed by the availability of teacher surveys in that year. Students are only included if their math or English teachers were interviewed in the 1990 wave of data collection. Table 4.1 contains descriptive statistics for all variables used in the analyses.

Two dependent variables will be used in the analyses. These are students’ math grades in 1990 and their English grades in 1990. These variables are available in the restricted-use version of NELS:88 and were derived from transcript data provided by schools to the NCES. The two subjects are used as dependent variables in separate, but similar, models to examine similarities and differences in the factors that impact grades across the two subjects. Although curricular differences are not a focus of the study, noting differences in grading practices across the two types of courses may reveal variations that are important from a theoretical perspective.

Students’ sophomore (1990) grades are used in the analyses for several reasons. First, transcript grades are not available for the 1988 wave of the survey when the students were in 8th grade. In the students’ senior year (1992), no English teachers were
### TABLE 4.1

**DESCRIPTIVE STATISTICS**

**FOR VARIABLES IN THE ANALYSES**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DEPENDENT VARIABLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10th Grade Math Grade</td>
<td>6.760</td>
<td>(3.346)</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>10th Grade English Grade</td>
<td>7.279</td>
<td>(3.229)</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td><strong>CONTROL VARIABLES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>.523</td>
<td>.521</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>.075</td>
<td>.089</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>.100</td>
<td>.109</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Asian</td>
<td>.071</td>
<td>.070</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Native American</td>
<td>.008</td>
<td>.009</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Socioeconomic Status (Math)</td>
<td>.038</td>
<td>(.771)</td>
<td>-2.272</td>
<td>2.753</td>
</tr>
<tr>
<td>Socioeconomic Status (English)</td>
<td>.025</td>
<td>(.770)</td>
<td>-2.272</td>
<td>2.541</td>
</tr>
<tr>
<td>10th Grade Math Test Score</td>
<td>46.319</td>
<td>(13.341)</td>
<td>16.99</td>
<td>72.76</td>
</tr>
<tr>
<td>10th Grade English Test Score</td>
<td>31.743</td>
<td>(9.521)</td>
<td>10.47</td>
<td>48.80</td>
</tr>
<tr>
<td>8th Grade Grade</td>
<td>4.089</td>
<td>(.930)</td>
<td>4.079</td>
<td>(.879)</td>
</tr>
<tr>
<td><strong>INSTRUCTIONAL EXPERIENCES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honors level class</td>
<td>.137</td>
<td>.200</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Academic level class</td>
<td>.568</td>
<td>.371</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Vocational class</td>
<td>.022</td>
<td>.015</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Other type of class</td>
<td>.018</td>
<td>.030</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>SOCIAL INTERACTIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student talks with teacher outside of class</td>
<td>.334</td>
<td>.401</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Student relates well to others</td>
<td>.905</td>
<td>.895</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>STUDENT BEHAVIOR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Classroom Engagement</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often student does homework</td>
<td>4.023</td>
<td>(.932)</td>
<td>4.042</td>
<td>(.936)</td>
</tr>
<tr>
<td>How often student is absent</td>
<td>2.183</td>
<td>(.654)</td>
<td>2.192</td>
<td>(.661)</td>
</tr>
<tr>
<td>How often student is tardy</td>
<td>1.525</td>
<td>(.708)</td>
<td>1.562</td>
<td>(.738)</td>
</tr>
<tr>
<td><em>Classroom Demeanor</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often student is attentive</td>
<td>3.995</td>
<td>(.828)</td>
<td>3.955</td>
<td>(.822)</td>
</tr>
<tr>
<td>Student is passive</td>
<td>.080</td>
<td>.089</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Student is disruptive in class</td>
<td>1.570</td>
<td>(.814)</td>
<td>1.618</td>
<td>(.868)</td>
</tr>
<tr>
<td>Student usually works hard</td>
<td>.696</td>
<td>.682</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>EDUCATIONAL FUTURE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student probably won't go to college</td>
<td>.226</td>
<td>.246</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Don't know if student will go to college</td>
<td>.167</td>
<td>.150</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>3547</td>
<td>4405</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SOURCE:** National Education Longitudinal Study of 1988

**NOTE:** Standard deviations for all non-dichotomous variables are presented in parentheses.
interviewed. Only data from math teachers would be available which would prevent a comparison of grades across the two subjects. In addition, many schools do not require that students take a math course during their senior year so the limited number of students who did take a math course and whose teachers were surveyed would be a selective group that is not representative of high school students.

The sophomore year grades do not have the drawbacks listed above. The NCES interviewed both math and English teachers in 1990 so a comparison of the two subjects is possible. In addition, most students are required to take English and math courses at that point in their high school careers so there should not be a bias in the type of student included in the analyses.

The tactics used for coding the grades warrant mention because the variable is presented in a different format than most other data in NELS:88. The transcript-indicated grades are provided in one variable which contains all grades for all available students for each year of high school. To narrow this down, grades from 1990 were selected. All grades that were not math or English were then removed, depending on which variable was being created. This was done by selecting for the Carnegie course codes that correspond with each subject from the variable provided in NELS:88. The next step was selecting for grades that students received for full year courses or in the second semester, third trimester, or fourth quarter of the year. This was done because teachers were interviewed in the spring semester so the final grade of the year is most likely to have been assigned by the teacher who was interviewed. After this step, a small number of students still had multiple grades for a subject. This may have been because there was an error in the transcript coding or the students were taking more than one course in a
particular subject. This was dealt with by deleting all grades except the first one that was listed for each student.

The analyses include common control variables for educational research. These are students’ reports of their gender, race/ethnicity, and a composite variable that indicates socioeconomic status. In addition, the score from a standardized test, administered by NCES, is included as a control for knowledge. The sophomore year (1990) tests are utilized in these analyses because they represent students’ knowledge in the relevant subject area close to the time that the grade used as the dependent variable was assigned. Finally, students’ eighth grade (1988) self-reports of their average grades in the subject being analyzed for sixth through eighth grade are included as a control for prior classroom outcomes. Transcript-indicated grades are not available for 1988, but the self-reported grades provide an estimation of earlier classroom assessments and account for students’ perceptions of feedback they have received in the subject.

All of the explanatory variables in the current analyses came from the 1990 teacher surveys of either math or English teachers. The variables are divided into four subsets of teacher reports – instructional experiences, social interactions, student behavior, and educational future – chosen to represent aspects of the theory presented in Chapter Three. Each variable will be described in detail in the appropriate section below.

The explanatory variables had a limited amount of missing data. The “ice” command in Stata was utilized to impute values for the missing data in the independent variables. This method takes existing data into account to create values for missing data. The imputations were restricted to students who had a teacher survey available. This limits the sample size but maintains the integrity of the data by not permitting teacher
data to be created if no teacher survey was associated with the student. No data was imputed for the dependent variables (transcript-indicated grades). Five datasets were created through the imputation process. For this reason, the “micombine” command is used with all analyses to create an average of the results of each dataset for the specified models. Ordinary least squares regression was used for all analyses. This is appropriate because the dependent variables are continuous – coded from 1 (F) to 13 (A+).

Prior to conducting multivariate analyses to examine the predictors of grades, the correlations between test scores and grades in the sample were analyzed. English test scores and grades have a correlation of .451 and the correlation for math scores and grades is .502. These results indicate that a strong relationship exists between the two measures. However, they also show that the two measures differ in important ways.

In order to make the relationship between individual students’ grades and test scores clearer, graphs are created representing the distribution of test scores by students’ grades for each subject. The graphs are presented as Figure 4.1 and Figure 4.2. Within the box plots, the line in the middle of the box represents the median test score for students with the grade represented on the x axis. The lines indicate that test scores follow a predictable pattern for both academic subjects. Median test scores increase as students’ grades increase. The boxes follow similar patterns in that students with lower grades tend to earn test scores that are lower than those received by students with higher grades.

While the graphs show a generally consistent pattern, the lines representing the adjacent values in the graph – those students who are at the upper and lower bounds of the distribution – illustrate that students within each grade category earn a wide range of
Figure 4.1: The graph presents the distribution of math test scores by grades. The data source is the National Education Longitudinal Study of 1988. The graph was created using Stata 10.

test scores. Although scores tend to cluster in predictable ways, many students are receiving test scores that are not matched with their grades in that the score is higher or lower than might be expected. While the frequent convergence of the two measures illustrates that they can provide similar assessments of students, the mismatches indicate that the two measures may provide evaluations of different skills and abilities. The following analyses and discussion will examine what might account for these differences.

Based on the theory presented earlier, it is hypothesized that regression analyses will illustrate that most of the explanatory variables in these models have significant relationships with grades. The first group of variables represents teachers’ reports of students’ ability group placement. The relationships between these variables and grades
Figure 4.2: The graph presents the distribution of English test scores by grades. The data source is the National Education Longitudinal Study of 1988. The graph was created using Stata 10.

are difficult to predict because the nature of the contextual aspect of grades is not yet fully understood. As was indicated previously, it is believed that students who have high-quality instructional experiences will gain more knowledge than students who receive low-quality instruction. However, grades provide a comparison of students within a classroom setting. Therefore, students who receive the same instruction may receive different grades based on their abilities, their engagement with opportunities to learn, and their dedication to classroom goals. For this reason, grades may vary more systematically within classrooms than across classrooms. Given the uncertainty regarding the assignment and distribution of grades, no firm hypotheses will be offered at
this point although it is predicted that ability group assignment will be related to grades because the assignments are powerful determinants of instructional experiences.

For the remaining three sets of variables, it is hypothesized that, if teachers report positive results for students, students will receive higher grades. Positive results are those which indicate that students are committed to classroom norms and educational goals. The first group of behavior variables includes teachers’ perceptions of students’ social interactions. It is predicted that students who have frequent and positive interactions will earn higher grades than their less engaged peers because the interactions indicate a positive engagement with members of the classroom community.

The next set of variables represents teachers’ perceptions of students’ behavior and is divided into two subsets. The first subset is classroom engagement which represents behaviors that teachers can objectively measure such as homework completion. These variables are hypothesized to be strongly related to grades because they indicate students’ engagement with opportunities to learn and will be recorded and tracked by teachers to be included in grades. The second subset is classroom demeanor which represents behaviors that teachers can perceive but not measure directly such as whether or not students are attentive or hard working. The behaviors represented in these variables are also hypothesized to be related to grades but not as strongly as the behaviors in the classroom engagement section because they cannot be measured directly and are less likely to be recorded by the teacher.

The final set of variables includes teachers’ reports of students’ educational futures. Because these reports represent teachers’ beliefs regarding students’
commitment to education, it is hypothesized that they will be strongly related to grades because they are representative of engagement with educational opportunities.

The results discussed below are presented in Table 4.2 and interpreted in terms of the hypotheses presented above. The first row in the table presents the results with math grades as the dependent variable and the second row uses English grades as the dependent variable.

The control variables in these models are included not for substantive results, but to ensure that systematic differences across students do not cloud the results for the explanatory variables. For that reason, the results for the control variables will not be discussed at length. However, it is important to note that the results indicate that persistent differences exist in assigned grades across demographic groups. This finding deserves further exploration as it mirrors durable gaps in test scores that also exist between groups. Future research should address the causes of these gaps and explore how they may or may not reflect similar processes as those that create disparities in test scores.

It is also important to emphasize the presence of test scores and prior grades as control variables. Test scores are conceptualized to provide an approximation of student subject knowledge around the time the 1990 grade was assigned. Students’ self-reports of eighth grade grades represent prior classroom experiences and educational feedback. As expected, both of these factors have strong and positive relationships with the dependent variable in each model. More important, though, is that these variables are held constant in the model and permit a less ambiguous exploration of the relationship between grades and the explanatory variables.
### Table 4.2

**Regression of Math and English Grades on Teachers’ Reports of Students’ Instructional Experiences, Social Interactions, Behaviors, and Educational Futures**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Female</strong></td>
<td>-.137 (.082)</td>
<td>.209** (.082)</td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td>-.640*** (.156)</td>
<td>-.983*** (.122)</td>
</tr>
<tr>
<td><strong>Hispanic</strong></td>
<td>-.459** (.136)</td>
<td>-.046 (.112)</td>
</tr>
<tr>
<td><strong>Asian</strong></td>
<td>-.015 (.156)</td>
<td>.413** (.135)</td>
</tr>
<tr>
<td><strong>Native American</strong></td>
<td>-.290 (.445)</td>
<td>.093 (.356)</td>
</tr>
<tr>
<td><strong>Socioeconomic Status</strong></td>
<td>.126* (.059)</td>
<td>.189*** (.051)</td>
</tr>
<tr>
<td><strong>10th Grade Test Score</strong></td>
<td>.070*** (.004)</td>
<td>.057*** (.004)</td>
</tr>
<tr>
<td><strong>8th Grade Grade</strong></td>
<td>.276*** (.048)</td>
<td>.374*** (.044)</td>
</tr>
</tbody>
</table>

**INSTRUCTIONAL EXPERIENCES**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honors level class</td>
<td>-.693*** (.154)</td>
<td>-.050 (.107)</td>
</tr>
<tr>
<td>Academic level class</td>
<td>-.716*** (.104)</td>
<td>-.138 (.085)</td>
</tr>
<tr>
<td>Vocational class</td>
<td>.703* (.287)</td>
<td>-.133 (.280)</td>
</tr>
<tr>
<td>Other type of class</td>
<td>.429 (.306)</td>
<td>.584** (.204)</td>
</tr>
</tbody>
</table>

**SOCIAL INTERACTIONS**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student talks with teacher outside of class</td>
<td>-.002 (.086)</td>
<td>.041 (.071)</td>
</tr>
<tr>
<td>Student relates well to others</td>
<td>-.267 (.152)</td>
<td>-.142 (.122)</td>
</tr>
</tbody>
</table>

**STUDENT BEHAVIOR**

*Classroom Engagement*

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often student does homework</td>
<td>.953*** (.066)</td>
<td>.932*** (.058)</td>
</tr>
<tr>
<td>How often student is absent</td>
<td>-.253*** (.068)</td>
<td>-.366*** (.059)</td>
</tr>
<tr>
<td>How often student is tardy</td>
<td>-.057 (.063)</td>
<td>.026 (.052)</td>
</tr>
</tbody>
</table>

*Classroom Demeanor*

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often student is attentive</td>
<td>.572*** (.075)</td>
<td>.177** (.062)</td>
</tr>
<tr>
<td>Student is passive</td>
<td>-.131 (.164)</td>
<td>-.088 (.133)</td>
</tr>
<tr>
<td>Student is disruptive in class</td>
<td>.267*** (.060)</td>
<td>.097* (.048)</td>
</tr>
<tr>
<td>Student usually works hard</td>
<td>1.046*** (.124)</td>
<td>1.338*** (.101)</td>
</tr>
</tbody>
</table>

**EDUCATIONAL FUTURE**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student probably won’t go to college</td>
<td>-.634*** (.131)</td>
<td>-1.057*** (.110)</td>
</tr>
<tr>
<td>Don’t know if student will go to college</td>
<td>-.994*** (.119)</td>
<td>-1.621*** (.108)</td>
</tr>
</tbody>
</table>

N=3547 N=4405

*Note: Standard deviations for all coefficients are presented in parentheses.*

*Note: *p<.05; **p<.01; ***p<.001*
In addition to the control variables shown, models were calculated which included administrators’ reports of school characteristics as controls. The reports included the percentage of students in the school taking remedial math or English classes, the percentage of students in advanced placement courses, and the percentage of graduates from the previous year who attended a four-year college. These variables were included in the models to determine if the school context in which the student earned the grade impacts the outcome. For instance, students in schools with a high percentage of advanced placement students may be less likely to earn high grades because they are compared to more academically-capable students than in schools with fewer students in rigorous courses. The results for these variables are not shown because the coefficients were not significant and did not alter the coefficients for other variables.

The first set of explanatory variables represents students’ instructional experiences by examining differences in grades across ability groups. Teachers were asked to report the track of the class to which the student who was the subject of the survey was assigned. It is assumed that teacher reports correspond closely to official school classifications of the courses and, therefore, represent a valid measure of students’ instructional placements. The teachers’ reports were used to create four dummy variables for ability group: honors, academic, vocational, and other. The omitted category is “general.” It is impossible to say with certainty what courses of study “other” represents but it may include a preponderance of students who participated in remedial classes, special education, or were English Language Learners.

Unlike the other subsets of variables in the models which are discussed below, significant differences appear between math and English grades in terms of the impact of
ability group experiences. For that reason, I will discuss the two sets of results separately, beginning with math.

The analyses indicate that ability group placement is strongly related to math grades. Students in honors and academic courses receive significantly lower grades than students in general courses. Students in vocational courses receive slightly higher grades than those in general courses. There is no effect for the “other” group in terms of math grades. The possible variability of courses within this group, as noted above, may account for the lack of a consistent relationship with grades.

The results regarding the relationship between ability groups and math grades appear reasonable when taking into consideration that test scores are held constant in the analyses. Given that math courses are often cumulative in nature, building on knowledge from prior courses, students who take an advanced level course would be expected to have some knowledge gaps compared to peers in the same class who received higher test scores. It is also to be expected that if the students with lower scores enrolled in a less academically challenging math course, they would have found a more suitable match between their knowledge and the curriculum and would have received a higher grade.

A discrepancy between knowledge and course content may reduce math grades because teachers explicitly or implicitly rank students within the classroom. Teachers may directly compare students in the course and assign grades on a scale that ranks students according to their classmates’ performance. More likely, however, is that teachers implicitly rank students by offering instruction, assigning materials, and designing assessments based on the average level of achievement in the classroom. This tactic would ease instruction by allowing teachers to focus on students at one level of
achievement rather than forcing them to modify instruction to meet the needs of each student. While the practice of ability grouping is designed to, ideally, provide teachers with a homogeneous classroom in terms of academic achievement, improper placements do occur. Students who receive an inappropriate group assignment will be disadvantaged through comparison to their peers and are likely to receive correspondingly lower grades.

An additional effect of an inappropriate ability group assignment may occur if students become frustrated with their inability to perform at the academic level demanded by the teacher and cease to put effort into their coursework. However, this possibility is generally accounted for by including student behavior variables in these models which will be discussed later. Therefore, it appears that the effect of ability group placement on math grades is related primarily to knowledge gaps and mismatches between student ability and course demands.

The effect of vocational math placement on grades is positive compared to general courses, but smaller than the negative effect associated with academic or honors courses. This may be because content coverage in vocational and general classes is fundamentally different in that vocational courses focus on types of math that are technically oriented and do not require the same initial knowledge base and cumulative knowledge development of the academic courses. For that reason, comparisons across vocational and academic math courses may not measure the same types of curricula. Additionally, it is important to note that few teachers reported teaching vocational courses and the sample is small which reduces statistical reliability.

The results for English grades are noticeably different than the results for math grades. Students in honors, academic, and vocational English courses do not receive
significantly different grades than students in general courses. Only students in “other” English courses have consistent differences in grades compared to students in general courses. Given the uncertain nature of the “other” ability group category, it is difficult to develop a theoretical understanding of the observed difference.

Of more interest than the one significant effect is the lack of significance for the three other variables representing ability group placement. The results for math ability groups revealed a pattern that fit a narrative of increasing academic demands and the ability (or inability) of students with various levels of knowledge to meet those demands. The results for English ability groups defy that narrative and indicate that curricular differences exist across the two academic domains. To remain consistent with the interpretation of differences in math grades across ability groups, two explanations for the lack of such an effect across English ability groups are offered below, one or both of which may account for the differences between courses.

The first explanation is that content on standardized tests of English ability is not tightly linked to the curriculum of English courses upon which grades are based. This lack of association would prevent any consistent relationship between scores on English tests and English grades because the two forms of assessment would not measure the same underlying construct. This interpretation fits with the common perception of English courses being more subjective than math courses in terms of desired outcomes and would explain why students with similar English test scores can earn similar grades even if they are in different ability groups.

The second explanation is similar but focuses on the link between content across ability groups rather than across English courses and standardized tests of English
knowledge. The explanation is that content differences across English ability groups may be more horizontal than vertical. Whereas math courses tend to rely on cumulative knowledge and more rigorous ability groups will assume that students already possess knowledge being taught in less rigorous groups, English classes may not have the same hierarchical construction across groups. The lack of a definite connection in curricular content across groups could also lead to an absence of a consistent relationship between outcomes on assessment measures across groups. In the situation described above, students would be judged based upon their immediate performance and would not be penalized for a deficiency in preparatory knowledge, as they might be in math courses. This would also limit comparisons of students within a classroom because knowledge gained prior to entering the classroom would not be a valid point of comparison.

Overall, the results indicate that instructional experiences can have strong relationships with grades. However, those effects are dependent on location and context. The apparently segmented determinants hint at the complexity of grades as a measure of educational outcomes. Assignment of grades appears to be dependent on academic comparisons to other students within the classroom in some situations but relatively independent in other courses. The variations may be related to the nature of course content across ability groups. In order to determine if grades are affected differentially by student behaviors across ability groups, analyses were performed that included interactions of the ability group variables with the classroom engagement and classroom demeanor variables discussed below (results not shown). The majority of the interactions were insignificant, indicating that academic requirements and standards may vary across ability groups, but behavioral standards are relatively constant across groups.
The next section includes two variables that represent students’ social interactions. Teachers reported how often the student interacted with them outside of class and how well the student related to others. These variables are included in the analyses to address the theoretical component on social interactions by assessing the frequency and tenor of students’ interactions with others while in school.

The results indicate that students’ social interactions, as measured by teachers’ reports of the interactions, are not related to grades. The results are consistent across both subject matter areas in that neither variable is significantly related to grades for either math or English. These findings indicate that students’ levels of sociability do not directly impact their final grades.

The hypothesis presented earlier predicted that students who interact with their teachers more often than do their peers and have positive relationships with members of the school community would receive higher grades than students who had fewer or negative interactions. Frequent and positive interactions are thought to ease classroom instruction, provide for further transmission and reinforcement of classroom norms, allow students to receive informal feedback, and indicate students’ acceptance of classroom goals. All of the factors mentioned should have positive impacts on grades. However, a closer look at the theory reveals that it may not be the frequency and tone of interactions themselves that matter – it is the academic behaviors that are connected to them that will directly influence assessments. To test this, analyses were conducted to determine if removing the variables representing student behavior from the model would reveal a relationship between social interactions and grades. The results (not shown) confirmed a positive relationship between both variables representing social interactions and grades.
This indicates that there is an indirect effect of social interactions that works through student behavior to influence grades.

The next group of variables was chosen to represent teachers’ impressions of student behavior. The section is divided into two subsets. Both subsets represent teachers’ perceptions of students’ acceptance of the norms of achievement and independence. The first represents classroom engagement and includes variables that can be objectively measured or observed by teachers. The second section analyzes teachers’ reports of classroom demeanor. The variables in that section are more subjective and address attitudinal impressions.

The classroom engagement section includes three variables – frequency of homework completion, absences, and tardiness – that reflect teachers’ reports of students’ engagement with the course. These variables indicate students’ presence in the class in terms of their willingness and ability to meet minimal requirements of school and classroom demands. The actions measured by these reports also express students’ levels of acceptance of the norms of achievement and independence to teachers. All three indicators are relatively objective measures for which the teachers may maintain records which they could take into consideration when assigning grades.

The results for the variables in the classroom engagement section are generally consistent across both math and English so the two courses will be discussed simultaneously. The first variable is teachers’ reports of how often the student in question completes homework. This variable was chosen for inclusion in the analyses because homework is an independent activity – students must complete it through their own initiative. Therefore, homework represents a concrete opportunity for students to
illustrate acceptance of the norms of independence and achievement by working independently to achieve the educational goals set by teachers.

As hypothesized, the variable has a significant positive relationship with grades for both courses. The positive impact of completing homework on grades does not appear to be mediated through knowledge gains because standardized test scores are controlled for in the model. However, the effect could be related to course-specific knowledge gains that would not be applicable to standardized tests. If this were true, it might be expected that the size of the homework coefficient would be smaller for math than for English because it is believed that the content of math courses is more tightly linked to standardized test content so the benefits of doing homework would be partially soaked up by the standardized test score. This is not the case – the coefficients are almost exactly the same size. Instead, the connection between homework and grades may occur because completing independent assignments signals to teachers that students have accepted the goals of the classroom and are working to achieve them. For this reason, teachers may reinforce the importance of these goals by assigning credit to students simply for completing homework assignments or assigning partial credit to reward perceived effort even if students do not correctly finish the homework.

The next variable is teachers’ reports of how often the student is absent. These reports have a consistent negative effect on grades in both models. Because standardized test scores are controlled for in the model, the relationship between grades and number of days absent is not due solely to lost opportunities to learn and subsequent knowledge gaps. The negative impact of missing school that accrues to grades may result from students missing course-specific knowledge on those days, not handing in homework, not
making up work, or because the teacher penalizes students who miss class. Regarding each of these explanations, the grade penalty for missing school does not appear to be completely related to the potential to fall behind in terms of course content because students and teachers could work together to ensure that students are able to catch up after missing classes and avoid negative academic consequences. For that reason, it could be a partially non-academic grade reduction that penalizes students for illustrating a lack of engagement with educational goals or that is designed to discourage students from missing class because it places additional burdens on teachers.

It is also important to note, in regard to teachers’ reports of students’ absences, teachers were not asked to specify why students missed classes. It is possible students who miss class for different reasons, e.g. being sick versus being truant, may experience different outcomes in terms of how teachers react to their absences and the possible effects on their grades. More detailed questions might help to clarify the results and determine whether the reduction in grades related to missing classes is simply a result of student absence or if it is a penalty for illustrating a rejection of school norms.

The final variable in this subset is teachers’ reports of how often the student is tardy for class. These reports have no significant relationship to grades. The lack of a consistent relationship indicates that teachers do offer some behavioral leeway within the classroom and that many are willing to overlook minor infractions such as being late to class. Teachers may refrain from systematically reducing grades on the basis of small infractions in favor of focusing on behaviors that have more potential to directly impact classroom achievement, such as those described above. In addition, maintaining records of rule infractions is time consuming and teachers may choose to overlook problems that
they consider relatively inconsequential in favor of focusing time for administrative tasks on more substantial classroom issues.

The results described above hint at an important facet of grades as an indicator of educational outcomes – teachers may penalize students for engaging in actions that have the potential to reduce achievement or disrupt instruction. The rewards that students receive for completing homework and the deductions they receive for missing classes do not appear to be directly related to academic achievement. Instead, the variations in grades based on these criteria are related to possible academic outcomes and perceptions of student commitment to classroom goals.

The classroom demeanor subset of student behavior variables represents teachers’ impressions of students’ classroom behavior. The section contains four variables – whether or not student is attentive, extremely passive, disruptive, and whether or not the student works hard. These reports are more subjective than those used in the classroom engagement section – they cannot be directly measured by teachers. The rely more on teachers’ perceptions and interpretations of students’ behavior and level of commitment than do the variables in the previous section.

The first variable is teachers’ reports of how often the student is attentive in class. As hypothesized, teachers’ reports of student attentiveness have a positive relationship to grades for both courses, although the effect is stronger for math courses than for English courses. Student attentiveness is impossible for teachers to measure directly and it is conceivable for students to feign attentiveness to some degree. Therefore, positive ratings may be based on proxies for attentiveness such as successful completion of assignments and tests, frequent eye contact during lectures, ability to respond to questions
in class, note taking, and instances of student-initiated participation in class discussion. All of those markers may indicate levels of student attentiveness, but they are not guaranteed to be accurate. However, all of the proxies listed above do facilitate classroom instruction and serve as indicators of student commitment to the goal of academic achievement. Therefore, benefits may accrue to students who successfully appear to be attentive partially because teachers appreciate the behaviors associated with attentiveness and judge attentive students to also be well-socialized students.

There are two primary ways that the benefits of attentiveness are translated into student grades. Academically, students may directly benefit from attentiveness by gaining course-specific knowledge (general knowledge is controlled for through standardized test scores). Beyond the academic benefit, many teachers also include an explicit participation element in grades. This grading domain may be where students receive the primary benefits of appearing attentive because several of the proxies for attentiveness are elements of participation. In fact, participation may often be a component of grades for the reasons noted in the previous paragraph: Participation facilitates instruction and provides a method for teachers to judge students’ achievement orientation.

The next variable is teachers’ reports of whether or not the student is exceptionally passive. No significant relationship exists between these reports and grades for either course. As was noted in the theory section, teachers probably appreciate a degree of student passivity because it allows teachers to control the class more easily. However, an “exceptionally” passive student would appear disengaged from the course and the student might garner academic deductions for lack of involvement. Ultimately,
though, the lack of a consistent relationship between reports of passivity and grades may
result from the fact that passivity is not directly related to acceptance or rejection of
educational norms. Although passive students may alternately ease classroom instruction
by being willing followers or impede instruction by failing to participate, this
characteristic of students does not necessarily reflect their achievement orientation and
willingness to work independently. Therefore, it is neither accounted for in their
academic performance nor in teachers’ assessments of students.

The next variable is teachers’ reports of how often the student is disrupt
tive in class. Counterintuitively, the variable is positively related to grades in both courses with
a stronger relationship to math than English grades. It was hypothesized that this variable
would be negatively related to grades because the behavior would represent, to teachers,
a rejection of educational norms. The norms specifically at issue include those regarding
when and how to participate in the classroom setting and the prediction of a negative
effect is predicated on the assumption that the majority of the disruptions are non-
academic and impede instruction. Based on these presumptions, it is difficult to
understand how disruptive students could receive higher grades than their less disruptive
peers, particularly when considering the theory developed earlier.

A reconceptualization of how teachers interpreted “disruptive” when completing
the survey could bring the finding more in line with the theory. Teachers potentially
included a wider variety of behaviors under this category than those described above.
Perhaps disruptions do not have to be synonymous with negative classroom experiences
and can include boisterous students who ask academically-oriented questions out of turn
or insecure, but academically-capable, students who repeatedly request clarifications of
instruction and assignments. In those situations, it is possible that the students disrupt instruction but still illustrate an orientation to academic achievement that may be rewarded. Additionally, inappropriate, but achievement-related, intrusions into instruction may benefit the student through participation grades, additional clarifications of academic demands, and extra instructional attention from the teacher. This might explain why, unexpectedly, the analyses indicate that disruptive students are academically better off than passive students. However, this possibility is highly speculative and does not offer a strong explanation for the unexpected finding. It clearly warrants additional inquiry and thought.

The next variable is teachers’ reports of whether or not the student usually works hard. This has a strong positive relationship with grades for both types of courses. This relationship was hypothesized because the variable directly addresses teachers’ perceptions of students’ achievement orientation and acceptance of the goals of classroom education. Again, it is important to remember that the strong effect occurs despite controlling for test scores and previous grades. In addition, how often the student completes homework is accounted for in the model so the grade boost does not arise from completing assigned work. Therefore, students with similar knowledge levels, educational histories, and classroom work records are differentiated in their grades by their apparent dedication to work toward achieving educational goals.

Teachers’ reports of how hard students work represent an obscure category of teachers’ perceptions of students because it is not clear on what basis teachers assign the ratings. Because frequency of homework completion is accounted for, the grade benefits are not based solely on how often students finish assigned work. Other reports included
in the model, such as student attentiveness and frequency of interactions with the teacher, also account for possible methods that teachers could use to judge student effort. Instead, the grade benefit appears to be based on other indicators of student effort. These may include self reports from students, perceived effort during in-class work periods, and improvements in student work over time. Although it is unclear how teachers judge students’ effort, it is apparent from the results that those judgments matter and that teachers reward hard working students with higher grades.

The results described above are for variables that most people might assume would be related to grades. However, the behaviors measured by the variables are often thought to operate on grades by improving academic achievement. These analyses show that they have at least some independent effects that are separate from improvements in students’ academic work. Similar to the first section on classroom engagement, the findings indicate that teachers reward students for attitudes that might increase academic achievement and that indicate a dedication to achievement and independence rather than just for concrete improvements in academic work.

The next set of variables relates to students’ educational futures. Teachers were asked if they thought their students would go to college. Two dummy variables were created from these reports – the student probably will not go to college and the teacher does not know if the student will go to college. The omitted category is that the student probably will go to college. These variables were included because they represent teachers’ perceptions and understandings of student commitment to education and continuing education.
As hypothesized, both of the variables are negatively related to grades for both math and English. This indicates that students receive lower grades if their teachers perceive they are not academically prepared for college or they do not plan to attend college or if the teachers are unable to determine students’ academic potential or plans. Because academic achievement is controlled for in the models through the inclusion of test scores, prior grades, and other variables representing student performance and effort, it is unlikely that the negative relationship these variables have with grades is due predominantly to academic performance or teachers’ understandings of students’ academic abilities. Therefore, alternative avenues must be explored.

First, it is important to consider how teachers develop their perceptions of students’ future educational plans. In terms of teachers’ perceptions that students probably will not attend college, they may be based on discussions with students in which they stated their plans or teachers may have formed their opinions based on perceptions of student refusal to pursue academic achievement or disengagement from academic goals. Teachers who are unsure if students plan to attend college may simply not have meaningful interactions with the students in question. They also may believe that the students are not prepared for college even though the students plan to attend and, therefore, be unsure of the outcome of the students’ aspirations.

How these perceptions translate into grade reductions is an open question. Some clues may be found in the potential explanations for how teachers develop their perceptions of students’ college plans. If teachers’ reports in this section are based upon impressions of student engagement or on the strength of the relationship between teacher and student, grade variations may arise from some of the similar mechanisms mentioned
in earlier sections, such as participation grades. However, the strong effect of these perceptions, despite the inclusion of several variables representing student behavior, implies that something more substantial contributes to the effects. Although it is unclear at this point how the relationship between grades and educational plans is enacted, what is clear is that perceived educational aspirations have powerful and consistent effects on grades which are not completely based on academic performance. Once again, students who appear to value education, irrespective of how well they do on standardized tests, are rewarded in terms of assessments.

The analyses described above clarify several aspects of grades as a measure of educational outcomes. They simultaneously raise questions. This conclusion will address both aspects of the results, beginning with the clarifications.

First and most important, the analyses clearly indicate that grades are different than standardized test scores as a form of educational assessment. They account for factors that cannot be measured by standardized tests. This conclusion strikes a cautionary note for researchers and policymakers who may conceptualize grades and standardized test scores as similar measures of academic achievement and use them interchangeably. In addition, the results provide the initial supports for a theory of grades that can improve the use of grades in research and offer promising new avenues of exploration in terms of educational outcomes.

Also clarified by the analyses – and implicitly stated in the previous paragraph – is the fact that grades include considerations of factors other than academic achievement. As was noted earlier, this possibility led researchers and policymakers to be wary of using grades to assess educational outcomes. However, the theoretical understanding of
grades developed earlier highlights the importance of the non-academic factors included in grades for success both in educational and career settings.

Regarding specific variables in the analyses, the findings supported the theory in several important ways while also adding nuance to the initial conceptualization of grades. Instructional experiences were found to be related to grades, although in a more complex way than was initially expected. Social interactions were also important although the indirect nature of their impact was undervalued in the theoretical development. Teachers’ perceptions of students’ behaviors – both objectively measured actions and subjectively interpreted attitudes – can strongly influence assignment of grades, as was predicted. Finally, students’ educational aspirations, as understood by teachers, illustrated a strong relationship with grades which indicates the importance of students’ commitment to education.

Tacitly present in the findings described above are two aspects of the theory that were either not highlighted in the discussion of the results or directly addressed in the analyses. First, teachers matter. This is an overarching aspect of the analyses that pervades the findings to the point that it might be overlooked. While few people would believe that teachers do not matter at all for students, teachers’ influence is often characterized in terms of whether or not they can effectively improve students’ academic achievement. The other side of the equation, teachers’ roles in evaluating student performance and providing feedback, is likely to be undervalued in the sociological literature.

The second aspect of the theory that is alluded to in the findings is that grades are a product of interactions and negotiations between students and teachers. Although this
cannot be tested directly with the data in these analyses, a majority of the teachers’ accounts report actions and behaviors that must be observed over time. Teachers develop their beliefs about students’ achievement orientation through a two-way process in which students and teachers can impact each other. Therefore, it is important to consider the role of students in influencing educational outcomes and to remember that, although teachers ultimately assign grades, students are not simply bystanders in the process.

The primary question raised by the analyses is one of mechanisms. Although several possible explanations were offered within the discussion of results, it is still unclear exactly how some of the relationships revealed in the analyses operate. More detailed data may be able to provide further clarification of how student behavior is translated into grades. For now, the results in this chapter provide a solid base for understanding the composition of grades which can inform further efforts to understand how grades are created within a classroom setting.

The following chapters expand upon the analyses presented here and provide an opportunity to strengthen the theoretical understanding of grades developed earlier. First, analyses will be run with standardized test scores as the dependent variable and the same explanatory variables used for the analyses discussed in this chapter. These results will be used to clarify the role of teachers’ perceptions in academic outcomes and to evaluate if test scores are impacted differently than grades by student behavior in order to verify that the grades and tests represent different outcomes. Another set of analyses will be conducted in which students’ reports of their own behaviors will be included with teachers’ reports as predictors of grades. It is expected that teachers’ reports will have stronger relationships to grades than students’ reports, thereby highlighting the
importance of teachers’ perception in determining academic outcomes. Finally, classroom observations will be analyzed to examine the interactions and negotiations which occur in the classroom and to identify mechanisms that may work to translate students’ behaviors into grades.
CHAPTER 5:

RELATIONSHIPS BETWEEN TEACHERS’ REPORTS OF STUDENTS’ BEHAVIORS AND TEST SCORES

This section presents the results of analyses which use the same explanatory variables as the previous chapter with 10th grade standardized test scores, rather than transcript-indicated grades, as the dependent variable. The only change in the composition of the model is in the control variables. Rather than controlling for 10th grade standardized test score and students’ 8th grade self-reports of grades, students’ 8th grade standardized test score is used as a control for prior achievement. As expected, this variable has a strong positive relationship with 10th grade standardized test scores.

The supplementary analyses described in this chapter were conducted in order to provide a comparison to the analyses of the determinants of grades and examine if grades and standardized test scores do measure separate academic outcomes or if the effects found in the previous chapter were simply accounting for student knowledge in an unidentified way. Because of the difference in control variables noted above, the explanatory variables in these analyses do not represent exactly the same effects as they do in the analyses with grades as the dependent variable. In those analyses, the explanatory variables indicate a more immediate effect on the dependent variable because current test scores are controlled in the models. Due to these differences the results will
be interpreted cautiously. They are not assumed to represent an ideal comparison between the determinants of grades and test scores.

The analyses are expected to support the theory by illustrating the differences between predictors of grades and predictors of test scores. Based on the results for grades and the theoretical understanding of grades and test scores, several hypotheses were developed regarding predictors of test scores. Assignment to honors or academic ability group (compared to a general ability group) is expected to positively predict test scores because these assignments increase students’ opportunities to learn and test scores are theorized to provide a relatively objective measure of knowledge. Teachers’ reports of students’ social interactions and behaviors are anticipated to have small or no effects on test scores because standardized tests do not account for these aspects of education. However, some of the more objective behavior variables such as homework completion and number of days absent may be related to test scores because they represent student engagement with learning opportunities. Finally, it is unclear what to expect regarding the variables representing teachers’ perceptions of students’ educational futures. Because the mechanism that connects these variables to grades is not fully conceptualized, this linkage cannot be used to predict how they might be connected to test scores. It is hoped that the results of the analyses of test scores might help to clarify the connection between grades and teachers’ reports of students’ postsecondary plans or options.

The results for the section on instructional experiences, as shown in Table 5.1, show a significant relationship between ability group and students’ outcomes on standardized tests. As predicted, students in honors and academic classes earn higher test scores than students in general classes, the test scores of students in vocational classes are
TABLE 5.1
REGRESSION OF MATH AND ENGLISH TEST SCORES
ON TEACHERS’ REPORTS OF STUDENTS’ INSTRUCTIONAL EXPERIENCES,
SOCIAL INTERACTIONS, BEHAVIORS, AND EDUCATIONAL FUTURES

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-.507*</td>
<td>-.507**</td>
</tr>
<tr>
<td>Black</td>
<td>-2.504***</td>
<td>-.846**</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-.834*</td>
<td>-.461</td>
</tr>
<tr>
<td>Asian</td>
<td>-1.17</td>
<td>-102</td>
</tr>
<tr>
<td>Native American</td>
<td>-2.268</td>
<td>-.342</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>.659***</td>
<td>.573***</td>
</tr>
<tr>
<td>8th Grade Test Score</td>
<td>.807***</td>
<td>.772***</td>
</tr>
</tbody>
</table>

INSTRUCTIONAL EXPERIENCES

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honors level class</td>
<td>4.181***</td>
<td>1.644***</td>
</tr>
<tr>
<td>Academic level class</td>
<td>2.459***</td>
<td>.807***</td>
</tr>
<tr>
<td>Vocational class</td>
<td>-.642</td>
<td>-.901</td>
</tr>
<tr>
<td>Other type of class</td>
<td>-1.883*</td>
<td>-1.726**</td>
</tr>
</tbody>
</table>

SOCIAL INTERACTIONS

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student talks with teacher outside of class</td>
<td>-.126</td>
<td>.033</td>
</tr>
<tr>
<td>Student relates well to others</td>
<td>-.106</td>
<td>-.442</td>
</tr>
</tbody>
</table>

STUDENT BEHAVIOR

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom Engagement</td>
<td>.318</td>
<td>.135</td>
</tr>
<tr>
<td>How often student does homework</td>
<td>.318</td>
<td>.135</td>
</tr>
<tr>
<td>How often student is absent</td>
<td>-.317</td>
<td>-.031</td>
</tr>
<tr>
<td>How often student is tardy</td>
<td>-.343</td>
<td>-.087</td>
</tr>
<tr>
<td>Classroom Demeanor</td>
<td>.215</td>
<td>.497**</td>
</tr>
<tr>
<td>How often student is attentive</td>
<td>.215</td>
<td>.497**</td>
</tr>
<tr>
<td>Student is passive</td>
<td>-.217</td>
<td>.126</td>
</tr>
<tr>
<td>Student is disruptive in class</td>
<td>-.293</td>
<td>-.206</td>
</tr>
<tr>
<td>Student usually works hard</td>
<td>-.070</td>
<td>-.024</td>
</tr>
</tbody>
</table>

EDUCATIONAL FUTURE

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student probably won’t go to college</td>
<td>-3.545***</td>
<td>-2.241***</td>
</tr>
<tr>
<td>Don't know if student will go to college</td>
<td>-1.753***</td>
<td>-0.958***</td>
</tr>
</tbody>
</table>

N=3547                     N=4405

NOTE: Standard deviations for all coefficients are presented in parentheses.
NOTE: *p<.05; **p<.01; ***p<.001
not significantly different, and students in “other” ability groups receive significantly lower test scores. These results are consistent across both math and English.

The results described above are in marked contrast to the analyses of grades. In those models, math students in progressively more challenging ability groups earned lower grades than their peers in less rigorous ability groups. There were no consistent effects on grades by ability group for English.

These results may initially appear contradictory. However, they are in keeping with the theoretical understanding of the differences between grades and test scores developed earlier. In terms of test scores, the findings indicate that students in more challenging ability groups, on average, accumulate more subject-related knowledge than their peers in less challenging groups. Test scores reflect these differences and provide a method for comparing students across groups.

Grades represent a different measurement than test scores. In math classes, they appear to compare students within ability group rather than across ability group. Therefore, a student in an honors math class may gain more knowledge than a similar peer in a general math class and subsequently earn a higher score on a standardized test. However, the honors student may not be able to perform as well as the other students in his or her class and may receive a lower grade than the peer in the general class who is able to complete work at or above the average achievement level within the class.

The results for English test scores help to clarify the findings for English grades. The analyses of test scores indicate that, on average, students in honors and academic ability groups gain more knowledge than their peers in less challenging groups. However, when examining grades and holding current test scores constant, students
across English ability groups (excluding “other”) earn similar grades. Two explanations were offered for the findings regarding grades in the previous chapter. The first was that the content on standardized tests of English was not tightly linked to the classroom content that was assessed in English classes. The second was that the curricula in English ability groups was not hierarchically constructed so that students in honors or academic groups were not exposed to more challenging material than students in general or vocational ability groups. The results for test scores tend to rule out the second explanation because they indicate that students in progressively more rigorous ability groups are gaining more knowledge than their peers.

Although the findings discussed here cannot directly confirm the alternative explanation – that assessed content in English classes is not tightly linked to content on standardized tests – they do lend some support to that explanation by eliminating the other proposed option for the relationship between ability group and grades. The results for English grades and English standardized test scores, when taken together, suggest that students in English classes are assessed by their teachers on material that deviates from the content on standardized tests. One reason for this departure may be that English teachers are likely to provide instruction and evaluate students on course-specific knowledge that is not on standardized tests. The disconnect may also occur because the types of assessments that teachers use in these courses are different than those used on standardized tests. For instance, grades in English classes may rely heavily on critical analysis, essays, or creative writing while test scores are based more on reading comprehension and vocabulary. The observed differences in outcomes suggest that
particular caution should be used when considering whether test scores or grades will be used to assess students’ outcomes in English courses.

As predicted, the results for the variables representing social interactions are similar to the models for grades in that neither variable has a significant direct effect on standardized test scores. The results for the full set of student behavior variables indicate that only one of the reports is significantly related to students’ test scores: Reports of student attentiveness positively predict test scores, but only in English classes. This finding indicates that students who are perceived to focus on instruction in English classes receive benefits through an increase in subject-related knowledge. The remaining six variables are not related to either math or English standardized test scores.

The most interesting outcome in this section is the overall lack of relationship between teachers’ reports of student behavior and test scores. Although it was predicted that the relationships would be weaker than with grades or that fewer relationships would be found, a difference of this magnitude was unexpected. In the previous chapter, five of the seven behavior variables had strong to moderate relationships with grades for both types of courses. These results provide additional support for the presence of systematic differences between grades and test scores. And they highlight the role of teachers in grading and teachers’ ability to reward or punish students for their behavior through grades, in addition to or in spite of students’ academic performance.

The last two variables, teachers’ reports that they believe the student will not go to college or that they do not know if the student will go to college, have similar relationships to test scores as they do to grades. Both variables are strongly and negatively related to test scores. The similarity of these results for both grades and test
scores indicates that whatever is connecting these reports to the outcomes might not be working through teachers’ perceptions and their subsequent evaluations of students – but might instead be capturing some underlying element or indirect effect that impacts students’ outcomes in terms of knowledge gains. This factor may be a type of effort or commitment that is not captured in the behavior variables. It may also relate to outside forces such as parental involvement with education, which are not accounted for in the model. It is also possible that teachers are basing their reports of students’ likelihood of attending college on students’ academic outcomes, thereby creating a causality issue because the report is based on the outcome that is being examined.

In order to explore the possible causality problem mentioned above, a new test score model was estimated that did not include teachers’ reports of students’ educational futures. The results are presented in Table 5.2. Similar analyses were also conducted for the models for grades discussed in Chapter Four. These results are not presented because only two small substantive changes in coefficients were found – for English, “other” ability group becomes slightly less significant and reports of students’ disruptive behaviors lose significance. These results indicate that teachers’ perceptions of students’ educational futures have effects on grades that are independent of the behaviors accounted for in the models.

The models for test scores indicate, in contrast to those for grades, that teachers’ reports of students’ educational plans are obscuring behavioral factors that affect test scores. While the models show only a few small changes in the results for instructional experiences and social interactions, several variables in the behavior section gain significance compared to the models discussed earlier in this chapter. In the classroom
TABLE 5.2
REGRESSION OF MATH AND ENGLISH TEST SCORES
ON TEACHERS’ REPORTS OF STUDENTS’ INSTRUCTIONAL EXPERIENCES,
SOCIAL INTERACTIONS, AND BEHAVIORS

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-.511*</td>
<td>-.567**</td>
</tr>
<tr>
<td>Black</td>
<td>-2.395***</td>
<td>-.743*</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-.704</td>
<td>-.357</td>
</tr>
<tr>
<td>Asian</td>
<td>-.025</td>
<td>.031</td>
</tr>
<tr>
<td>Native American</td>
<td>-2.181</td>
<td>-.350</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>1.039***</td>
<td>.806***</td>
</tr>
<tr>
<td>8th Grade Test Score</td>
<td>.835***</td>
<td>.792***</td>
</tr>
</tbody>
</table>

INSTRUCTIONAL EXPERIENCES
Honors level class     4.702*** (.425) 1.975*** (.263)
Academic level class   3.004*** (.279) 1.036*** (.206)
Vocational class       -.972 (.773) -1.329 (.759)
Other type of class    -1.971* (.877) -2.169*** (.528)

SOCIAL INTERACTIONS
Student talks with teacher outside of class .012 (.238) .078 (.181)
Student relates well to others .201 (.485) -.179 (.304)

STUDENT BEHAVIOR

Classroom Engagement
How often student does homework .537** (.182) .326* (.145)
How often student is absent -.423* (.187) -.120 (.149)
How often student is tardy -.379* (.180) -.082 (.132)

Classroom Demeanor
How often student is attentive .319 (.212) .552*** (.156)
Student is passive -.444 (.455) .083 (.338)
Student is disruptive in class -.382* (.180) -.211 (.121)
Student usually works hard .337 (.373) .307 (.252)

N=3547 N=4405

NOTE: Standard deviations for all coefficients are presented in parentheses.
NOTE: *p<.05; **p<.01; ***p<.001
engagement section, all of the coefficients for math gain marginal significance. Students who complete math homework earn higher scores on standardized tests than those who do not. Students who are frequently absent from or late for class receive slightly lower scores than their peers who are more often present. For English classes, only reports of how often students complete homework have a marginally positive relationship with test scores. In the student demeanor subsection, reports of how often students are disruptive are slightly and negative related to math test scores. Reports of how often students are attentive were significantly positively related to English grades in the previous model and gain additional significance in the current model.

The results presented in Table 5.2 offer further evidence of differences between grades and test scores. First, although several behavior variables did gain significance in these models, student behavior still does not have as strong of a relationship with test scores as it does with grades. Second, different types of behaviors appear to predict grades and test scores. Reports of student tardiness are not related to grades but they are negatively related to math test scores. Reports of whether or not students are attentive or work hard are strongly related to grades for both subjects but only student attentiveness is related to English test scores. And, in a reversal of earlier results, disruptive students receive lower scores on math tests than less disruptive peers. In the models predicting grades, disruptive students earned higher grades.

The findings noted above indicate that different behaviors matter for each measure of educational outcomes. They also illustrate that for test scores, unlike for grades, different behaviors matter for test results across subject. Only one classroom engagement variable is significant for English while all three are significantly related to
test scores for math. The differences hint that engagement with opportunities to learn presented in the classroom is more important for math than English test scores. This difference is consistent with the interpretation of the effects of instructional experiences discussed earlier. If course content in math is more tightly linked to standardized tests than the content of English courses, it makes sense that students who are more engaged in math courses would receive higher test scores than their peers and that the effect would be stronger than it is for English courses. In the classroom demeanor section, one variable is significantly related to test scores for each subject. For math, disruptive students earn lower test scores and, for English, attentive students receive higher test scores. Again, these results provide evidence that different aspects of behavior are more important for success on standardized test across the two subject areas.

The findings in Table 5.2 show that behavior does matter for test scores. However, the relationships between behaviors and test scores are not as strong as the relationships with grades. And the relationships appear to operate differently. For math, student engagement with opportunities to learn appears to be the most important element of behavior for test scores, probably because this behavior increases knowledge gains. For English, student attentiveness is the single most important factor, indicating that the most vital behavioral element for improving test scores is that students care about and are interested in course content. With grades, unlike test scores, teachers can directly account for behaviors, which helps to explain the similarity and strength of behavioral effects across subjects.

The results presented in this chapter present a basis for validating theoretical elements. When the results in this and the previous chapter are taken together, they offer
support for the contention that grades and standardized test scores measure different types of academic achievement. The findings reinforce the theoretical position that, unlike grades, standardized test scores provide mainly a representation of student knowledge and engagement with opportunities to learn and do not offer a strong account of behaviors associated with the norms of achievement and independence.

The findings also indicate that teachers play an important role in determining students’ outcomes for both measures. The variables representing ability groups act as a composite of instructional factors that mediate students’ knowledge gains. Teachers are an integral element of this composite. The strong relationship between ability group and test scores illuminates the importance of teachers for offering instruction and insuring that students are provided with opportunities to learn which are appropriate for students’ ability levels and academic preparation.

Teachers perform an equally important, but more direct and complex, function in determining students’ grades. Teachers’ roles in grading are more direct than with test scores because they both assign grades and primarily determine what will be assessed by grades whereas the composition and content of standardized tests is outside of their control. Teachers’ roles in grading are more complex than with test scores because, rather than simply providing opportunities to learn and evaluating knowledge levels, the analyses in the previous chapter illustrate that teachers also account for academic performance and commitment to educational norms. In terms of grades, teachers both mediate and judge student performance within the classroom environment.

The findings in this chapter and the previous chapter point toward an interesting link between grades and test scores. It is possible that one of the primary functions of
including behavioral factors in grades is to enforce standards that will facilitate improvements in knowledge acquisition and subsequently raise test scores. The behaviors that teachers assess through grades are not arbitrary – they are related to classroom instruction and might lead to improvements in academic achievement, as discussed in the previous chapter. Students who complete homework, attend class regularly, pay attention, and participate in discussions facilitate instruction. Individual students who meet these classroom standards are rewarded with better grades. However, the behaviors for which students are rewarded can increase and improve learning opportunities for all students in the class. Therefore, some of the instructional benefits of the assessed behaviors can diffuse to all students in the classroom and lead to classroom-wide improvements in test scores. While the individual student benefits from engaging in approved behavior through an improved grade, the whole class might benefit through better test scores.
CHAPTER 6:
RELATIONSHIPS BETWEEN STUDENTS’
REPORTS OF THEIR BEHAVIORS AND GRADES

This chapter presents the results of analyses which use students’ reports of their own behaviors and attitudes to predict their grades in math and English. Two models are presented for each academic subject. The first includes the control variables used in models in previous chapters, teachers’ reports of the ability group of the course for which the grade was assigned, and students’ reports of their behaviors and attitudes. The second model includes all of the variables in the first model, plus teachers’ reports of students’ behaviors and attitudes that were used as explanatory variables in the analyses presented in Chapters Four and Five. The analyses are designed, primarily, to determine if teachers’ perceptions of students outweigh students’ own reports in terms of predicting grades.

All of the students’ reports come from the sophomore year (1990) surveys to ensure that the reports represent the same timeframe as the teacher reports and the grades used as dependent variables. Descriptive statistics for all student variables are presented in Table 6.1. The sample size for both models is limited to students for whom teacher reports are available in order to maintain consistency with the sample used for all other analyses. The student surveys asked students to report on their behaviors and attitudes in
### TABLE 6.1

**DESCRIPTIVE STATISTICS**

**FOR STUDENT VARIABLES IN THE ANALYSES**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>English</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SOCIAL INTERACTIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent feels put down by teachers</td>
<td>1.927</td>
<td>1.929</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Most teachers listen to respondent</td>
<td>2.752</td>
<td>2.752</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>STUDENT BEHAVIOR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classroom Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time spent on homework</td>
<td>4.694</td>
<td>4.725</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>Often go to class without homework done</td>
<td>2.020</td>
<td>2.012</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>How many days absent</td>
<td>2.846</td>
<td>2.843</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>How often did respondent skip classes</td>
<td>.511</td>
<td>.507</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Classroom Demeanor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How important are good grades</td>
<td>3.442</td>
<td>3.449</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>How often did respondent get in trouble</td>
<td>3.296</td>
<td>3.310</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>It is ok to disobey school rules</td>
<td>.575</td>
<td>.598</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>It is ok to talk back to teachers</td>
<td>3.414</td>
<td>3.412</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>EDUCATIONAL FUTURE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondent plans to attend college</td>
<td>.814</td>
<td>.806</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

N=3547                  N=4405

**SOURCE:** National Education Longitudinal Study of 1988

**NOTE:** Standard deviations for all non-dichotomous variables are presented in parentheses.

general so the student reports are not specific to the class for which teacher reports or grades are available.

In Tables 6.2 and 6.3, the explanatory variables that come from the student surveys are denoted with an “S” before the variable label in order to clearly differentiate them from teachers’ reports. The variables representing students’ reports were chosen to match as closely as possible the content of reports provided by teachers. Table 6.2


### TABLE 6.2

REGRESSION OF MATH GRADES ON STUDENTS’ AND TEACHERS’ REPORTS OF STUDENTS’ INSTRUCTIONAL EXPERIENCES, SOCIAL INTERACTIONS, BEHAVIORS, AND EDUCATIONAL FUTURES

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>.321**</td>
<td>-.072</td>
</tr>
<tr>
<td>Black</td>
<td>-.766***</td>
<td>-.772***</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-.445**</td>
<td>-.539***</td>
</tr>
<tr>
<td>Asian</td>
<td>.101</td>
<td>-.100</td>
</tr>
<tr>
<td>Native American</td>
<td>-.490</td>
<td>-.343</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>.251***</td>
<td>.125*</td>
</tr>
<tr>
<td>10th Grade Test Score</td>
<td>.100***</td>
<td>.070***</td>
</tr>
<tr>
<td>8th Grade Grade</td>
<td>.483***</td>
<td>.243***</td>
</tr>
</tbody>
</table>

**INSTRUCTIONAL EXPERIENCES**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>Honors level class</td>
<td>-.676***</td>
<td>-.725***</td>
</tr>
<tr>
<td>Academic level class</td>
<td>-.755***</td>
<td>-.719***</td>
</tr>
<tr>
<td>Vocational class</td>
<td>.808*</td>
<td>.722*</td>
</tr>
<tr>
<td>Other type of class</td>
<td>.187</td>
<td>.422</td>
</tr>
</tbody>
</table>

**SOCIAL INTERACTIONS**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: Respondent feels put down by teachers</td>
<td>-.221**</td>
<td>-.131*</td>
</tr>
<tr>
<td>S: Most teachers listen to respondent</td>
<td>.244**</td>
<td>.144*</td>
</tr>
<tr>
<td>Student talks with teacher outside of class</td>
<td>-.023</td>
<td>.086</td>
</tr>
<tr>
<td>Student relates well to others</td>
<td>-.273</td>
<td>(.151)</td>
</tr>
</tbody>
</table>

**STUDENT BEHAVIOR**

**Classroom Engagement**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: Time spent on homework</td>
<td>.009</td>
<td>-.024</td>
</tr>
<tr>
<td>S: Often go to class without homework done</td>
<td>-.329***</td>
<td>-.151**</td>
</tr>
<tr>
<td>S: How many days absent</td>
<td>-.239***</td>
<td>-.084*</td>
</tr>
<tr>
<td>S: How often did respondent skip classes</td>
<td>-.117*</td>
<td>.077</td>
</tr>
</tbody>
</table>

**Classroom Demeanor**

<table>
<thead>
<tr>
<th></th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: How important are good grades</td>
<td>.714***</td>
<td>.413***</td>
</tr>
<tr>
<td>S: How often did respondent get in trouble</td>
<td>-.107</td>
<td>.073</td>
</tr>
<tr>
<td>S: It is ok to disobey school rules</td>
<td>-.255**</td>
<td>-.160*</td>
</tr>
<tr>
<td>S: It is ok to talk back to teachers</td>
<td>-.064</td>
<td>-.096</td>
</tr>
<tr>
<td>How often student is attentive</td>
<td>.550***</td>
<td>(.074)</td>
</tr>
<tr>
<td>Student is passive</td>
<td>-.104</td>
<td>(.163)</td>
</tr>
<tr>
<td>Student is disruptive in class</td>
<td>.249***</td>
<td>(.060)</td>
</tr>
<tr>
<td>Student usually works hard</td>
<td>1.013***</td>
<td>(.124)</td>
</tr>
</tbody>
</table>

N=3547
### TABLE 6.2 (CONT'D.)

<table>
<thead>
<tr>
<th>EDUCATIONAL FUTURE</th>
<th>Math</th>
<th>Math</th>
</tr>
</thead>
<tbody>
<tr>
<td>S: Respondent plans to attend college</td>
<td>-.146 (.134)</td>
<td>-.202 (.115)</td>
</tr>
<tr>
<td>Student probably won't go to college</td>
<td>-.626*** (.132)</td>
<td></td>
</tr>
<tr>
<td>Don't know if student will go to college</td>
<td>-.991*** (.118)</td>
<td></td>
</tr>
</tbody>
</table>

N=3547 N=3547

NOTE: Standard deviations for all coefficients are presented in parentheses.

NOTE: *p<.05; **p<.01; ***p<.001

### TABLE 6.3

**REGRESSION OF ENGLISH GRADES ON STUDENTS’ AND TEACHERS’ REPORTS OF STUDENTS’ INSTRUCTIONAL EXPERIENCES, SOCIAL INTERACTIONS, BEHAVIORS, AND EDUCATIONAL FUTURES**

<table>
<thead>
<tr>
<th>English</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>.600*** (.082)</td>
</tr>
<tr>
<td>Black</td>
<td>-1.292*** (.142)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-.243 (.130)</td>
</tr>
<tr>
<td>Asian</td>
<td>.574*** (.155)</td>
</tr>
<tr>
<td>Native American</td>
<td>-.422 (.410)</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>.402*** (.058)</td>
</tr>
<tr>
<td>10th Grade Test Score</td>
<td>.087*** (.005)</td>
</tr>
<tr>
<td>8th Grade Grade</td>
<td>.657*** (.050)</td>
</tr>
</tbody>
</table>

**INSTRUCTIONAL EXPERIENCES**

| Honors level class           | -.034 (.122)                 | -.152 (.105)                 |
| Academic level class         | -.149 (.094)                 | -.186* (.083)                |
| Vocational class             | -.355 (.319)                 | -.142 (.273)                 |
| Other type of class          | .666** (.240)                | .597** (.201)                |

**SOCIAL INTERACTIONS**

| S: Respondent feels put down by teachers | -.281** (.063) | -.165** (.054) |
| S: Most teachers listen to respondent  | .154* (.065)   | .061 (.057)    |
| Student talks with teacher outside of class | .036 (.071)   |              |
| Student relates well to others       | -.174 (.120)   |              |

N=4405 N=4405
**TABLE 6.3 (CONT'D.)**

<table>
<thead>
<tr>
<th>STUDENT BEHAVIOR</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classroom Engagement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S: Time spent on homework</td>
<td>.055*** (.016)</td>
<td>.034* (.014)</td>
<td></td>
</tr>
<tr>
<td>S: Often go to class without homework done</td>
<td>-.246*** (.057)</td>
<td>-.106* (.049)</td>
<td></td>
</tr>
<tr>
<td>S: How many days absent</td>
<td>-.263*** (.033)</td>
<td>-.090** (.029)</td>
<td></td>
</tr>
<tr>
<td>S: How often did respondent skip classes</td>
<td>-.283*** (.047)</td>
<td>-.085* (.041)</td>
<td></td>
</tr>
<tr>
<td>How often student does homework</td>
<td></td>
<td>.862*** (.056)</td>
<td></td>
</tr>
<tr>
<td>How often student is absent</td>
<td>-.271*** (.062)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How often student is tardy</td>
<td>.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Classroom Demeanor</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S: How important are good grades</td>
<td>.752*** (.064)</td>
<td>.445*** (.055)</td>
<td></td>
</tr>
<tr>
<td>S: How often did respondent get in trouble</td>
<td>-.190*** (.054)</td>
<td>-.074 (.047)</td>
<td></td>
</tr>
<tr>
<td>S: It is ok to disobey school rules</td>
<td>-.194** (.069)</td>
<td>-.165** (.059)</td>
<td></td>
</tr>
<tr>
<td>S: It is ok to talk back to teachers</td>
<td>-.154* (.061)</td>
<td>-.147** (.052)</td>
<td></td>
</tr>
<tr>
<td>How often student is attentive</td>
<td>.167** (.061)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student is passive</td>
<td>-.016 (.131)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student is disruptive in class</td>
<td>.107* (.048)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student usually works hard</td>
<td>1.282*** (.100)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>EDUCATIONAL FUTURE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S: Respondent plans to attend college</td>
<td>.218* (.111)</td>
<td>-.083 (.097)</td>
<td></td>
</tr>
<tr>
<td>Student probably won't go to college</td>
<td>-.979*** (.110)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Don't know if student will go to college</td>
<td>-.567*** (.106)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N=4405 N=4405

NOTE: Standard deviations for all coefficients are presented in parentheses.

NOTE: *p<.05; **p<.01; ***p<.001

contains regressions of variables on math grades and Table 6.3 represents the results for regressions on English grades.

Two main hypotheses were developed for this section, based on the theory developed earlier and the results in previous chapters. First, it is hypothesized that several of the students’ reports will have strong relationships to grades prior to the inclusion of teachers’ reports in the model and that the pattern of relationships will generally follow those found using teachers’ reports. This prediction is based on the expectation that students’ reports will mirror teachers’ reports in important ways and
capture the indirect effects of students’ behaviors and attitudes on grades. Specifically regarding students’ reports of their social interactions, it is predicted that the results will have weak or no relationships with grades because the effects of social interactions will operate through behavior to impact grades, similar to the results for the teacher model. Students’ reports of their behaviors and educational plans are expected to have strong relationships with grades. Consistent with the models using teacher reports, students who engage in behaviors that indicate a commitment to achievement and independence and who aspire to achieve more education will receive higher grades than their peers.

The second hypothesis for the analyses in this chapter is that the inclusion of teachers’ reports in the model will reduce or eliminate the relationships between students’ reports and grades. This expectation is derived from the theory which posits that teachers ultimately assign grades and, therefore, their interpretations of students’ behaviors and attitudes will matter more than students’ understandings of their own educational experiences in determining grades. Teachers’ reports will have a direct effect on grades, whereas students’ reports have an indirect effect that works at least partially through teachers’ perceptions. Because teacher and student reports are not perfectly matched, the student reports are expected to maintain some power in predicting grades because they will represent somewhat different aspects of students’ behaviors and attitudes than teacher reports. In addition, students’ reports may capture elements of effort or commitment to educational norms that teachers’ reports do not.

All four of the models include the same control variables as the models predicting grades in Chapter Four. Teachers’ reports of ability group were also included as control variables. The teacher reports were used because no similar reports were available from
students in the 1990 survey and because teachers are expected to have an accurate understanding of the instructional categorizations for the classes they teach. The variables are included in the analyses as control variables because ability group assignment can mediate the effects of students’ academic output and behavior on grades.

In all four models, the relationships between ability group assignment and grades remain almost the same as in the models which included only teachers’ reports as explanatory variables. The only significant difference is that, in the model with both student and teacher reports, students in academic level English classes receive slightly lower grades than students in general courses (the omitted category). The effect is rather small and, substantively, represents a minor change from the previous models. The stability of the effects of ability groups in these models reaffirms the importance of these assignments in determining grades and the persistence of differences in the determinants of grades across math and English ability groups.

Although the results for the explanatory variables representing teachers’ reports are not the focus of this chapter, the importance of teachers’ perceptions of students is reinforced by the results. The size and strength of the relationships between teachers’ reports and grades remain consistent with the findings in Chapter Four, despite the inclusion of the student reports. This indicates that these are robust effects and that including only students’ reports of their educational experiences when predicting grades might be a misspecification because teachers’ interpretations of student behavior and the translation of those perceptions into evaluations is of utmost importance in determining grades.
The first subset of explanatory variables represents social interactions. The two variables chosen to represent students’ reports of their social interactions are whether or not they feel put down by teachers and whether or not teachers listen to them. Both of the student reports of social interactions have weak to moderate relationships with grades for both types of courses prior to entering teachers’ reports. The analyses indicate that students who feel put down by teachers get lower grades and those who report that teachers listen to them receive higher grades. After adding teachers’ reports, all of the effects are reduced but remain somewhat significant, except students’ reports of whether or not teachers listen to them. The relationship between that variable and English grades is eliminated by the addition of teacher reports.

The presence of significant relationships between students’ reports of interactions with teachers and grades was not expected. These interactions were predicted to work through other behaviors to impact grades, similar to teachers’ reports of students’ social interactions. The continued strength of these variables may be due to the nature of the variables in that they represent the quality of classroom interactions while the teachers’ reports address social interactions more broadly. Rather than simply accounting for interactions, students are reporting on how they feel about the interactions. Their perceptions of and, perhaps more importantly, their subsequent reactions to teachers may have durable effects that directly impact grades or that work through behaviors that are not accounted for in the model to impact grades.

The next group of variables includes students’ reports of classroom engagement behaviors that can be measured objectively. The four student reports included in this section are the amount of time they spend on homework (both in and out of school), how
often they go to class without completing homework, how many days they were absent from school during the first half of the school year, and how many days they skipped classes in the first half of the school year.

When the four student variables are included in the model without teachers’ reports, several of them have strong relationships with grades. For math, students’ reports of how often they go to class without completing homework are negatively related to grades as are their reports of how many days of class they missed. Reports of how often they skip class have a weak negative relationship with grades. All four variables are strongly related to English grades in the expected directions. After including teachers’ reports, all student reports are reduced and one – the relationship between skipping class and math grades – is eliminated.

As predicted, students’ reports of classroom engagement have significant relationships with grades prior to the inclusion of teacher reports. The weakened relationships between students’ reports and grades after the inclusion of teacher reports, coupled with the stability of teachers’ reports as predictors, are consistent with the theoretical understanding that teachers’ reports may be more important than students’ reports in determining grades. The findings clearly indicate that teachers’ understandings of students’ behaviors partially mediate the effects of student behavior on grades. However, more relationships were expected to be eliminated by the addition of teacher variables. This is particularly true because the behaviors which are reported in these variables represent activities that teachers are expected to record and refer back to when grading. Therefore, teachers should have objective reports to reference when factoring these behaviors into grades.
The remaining relationships between students’ reports and grades indicate that students’ reports are capturing elements of their academic performance, behavior, attitude, or effort that are not included in teachers’ reports but do impact their grades. The behaviors represented by the variables in this section provide students with opportunities to illustrate a commitment to the norms of achievement and independence. It was shown in Chapter Four that teachers reward students for illustrating this commitment through grades. But beyond the direct effect of teachers’ perceptions of student behavior on grades, student behavior can have an additional effect on grades. This effect may occur for academic reasons. Students who are often absent or do not complete homework may lack course-specific knowledge and receive a grade penalty. The behaviors mentioned above are also symptomatic of a lack of effort which may depress students’ academic outcomes. The remaining relationships may also be behavioral in that students who are absent or do not complete schoolwork may be less willing to participate in class or may act out in other ways for which teachers will penalize them.

The four student reports included to represent students’ classroom demeanor – the subjective aspect of classroom behavior – are students’ ratings of how important good grades are to them, how often they got in trouble at school, whether or not they think it is ok to disobey school rules, and whether or not it is ok to talk back to teachers. Two of the classroom demeanor variables are significantly related to math grades in the model without teacher reports. Students who report that good grades are important receive higher grades in math and those who report that it is ok to disobey school rules receive lower grades. All four variables are significantly related to grades in English in the
expected direction – although the negative relationship between students’ reports of whether or not it is ok to talk back to teachers and grades is weak. After adding teachers’ reports, the relationships between students’ reports of how often they get in trouble and their English grades is eliminated. The effects of all other variables are reduced except the relationship between whether or not it is ok to talk back to teachers and English grades, which increases slightly.

These results are consistent with the hypothesized reduction in the relationships between student reports of behavior and grades after the addition of teachers’ reports of student behavior. However, it was again expected that the relationships would be reduced more sharply or eliminated. Students’ reports regarding the importance of good grades remain particularly strong predictors of grades for both math and English courses. This finding reinforces the importance of student acceptance of the norm of achievement for good grades because students who are committed to doing well in school are advantaged in academic outcomes over academically similar peers who may not share the same level of commitment. The benefit may accrue to students’ grades through academic or behavioral factors similar to those mentioned in the section on classroom engagement variables.

One student report was included in the section representing students’ educational futures. The variable is a dummy variable indicating whether students plan to attend college in the future. The reference category includes students who do not plan to attend college or who do not know if they will attend college. The variable is not related to math grades and has only a weak positive relationship with English grades. The addition
of teachers’ perceptions of students’ educational futures eliminates the weak relationship between students’ reports and English grades.

The persistent power of teachers’ perceptions of students’ educational futures to predict grades – especially in light of the lack of relationship between students’ own plans and their grades – again indicates that this finding deserves further attention. The differences in the predictive power of reports by students and teachers may result from the bases for each group’s responses. Students are likely to overestimate their likelihood of attending college. They may report that they plan to attend college in the future because they aspire to that goal – not necessarily because they are academically prepared to achieve the goal or because they are taking the steps necessary to achieve the goal. Teachers probably base their responses on understandings of students’ actual academic trajectories.

Additional analyses (not shown) were performed to determine if specific subsets of teacher variables account for more of the reductions in the relationships between student reports and grades than others. Four models were run for each dependent variable with one subsection of teacher variables removed in each model. The results indicate that all four groups of variables play a part in mediating the relationships between students’ reports and grades. Two groups of teacher reports are more powerful than the others, though. The most influential set of teacher reports, for both math and English, was the group of variables representing classroom engagement. When these variables were removed from the models, the coefficients for several student reports gained significance compared to the final model and the models with other subsections of teacher variables removed. Teachers’ reports of students’ classroom demeanor also
proved to be powerful in reducing relationships, particularly for math grades. The results reinforce the expectation that teachers rely heavily on objectively measured indicators of students’ attitudes to factor students’ acceptance of the norms of achievement and independence into grades.

Considered together, the results described in this chapter offer support for both hypotheses outlined at the beginning of the chapter. Students’ reports of their behaviors and attitudes do have significant relationships with their grades. The addition of teachers’ reports of students’ behaviors and attitudes in the model successfully reduced many of these relationships, indicating that teachers’ perceptions of students partially mediate the effects of student behavior and attitudes on grades.

However, the addition of teachers’ reports did not completely eliminate the relationships. The remaining relationships indicate that students’ behaviors and attitudes have direct effects on their grades. These effects may operate through academic performance in that students who are not committed to educational norms gain less course-specific knowledge than their academically-oriented peers or they put less effort into coursework and do not perform well in comparison to peers. Therefore, students who do not accept the norms of achievement and independence face a double grade penalty – they will be penalized for sub par academic work and for behaving in ways that indicate a lack of commitment to educational goals. Teachers directly assess students’ behaviors and acceptance of academic norms in grades. Behaviors are then indirectly evaluated by teachers through their impact on academic work.

The results discussed above and in previous chapters highlight the complexity of grades as a measure of educational outcomes. Unlike standardized test scores, they are a
negotiated outcome based on interactions between teachers and students that account both for academic and behavioral factors. While the results thus far have helped to clarify what grades measure, questions still remain regarding how students’ behaviors and attitudes are linked to grades. The next chapter will present analyses of classroom observations in an attempt to further clarify how behavior and attitudes are translated into academic achievement and grades.
CHAPTER 7:
ANALYSES OF CLASSROOM OBSERVATIONS

This chapter presents analyses of field data collected through observations conducted in high school classrooms. The analyses expand on the existing theoretical framework and supplement the preceding quantitative analyses by addressing the intersection of instruction and socialization within high school classrooms. The examinations in this chapter support the overall goal of the research by clarifying the importance of accounting for behavior in students’ grades.

The findings presented in this chapter show that instruction and student behavior are intimately connected and difficult to consider separately when analyzing academic achievement. The classroom characteristics at issue highlight aspects of classroom instruction that are undervalued in earlier chapters. They indicate that learning is a complex process that is affected by variables related to the teacher, the student, and the classroom situation. While the findings in earlier chapters point to the fact that teachers do account for students’ behaviors in grades, these analyses indicate that doing so is appropriate and necessary by highlighting the importance of behavior for learning.

The analyses capture the importance of student classroom behavior by focusing on moments when teachers do not have control of instruction. Teachers may lose control of instruction or they may ask students to take control in an effort to meet educational goals. During these times, students’ behaviors and reactions to the situations have the
most potential to impact their own and their classmates’ academic outcomes. The effects of students’ behaviors on academic outcomes are mediated in part by the classroom composition and by teachers’ reactions to students’ behaviors. Moments when teachers do not have control of instruction illustrate the consequential nature of student behavior for academic outcomes and highlight the important role of teachers in negotiating the interaction between behavior and instruction within the classroom and accounting for behavior in grades.

7.1 Data and Methods

The field notes come from approximately forty hours of observations and interviews in a public high school located in a mid-size Midwestern city, which will be referred to as City High School. Three English classes were observed throughout the second half of the spring semester – referred to as Courses A, B, and C. Course A was a regular-level sophomore course. The teacher was an African American woman who was approximately 35 years old and will be referred to as Ms. Gorseth. The class enrollment approached 30 students and 25 were in attendance most days. Slightly over half the students were African American and one student appeared to be Indian American. The remaining students appeared to be European American.

Courses B and C were taught by the same teacher - a European American woman who appeared to be in her late 40s and will be referred to as Ms. Farmer. Again, approximately 25 students were in attendance in each course on most days. The majority of the students in the courses appeared to be European American with one or two African American students and two to four Asian American students in each class. Both courses
were Advanced Placement/International Baccalaureate (AP/IB) courses designed for juniors.

AP courses are high-level academic classes meant to prepare students for college-level courses and to pass tests which allow them to receive college credit. The IB students in the courses were participants in a magnet program run by City High School that drew students from throughout the urban area. In order to receive an IB diploma, students participated in a rigorous course of study that involved a required series of courses and a lengthened school day. Therefore, although the AP students in Courses B and C were intelligent and capable, the IB students illustrated a greater commitment to meeting academic goals and are assumed by the researcher to be more dedicated to educational norms and to value academic success more than their peers in the AP program.

The compositions of Courses B and C differed in terms of the academic program of the students. Course B had six IB students and the rest were AP. In contrast, Course C had a majority of IB students and only 4-6 AP students. Consequently, the students in Course C, on average, were considered more capable and motivated than those in Course B because they were more likely to be enrolled in the IB program. Ms. Farmer provided support for this contention because she observed several times during informal interactions with the researcher that she believed the students in Course C were brighter and more dedicated than those in Course B. Therefore, the supposed abilities and levels of dedication to educational success of the students increased progressively through each course – from A to B to C.
Each course was observed approximately 12 times. During visits to the field, the researcher focused primarily on interactions within the classroom and the interplay between instructional time, students’ behaviors, and teachers’ reactions to students. The researcher recorded field notes on these dynamics after visits to the field. Informal interactions between the researcher and the teachers or students were also recorded. In addition, each teacher was formally interviewed as were a small group of students who volunteered for interviews.

Consistent with the research presented in earlier chapters, student behavior is defined broadly as the actions and comportments of students within the observed classrooms for the purpose of analyzing field notes. While the quantitative analyses were limited in their scope by the available survey questions in terms of measuring students’ behaviors, the observations captured a full range of behaviors that are pertinent to achievement. These can include, but are not limited to, actions that worked toward achieving academic goals, conduct that prohibited instruction, or student refusal to engage with classroom processes. All types of behaviors, whether conceived of as action or a failure to act, were found to have the potential to alter instructional outcomes.

Although no math classes were observed, the earlier quantitative analyses indicate that there are few significant differences between math and English courses in terms of which behaviors are accounted for in grades. Therefore, it is not expected that observing math courses would have revealed significant differences. However, future observations of math courses may be beneficial in that they could confirm the similarities across courses or reveal differences that are not captured in the broad categories used to represent student behavior in the quantitative analyses.
The field data were analyzed using the extended case method (Burawoy 1991). This method relies on participant observation to build on existing theory and strengthen it. This is done by looking at anomalies in the observations that are not explained or are omitted from the theory and working to explain them. Through this process, the framework is strengthened and can account for a wider variety of observations.

7.2 Theoretical Framework for Analyzing Observations

In order to pursue the extended case method, it is important to re-visit the theory presented in Chapter Three which focuses on two primary goals of education – instruction and socialization. In particular this chapter will focus on strengthening Section 3.2 of that chapter, which addresses the intersection of instructional and socialization. While instruction is expected to vary across classroom settings because of ability grouping and student-teacher interactions, behavioral expectations are anticipated to be relatively constant across settings because there is a minimum behavioral standard that must be met by students in order for them to benefit from instruction. The requirements of instruction and socialization intersect because well-socialized students often behave in ways that allow teachers to advance instruction and provide opportunities for the students to gain academic knowledge.

Quantitative analyses in previous chapters provide support for these assertions and indicate that well-behaved students are rewarded through grades both because their behaviors facilitate instruction and because students are likely to benefit academically from their own behaviors. The analyses in Chapter Four show that students who engage in teacher-approved behaviors receive higher grades than academically similar peers who
do not behave in approved ways. Analyses in Chapter Six illustrate that students who behave in approved ways are likely to do better academically in terms of course-specific content and assessments than peers who do not behave well.

The observations also provide support for elements of the theory described in Chapter Three and the findings of prior quantitative analyses. Although instruction differed across the courses – both as a function of the academic abilities of the students in the courses and of the individual styles of the teachers – behavioral standards set by the teacher did prove to be similar in all three courses. Both teachers required that students pay attention, remain quiet during lectures, and complete assignments or the students could face grade penalties. Basic rules such as not talking to friends during instruction, no cell phone use, and treating others with respect were in force in both classrooms. The similarities in behavioral standards across classrooms are in keeping with the expectation that teachers would require similar behaviors despite instructional differences.

The majority of students in all three classrooms appeared to be well-socialized in that they frequently met behavioral demands and indicated that they valued educational pursuits. The descriptive statistics from Chapter Four also illustrate that teachers in the quantitative sample reported that most students engage in behaviors that indicate they are dedicated to achieving educational goals. Students in Course C participated in breaches of behavioral standards less frequently than students in Courses A and B. This is in keeping with the supposition that students who are more dedicated to educational pursuits also are better behaved – most likely because the behavior is a result of the commitment to educational goals that led them to enroll in the IB program. However, it is important
to note that the majority of students in all three courses met behavioral standards and appeared to have some level of dedication to meeting academic goals.

The theory and findings noted above provide an initial basis for examining the intersections of instruction and socialization. However, they fail to consider fully two aspects of classroom realities that were made clear through the observations and that directly relate to students’ behaviors and their impacts on academic outcomes. The first is that teachers do not always have control of instructional processes – either because students take control or the teacher relinquishes control to the students. While the theory acknowledges that teachers may need to negotiate with students to regain control of the classroom at times, it does not consider that teachers may willingly allow students a measure of control of instruction. By exploring the reasons behind and the outcomes of teachers’ loss or abdication of instructional control in this chapter, the theory regarding the production of grades will be stronger and better able to assess the use of grades as an indicator of both instructional and behavioral outcomes.

The second classroom reality undervalued in the theory is the presence of several students within each classroom. While this may sound obvious, that facet of the classroom environment is minimized in the theory and analyses which tend to focus on interactions between individual students and teachers. The quantitative data used for the analyses are not well-suited to understanding how classroom context can impact academic outcomes. The qualitative data provide a supplement that can offer insight into that aspect of the academic environment and how it relates to students’ academic opportunities and, therefore, grade production.
The contextual nature of classroom instruction is seen most clearly during moments when the teacher does not have control of instruction. During these times, students’ behaviors and reactions shape the classroom environment and the direction of instruction. For that reason, students’ actions have the most potential to impact their classmates’ learning opportunities in these moments. This provides further impetus to focus the analyses on situations when teachers do not control instruction.

7.3 Explanations for Teachers Not Having Control of Instruction

Understanding why teachers may not have control of instruction is a necessary first step in understanding the potential behavioral outcomes of these situations. For the purposes of these analyses, instructional time is defined as time when the teacher and students are meant to be focused on academic topics relevant to the course and/or working to complete assignments or projects for the course. Examples of non-instructional time would be any situation in which the teacher is not discussing course topics and has not provided the students with academic assignments to complete. For instance, the first few minutes of class are often non-instructional time because the teacher must take attendance and complete other administrative tasks and may not focus the students on academic tasks.

Frequently, teachers do attempt to control classroom instructional time and enforce limits on students’ behaviors. They ask that students arrive on time, sit in a desk, and pay attention so that they can benefit from instruction. Class time is often spent with teachers lecturing or going over materials. However, there are regular instances when teachers do not have control of instruction. These moments are revealing in terms of how
and why behavior is controlled in the classroom and the impacts of group context on individual students.

The most commonly cited and discussed example of a lack of teacher control in the sociology of education literature is an accidental loss of control (McFarland 2001; Hurst 1991; Cusick 1973). Examples of teachers accidentally losing control to students during instruction would be if students disrupted a lecture, talked to other students during instruction, failed to pay attention, or left the classroom. This is experienced by teachers as an unsanctioned loss of control that impedes instructional plans.

This chapter makes a contribution to the understanding of classroom social dynamics by examining a second situation in which teachers abdicate instructional control and ask students to guide instruction. Certain instances of abdications of control may be conceived of as student-centered instruction. However, that conceptualization does not fully describe the technique from the teacher’s perspective. It does not provide a strong impetus for understanding why teachers would turn control of instruction over to students and how they would manage the situation. Therefore, these analyses will focus on teachers’ decisions to abdicate control of instruction and their methods for managing the behavioral and academic outcomes of the abdication.

Abdication of instructional control was a regular component of classroom instruction in the observed classrooms. Common examples would be asking for volunteers to read, opening a discussion up to student participation, requiring student-led presentations and having students do group or individual work during class time. The situation may not be perceived as a lack of control by observers – particularly if it has a positive outcome in that students reach the goal defined by the teacher. However, in all
of these situations, the teachers asked students to play a role in guiding instruction while removing themselves from the process to a degree.

In all of the instances of abdication of control mentioned above, teachers desire an appropriate instructional outcome. Specifically, they have set an academic goal – whether they have made it explicit to the students or not – and they hope that the students will take action to achieve it. In addition, teachers have not relinquished behavioral control of the classroom at these times. Behavioral standards vary based on the situation, but they are active at all times and teachers may enforce the standards despite the fact that the teacher is not guiding instruction.

For example, when students are working individually to complete projects in the classroom, the teacher has set an academic goal of finishing a certain amount of work by the end of the time designated for the assignment. However, the teacher is allowing the students to determine how they will accomplish the work by not guiding them through the assignment as a whole-class project. During these times, students are expected to work diligently to meet the goals and ask for assistance from the teacher if necessary. While the teacher is not directly controlling student progress, he or she may intervene if the student does not exhibit appropriate behaviors. For instance, if the student is talking to another student or has his or her head down on a desk, the teacher is likely to ask the student to change the behavior because the student cannot achieve the academic goals while exhibiting those behaviors.

Because the classroom is often conceived of as a location where teachers desire to maintain control over students and manage instruction, it is important to understand why they would choose to allow students to have control over instructional endeavors. In
order to explore this, it is helpful to first understand teachers’ educational goals for
students. Teachers are required to manage a large group and help each member navigate
a set of obstacles and arrive at the final destination – which is, at the most basic level, the
successful completion of the course. Ultimately, it is in everyone’s best interest to arrive
at the final destination.

However, some students will not be capable of or committed to reaching the final
goal. The teacher is likely to try to help all students succeed, though. While it is possible
that some teachers do not have this sense of responsibility, that was not the case in the
observed courses in which the teachers appeared to be dedicated to achieving success for
every student. Ms. Farmer noted that she tries, “to give [students] ways not to fail.” Ms.
Gorseth behaved in ways that indicated the same level of dedication to helping students
successfully complete the course. This commitment means that teachers must put effort
into meeting the varied needs of each student in their classes.

The classroom observations indicate that the ideal method for achieving the goal
of successfully navigating course requirements, given a situation where not all members
of the group are dedicated to or capable of achieving the goal, is for teachers to be able to
give up control and have the students control themselves on occasion. From a
pedagogical perspective, teachers may ask students to guide instruction in order to
encourage student investment in learning and responsibility for academic outcomes.
Sociologically, teachers abdicate control of instruction for two primary reasons as noted
through classroom observations. The first involves resource limitations and classroom
management. Teachers have a limited amount of time and energy and they are working
with students with varying academic abilities and energy levels. In order to work within
these constraints and seek success for all students, the teacher may make instructional choices that involve students working on their own. This will allow teachers to utilize available resources most appropriately.

For instance, Ms. Gorseth frequently chose to have students work in groups. During these times, she was free to move from group to group and provide specific instructional help that met the needs and ability levels of each student. These situations also provided an opportunity for Ms. Gorseth to accomplish other tasks such as helping students who have missed class or discussing grades with students. Group and individual work time maximizes resources compared to a situation in which the teacher is trying to address differences in abilities and levels of preparation during whole class instruction. It was also observed that group work can provide a method for managing time constraints because the teacher and students can work to achieve separate, but necessary, goals during these times.

The other primary reason teachers give up control of instruction is because it provides them with an opportunity, other than formal assessments such as tests or homework, to examine students’ academic abilities and commitment to educational norms. For instance, by allowing students to guide discussion, teachers can assess whether students have done relevant readings and how well individual students understand them.

An incident in Ms. Farmer’s class illustrates the importance of this technique for providing feedback on student progress. During one class period, Ms. Farmer chose to forego discussion because several students reported that they had not completed the readings. While she could have continued with her plans for discussion and relied on the
students who had done the reading, she indicated that the activity would not have been as beneficial because she would not have had an opportunity to assess student understanding of the material. In addition, she was aware that the discussion would have been anemic and not benefited the students as much as it would in ideal circumstances.

7.4 Determinants of Students’ Classroom Behaviors

In all of the situations of student-guided instruction described above, the academic outcomes depend on how students behave. This is where socialization and the educational norms discussed in earlier chapters play a large role. Primarily because of variations across students in the acceptance of educational norms and dedication to educational goals, individual students will be more or less likely to try to take control during instruction and will react differently to teachers’ abdication of instructional control.

Student behavior can vary based on several variables other than commitment to educational norms and academic achievement, including general mental and physical state and current personal situation. Because of the number of variables involved, students can behave differently on a day-to-day basis. However, in the field notes, disruptive or uncooperative students were regularly disruptive and uncooperative while well-behaved students rarely breached behavioral standards. This indicates that the primary determinant of behavioral outcomes was consistent. Student acceptance of the norms of independence and achievement is a relatively stable factor whereas the other factors mentioned above are subject to variation and could lead students to behave
differently from day to day in the classroom. Therefore, acceptance of educational norms appeared to be the main predictor of student behavior.

The importance of the acceptance of educational norms in determining behavior was illustrated in several ways. Across courses, more students misbehaved or were reprimanded by the teacher in Course A than Course B, and in Course B than in Course C. This pattern is consistent with the educational demands of the academic programs represented by each course in that, as the academic program became more demanding, students were less likely to behave in ways that the teacher considered inappropriate.

The importance of the acceptance of educational norms in determining behavior was also illustrated within classes because of similarities in student behaviors across times. Informal interactions with students indicated that those who had college aspirations and valued academic success were less likely than their non-college bound peers to act out or misbehave. Two students – Marcus in Course A and Eddie in Course B – were noted as standouts in the field notes because they engaged in inappropriate behaviors during the majority of observed class periods. While Marcus’ educational aspirations were unclear, his unwillingness to complete assigned work indicated that he did not value the academic experiences provided in class. Eddie informed the researcher that he planned to attend a vocational postsecondary school and did not believe that an English course would benefit him. He did indicate on several occasions that he enjoyed the class and liked the teacher. In addition, the teacher informed the researcher that Eddie was intelligent and capable. However, he had no specific reason to perform well in the class, either academically or socially, and his behavior was not constrained by his commitment to educational goals.
7.5 The Impact of Students’ Behaviors on Instruction

The appropriateness of students’ behaviors, as primarily determined by their commitment to educational goals and classroom norms, is an important determinant of academic outcomes. For the purpose of this research, appropriate student behaviors during moments when teachers do not have control of instruction are defined as actions that teachers approve and that move instruction forward. As noted earlier, what is considered appropriate by the teacher varies according to the situation. For instance, an appropriate reaction to classroom discussion is for students to participate and indicate that they are engaged with the topic being discussed. An appropriate reaction to individual time for work in class is for students to quietly pursue the completion of work. While these behaviors are obviously different, they are appropriate responses to the task at hand and would indicate that students are both aware of and dedicated to meeting classroom demands.

Appropriate behaviors elicited by an abdication of control or accidental lack of control lead to different situations. Appropriate behavior during an abdication of control leads to the ultimate goal – student-guided learning. An example of this includes when students are engaged, active, and on topic during discussions. During times designated for group or individual work, appropriate behavior involves students actively working to complete assignments. At these moments, students are taking responsibility for their learning outcomes and showing interest in meeting academic goals set by teachers. By doing that, they are moving instruction forward and assisting the teacher in meeting goals.
On one occasion, Ms. Farmer had to leave the classroom to tend to an administrative task. She asked the students to begin reading the play they were working on in her absence. For several minutes after she left, the students failed to follow this directive. Eventually, a student named Paul took the initiative and announced, over the noise of the classroom, “We’re at the top of page 101.” He then began reading his part. At this point the other students joined in and student-guided instruction began. By the time Ms. Farmer returned to the classroom, the students were focused and on task. Because of Paul’s assistance and desire to achieve the goal of finishing the play, Ms. Farmer was able to attend to other tasks and still meet her instructional goals for the day.

A situation with an accidental lack of control and appropriate behavior is rare. It may involve a student disrupting instruction to address an element of instruction that the teacher was not going to cover or had not reached yet. These can be conceptualized as teaching moments during which the student takes control of instructional trajectory to a degree. They provide teachers with an opportunity to explore the material as it pertains to students’ interests but simultaneously disrupt existing plans for instruction.

The teacher may react positively to student curiosity and interest, as Ms. Gorseth did when Alex interrupted instruction to ask why a character in a play was talking to himself and to comment that it was strange. Ms. Gorseth noted that this was an “excellent observation” and spent a few moments discussing the conventions of playwrights. She then quickly returned to the pre-planned instructional trajectory. While she appreciated Alex’s curiosity, his behavior was unlikely to lead to instructional benefits from her perspective because it occurred during a moment when she had control of instruction and did not plan to discuss the topic that Alex raised.
Inappropriate behavior is defined as actions that are not approved by the teacher and do not facilitate instruction. This is the type of response to a teacher’s lack of control that researchers often discuss because the situation is identified through the inappropriate behavior outcome. The stereotype is students behaving inappropriately during an accidental loss of control. For instance, students may disrupt instruction to discuss non-academic topics, talk to other students during instruction, or arrive late for class and interrupt instruction. These behaviors can lead to distractions and sidetracking and are perhaps the most damaging in terms of instructional goals. They may prohibit instruction and temporarily or permanently derail instruction for the class period.

Marcus, a student in Ms. Gorseth’s class who regularly behaved inappropriately, brought instruction to a halt for ten minutes during one class because he received a text message from Darlene, another student in class. His phone rang during a lecture to indicate he had received the message. Ms. Gorseth had a strict policy regarding cell phone use in class and told Marcus that he had to give her his phone for the rest of the class period. He refused and they negotiated for several minutes until he finally relented. While the other students in class were clearly interested and engaged in this exchange, they also lost opportunities for instruction both during the exchange and while Ms. Gorseth worked to establish where they had left off prior to the incident.

Inappropriate behavior during an abdication of instructional control can be similar to the behaviors described above. These behaviors include students not participating or refusing to stay on topic during discussion or students not focusing on academic work during individual or group work time. During these times, instruction does not progress and instructional time is lost.
An example of this took place in Course B when Ms. Farmer, in an effort to conserve time, opted to have students begin reading and discussing a play on their own so she could work with a student who needed assistance. No one took action to start the reading during that time and the instructional time was lost. Upon returning her attention to the whole class, Ms. Farmer sarcastically commented, “Is this how we start independently? I’m so proud.” She was clearly frustrated with their unwillingness to assist her in instruction because their actions (or lack of action) led to spinning wheels – that is, instruction was not progressing.

7.6 Management of the Diffusion of the Effects of Students’ Behaviors

As shown above, moments when teachers do not have control of instruction can have positive or negative outcomes in terms of students’ willingness to guide or participate in instruction. These outcomes affect academic opportunities for students and illustrate the importance of student behavior in the classroom. While individual students are impacted by their own behaviors in terms of academic outcomes at these times, that behavior can also diffuse to classmates and affect their opportunities to learn.

During a test in Ms. Gorseth’s class, several students were behaving inappropriately by talking and moving around. Natasha, a student in the class, commented, “They need to be quiet because I cannot concentrate.” While it was unclear whether she was simply vocalizing her frustration or if she expected the disruptive students to hear her, this provides a basic example of how the behaviors of students can potentially impact other students’ academic outcomes, even if they are not participating
in the behaviors. Therefore, both individual and group dynamics matter within the classroom.

The potential academic impacts of students’ behaviors differ based on whether the class is in a state of full attention or divided attention. While teachers cannot accurately predict when they may accidentally lose control of instruction, they must account for the possible behavioral consequences, and subsequent academic consequences, of an abdication of control when deciding if and how they will relinquish control. Teachers appear to account for these possibilities in their decisions regarding abdications control of the classroom, as described below.

If the teacher gives up control of instruction during whole-class activities, all of the students in the class may receive rewards or penalties based on the behavior of other students in class. The most common situations in which this occurred in the observations was during discussion or student-led presentations of academic material. The students who participated or paid close attention were expected to benefit the most from these activities because they are directly involved and working through the material and are receiving the grade benefits of participation. However, if several students responded appropriately and participated in discussion, all students could benefit academically from the student-guided instruction even if they did not participate because they were exposed to discussion of the material covered in class.

Localized disruptions to discussion occurred when two or more students talked to each other, off topic, during discussion. The students participating in the inappropriate behavior were not benefiting from instruction at that moment. The impact of this inappropriate behavior could also affect other students in several ways. If the talk was
loud enough, all students might hear it and be distracted. The distraction may also be limited to students who were physically closest to those who were behaving inappropriately. The impact of the behavior also was determined in part by the teacher’s reaction. If the teacher chose to address the inappropriate behavior, it became a distraction for the whole class. If the teacher did not react, the impact might remain with only a few students. Teachers’ responses to inappropriate behavior will be discussed in more detail in the next section.

Ms. Farmer regularly utilized discussion as an instructional tool in her classroom. She appeared to choose this technique because it provided her with an opportunity to gauge students’ progress as is illustrated by the example mentioned earlier when she chose not to have students discuss a reading because many had not finished it and she knew the activity would not provide a gauge of their progress. In addition to the opportunity to assess student progress during discussion, the students in both Courses B and C were generally engaged and current with the material for discussion (the previous example is an exception to that rule) so there was low risk of inappropriate behavior during a discussion.

However, the same students were eager to participate in discussions on a regular basis which reduced Ms. Farmer’s opportunities to gauge the academic progress of the less enthusiastic students. She tried to overcome this problem by declining to call on frequent contributors and requesting that other students participate. She regularly made comments such as “Let’s get somebody whose voice I haven’t heard already.” or “Let’s hear a new voice.” This strategy was generally successful in that other students would volunteer to participate in discussion. Increased participation provided her with an
opportunity to assess additional students’ progress. When discussion did not go well, she easily switched to another activity such as lecture or individual work to allow the students to catch up with the material. Overall, this type of abdication of instructional control was relatively low risk and could provide rewards by engaging students with the material.

If students were working individually or in small groups, appropriate and inappropriate behaviors were different than if the whole class was participating in an activity. In these situations, talking to other students was generally acceptable and failure to engage with the assigned activity was unacceptable, while students could often refrain from joining in during discussion and not be reprimanded. The impacts of behavior were also different than if the whole class was participating in an activity. Inappropriate or appropriate behavior might impact only the individual student or the students in the small group where it was taking place. If students were actively disruptive, it could diffuse to other groups through distractions. Also, disruptive students might command the teacher’s attention, prohibiting him or her from providing help to other students.

However, individual or group work was a good method for alleviating some of the potential negative effects of students’ behaviors. Teachers can effectively utilize scarce resources such as time and individual instruction during these situations so the potential risks of inappropriate behavior during divided attention are reduced. Committed students can benefit more from individual or group work than they might from whole class discussion because they can work at their own pace and not be impeded by less capable or committed students. In addition, teachers can use this time to perform necessary tasks such as helping students catch up on assignments or providing guidance on assignments that must be accomplished individually. Because these tasks could not have been
completed during whole-class discussion, the risks of student non-cooperation during group work are reduced because there are no other viable means for achieving the goal.

Ms. Gorseth utilized group work more often than Ms. Farmer. This may be because the students in Ms. Gorseth’s class were more diverse in terms of academic abilities and educational commitment than the students in Ms. Farmer’s class. For this reason, she utilized group work as an effective method for personalizing instruction and conserving time and limiting the potential negative impacts of inappropriate behaviors. In addition, observing the groups provided her with an opportunity to judge students’ academic progress and motivation.

If Ms. Gorseth had used whole-class instruction during these times, she would have forced the capable students to languish and would have spent the majority of her time addressing the concerns of the least capable students. This would have wasted resources and maximized the negative impact of the least capable and most disruptive students on the other students. During group work, the least dedicated students may have made no or little progress, but their negative impact on more committed students was reduced.

7.7 Teachers’ Reactions to Students’ Behaviors

The final aspect of the intersection of instruction and socialization during moments when teachers do not have instructional control is how teachers choose to react to students’ behaviors. As was mentioned earlier, behavioral standards remain in force during times when teachers do not have instructional control, whether through accident or abdication. However, teachers can choose to selectively enforce these standards. While
some behaviors, such as fighting and yelling, were never tolerated and always elicited a response from the teachers, others provided more leeway and the teacher could decide if and how to react. These decisions often appeared to be based more on the situation and context of the behaviors than on the actual behaviors themselves.

When students react appropriately to teachers’ abdication of control, the teacher is likely to continue the situation because it meets instructional needs and helps students to meet the teacher’s academic goals. If students react appropriately to an accidental loss of control, the teacher may address it, but is unlikely to change instructional direction to continue it. Again, this situation is rare.

Teachers can react to inappropriate behavior in several ways. One common method is to ignore it. Ignoring the behavior can minimize the potential disruption. If students are talking, it may bother the students around them. But if the teacher says something during instruction or presentations, it disrupts the whole class.

Ms. Farmer regularly ignored behaviors that were clearly disruptive and inappropriate. She tended to choose this reaction at times when few students seemed to be involved or aware of the behaviors and when the instructional activity was relatively low-priority. For instance, during a discussion, Eddie (who regularly behaved inappropriately) and Jim went to the front of the classroom and had a mock swordfight. Ms. Farmer did nothing to acknowledge this activity and continued to listen to student-led discussion. She may have chosen to react in this way because the non-disruptive students appeared to be engaged in discussion and few seemed to pay attention to the disruptive activity. In addition, the discussion at the time was relatively low priority because it was supplementary to matters that had already been discussed. As will be
discussed below, in other situations she reacted more strongly to behavior that could be
defined as less disruptive and inappropriate. Her reaction appeared to be determined
more by the context of the situation than by the behaviors.

If teachers choose to react in an obvious way, there are two main options. They
can react in a way that will prohibit instruction from continuing or negatively impact
students’ academic outcomes if they do not behave appropriately. This is a zero sum
reaction. Conversely, teachers can negotiate with students or work to lead them to an
understanding of the situation. This is shepherding. It is similar to negotiation because
the teacher is trying to help the student reach an understanding of the reasons the
behavior is inappropriate. This is when the teacher works with the student as a partner to
reach a satisfactory conclusion and advance educational goals.

The two reactions can be combined at times. Teachers will indicate that
instruction will not continue or the student(s) will face academic penalties if the behavior
continues, but they will also work to help the student(s) gain an understanding of why it
is important that they change their behaviors. This may be done because the
inappropriate behavior occurs during an activity that is high-priority, thus eliciting the
zero sum reaction. However, zero sum reactions do not provide students with motivation
to complete the task other than fear of a reduced grade. By including the shepherding
element in the reaction, the teacher can hope to instill more intrinsic motivation in the
students by providing them with additional reasons to work toward the academic goal.

Teachers respond to inappropriate behavior with zero sum replies during an
accidental loss of control in order to reinforce behavioral limits. This may occur when
the behavior is known to be inappropriate and offering a more detailed explanation of the
problem would simply be repeating what the students already know and would be unduly disruptive to the task at hand.

For instance, students in Course C did presentations on an outside reading book. During one of the presentations, Clay and Tim were talking to each other. While the talk was loud enough that those around them could certainly hear it, it was unlikely that it was disruptive for the entire class. Ms. Farmer reacted by turning to them and saying, “Shut up!” in a voice that was loud enough for the whole class to hear. This reaction was clearly different than the one described earlier in which she ignored the seemingly more disruptive mock swordfight.

The different reactions can be explained based on the situation. This situation was more high-stakes than the one described earlier because it occurred during student-led presentations. The students were graded based on these presentations and distractions could reduce their grades which may be why Ms. Farmer felt compelled to respond. However, she did not elaborate on her response and shepherd the students, most likely because it was a behavior that was obviously inappropriate and elaborating would have taken up more time and further distracted the already nervous students who was presenting. By speaking loudly enough for everyone to hear her, she also took advantage of the opportunity to let the entire class know that breaches of behavioral standards would not be tolerated during presentations. If Clay and Tim had chosen to talk to each other during a class discussion of a low-priority topic, Ms. Farmer may have been more likely to simply ignore the behavior, as illustrated by the swordfighting example.

Teachers react with zero sum responses during an abdication of control if resources are limited or need to be conserved. This is also often the reason that they
decided to abdicate control. For instance, time may be limited and students are behind on assigned work. In that situation, teachers would more strictly enforce behavioral standards to ensure that resources are used appropriately than if they had abdicated control for other reasons.

Ms. Farmer abdicated control on one occasion so that she could have time to work individually with students and discuss their grades with them. In particular, she was meeting with students who were failing or were in danger of failing the course. During the abdication of control, the other students in the classroom failed to work on assignments and quickly began to talk to each other about non-academic topics. Ms. Farmer reacted by saying, “People, you aren’t being quiet enough. I gave you this time to get some work done because you are behind. And those of you who are most behind are the ones making the most noise.” Her response and disapproving tone indicated that academic outcomes were at stake and that she was not going to permit students to openly continue disregarding this opportunity to complete assignments.

Ms. Gorseth had a similar reaction when she scheduled times for students to do in-class presentations. She told the students that the deadlines for the presentations were firm and could not be renegotiated even if they were not prepared. If the students did not do their presentation at the scheduled time, they would not have an opportunity to do it later and receive partial credit. This response to students’ lack of preparation was based on the reality that there were no other times when the students could do the in-class presentations. If there had been more leeway in the schedule, she might have offered opportunities to make up the work as illustrated by other times when she allowed students to do make up work and receive at least partial credit.
Despite the need to react with zero sum consequences in many situations, teachers probably desire to react with shepherding in most situations. This is because shepherding is a more positive reaction that is less likely to alienate students and is more likely to encourage them to behave appropriately in the future. Teachers react with shepherding during an accidental loss of control when there are resources available and in order to motivate or strengthen relationships with students. They do this because of the possible long-term benefit that students may become better partners in the education process.

Ms. Gorseth frequently used shepherding during accidental losses of control. The observations indicate that she may have used this technique more often than Ms. Farmer because Ms. Farmer was faced with fewer losses of control. However, Ms. Farmer was also more likely to ignore inappropriate behavior, perhaps because the students in her classes were more dedicated to pursuing academic success so she did not need to enforce behavioral standards as strictly because inappropriate behaviors were less likely to spread to other students.

An example of Ms. Gorseth using shepherding to establish motivation in a student occurred on a day when she had abdicated instructional control so that students could work on individual assignments. She briefly regained control of the classroom in order to explain certain aspects of the assignment. During this time, a student continued to talk to a partner. Ms. Gorseth solicited her attention by pleasantly stating, “It is so important that you listen right now.” She then went on to explain why the activity they were completing at the time would benefit them in future classes. She was working to establish a larger goal for the academic work beyond just earning a good grade for the
assignment. She was also providing students with a reason to pay attention to the instructions and successfully complete the assignment.

Ms. Gorseth also mixed shepherding moments with zero sum consequences on occasion. Marcus interrupted a whole-class review session to say, “I have a question off the topic. Today is the last day to make up the test?” Ms. Gorseth replied that it was and that he would receive a grade of zero if he did not complete the test that day. This was a zero sum reaction to Marcus’ absence on the day the test was administered and his failure to make the effort to take the test within the allotted time. Marcus continued by asking, “Why do we have tests? I hate tests!” Ms. Gorseth replied “To show you and me what you have learned.” Marcus’ friend, Jimmy, joined in with her as she said this and Marcus also joined in, albeit sulkily, on the last few words. Their ability to parrot the words indicated that she had repeated this mantra frequently in an effort to motivate students to value examinations.

In both of the examples presented above, Ms. Gorseth chose to use shepherding despite the fact that she could have responded with zero sum consequences regarding students’ inappropriate behaviors that disrupted instruction. The situations indicate that she opted for shepherding because the surrounding instructional tasks were low-priority and could be interrupted for several minutes without serious losses. In addition, both instances of shepherding potentially provided benefits both for the individual students and for the rest of the class because they were an opportunity to reinforce intrinsic motivation for academic success.

Shepherding with abdication of control, similar to an accidental loss of control, occurs when teachers want to work toward developing instructional or behavioral goals.
In these situations though, the reaction is often directed specifically at encouraging students to take responsibility for their own work during the abdication of control.

Ms. Farmer’s efforts to encourage students to participate in discussion, as mentioned above, are an example of this type of activity. While she had relinquished control of discussion to the students, she would intervene in an attempt to guide who responded if she felt that not enough students were benefiting from the opportunity. While she did not completely refuse to let frequent contributors speak and she did not go so far as to call on specific students who rarely participated, she worked to ensure that students were engaged and to explain to students the importance of participation both for themselves and their fellow students.

The responses described above are the endpoints of negotiations between teachers and students regarding discrete instances of classroom behavior. These negotiations are undertaken repeatedly within classrooms. They illustrate the potential roles of outside forces and resource constraints in determining teachers’ responses to student behavior, indicating that academic outcomes are determined, in part, by a variety of factors that are not directly related to instruction. The negotiations are consequential because the examples illustrate that teachers’ reactions to students’ behaviors are important determinants of the academic and social consequences of the behaviors.

7.8 Conclusion

The analyses presented above extend the existing theoretical framework presented in Chapter Three by providing nuanced examinations of how behavior can impact academic outcomes in high school classrooms and why it is important for teachers to
attempt to control behaviors, both within the classroom and through formal feedback like grades. In particular, the observations supplement and expand on Section 3.2 of the theory, which discusses the intersection of instruction and socialization. They do this by focusing on the interactions between the two educational mandates and also expanding on that section to offer an account of contextual effects.

The observations indicate that student behavior is, at least in part, determined by acceptance of educational norms. They also illustrate how behavior is connected to academic outcomes by focusing on common situations in which behavior is most likely to affect instruction and opportunities to learn. These findings provide justification for both controlling and assessing student behavior. Teachers need to control student behavior because behavior plays a role in shaping the classroom environment and determining all students’ academic outcomes. Teachers can control behavior in the classroom by working to enforce behavioral standards. They can also attempt to control behaviors by making them a component of formal assessment, as most teachers do in grades. Students who value academic achievement and success will respond to the inclusion of behavior in grades by working to meet the standards and, subsequently, the students will play a role in ensuring successful instructional outcomes for the class.

The findings in this chapter can also be used to supplement and clarify the quantitative analyses in previous chapters. In relation to Chapter Four, the observations work to support the initial supposition that students are rewarded with grades by teachers for engaging in behaviors that may improve their own academic outcomes. The analyses of the observations attest to why these behaviors can have important impacts on academic achievement. They also provide clues to understanding how teachers account for
behavior in grades by focusing on moments of abdication of instructional control when teachers have opportunities to assess students’ academic progress through their classroom behaviors.

The analyses of test scores presented in Chapter Five indicated that students’ behaviors are not as strongly related to standardized test scores as they are to grades. This is partially because behaviors are not directly factored into test scores while they are factored into grades by teachers. But it was also hypothesized that academically-related behaviors may not be as strongly related to test scores for individuals because the effects of the behaviors on academic outcomes are somewhat diffused from the individual to the entire classroom. That is, the behaviors of individual students can impact opportunities to learn for all students within a classroom and, therefore, have an impact on all students’ results on standardized tests – similar to the way in which one student’s behavior has the potential to effect other students’ grades by mediating opportunities to learn. The qualitative analyses provide further support for this contention by emphasizing the contextual aspects of classroom instruction. Inappropriate behavior in the classroom had the potential to disrupt learning opportunities for all students in the observed classrooms. Conversely, classrooms in which the majority of the students were engaged with the material and behaving appropriately provided students with potential academic benefits. Therefore, standardized test scores may provide an indirect account of classroom behavior because they measure opportunities to learn which are determined, in part, by student behavior.

The importance of classroom context in shaping results on measures of educational outcomes prompts a consideration of the practice of ability grouping. While
teachers can work to limit the impact of individual students or groups of students on instruction, as illustrated above, the effects of behavior on instruction cannot be completely controlled or negated in modern classrooms. Because social behavior is, in many ways, driven by students’ dedication to academic achievement and norms, classrooms with students who are well-socialized are less likely to be subject to the negative impacts of inappropriate behavior and vice versa. Therefore, ability grouping is likely to group students not simply by academic factors, but also by behavioral characteristics. The classroom observations support this assertion because fewer inappropriate behaviors were noted in Courses B and C than in Course A. Although Courses B and C were technically the same class, students in Course C were slightly less likely to engage in inappropriate behavior – most likely because a higher proportion of students in that class participated in the academically-rigorous IB program.

One outcome of this type of grouping is that students in less academically-rigorous classes, such as remedial or basic courses, face disadvantages in opportunities to learn compared to their peers in more rigorous classes. While many students in these classes may be dedicated to educational norms (and the observations indicate that a majority were), they are likely to have classmates who do not share this focus and who behave inappropriately and subsequently reduce learning opportunities. The students in these classes then face a double penalty – they are already less capable or prepared than peers in more rigorous courses and they are then subject to a classroom context in which they do not receive as many opportunities to learn because of the inappropriate behavior of their classmates. The academic cycle created by these groupings may perpetuate or exacerbate existing academic differences between students in different ability groups.
While these conclusions indicate that it is appropriate for teachers to assess individual students based on their social behaviors because of the possible effects of those behaviors on students’ academic outcomes, they offer an additional rationale for teachers to take into account social behavior when computing grades. It is appropriate for teachers to reward students for behaviors that facilitate instruction even if they do not directly benefit the student academically. That may seem inappropriate because students are being rewarded simply for making teachers’ jobs easier. However, that focus also provides other students in the classroom with opportunities to learn, so facilitating instruction has additional academic benefits beyond easing instruction for teachers. Students who behave appropriately may improve learning outcomes for their peers by engaging in behavior that facilitates instruction, even if they do not directly benefit from the instruction themselves.

The next chapter will examine the predictive validity of grades for students’ future outcomes. The qualitative analyses provide important insights into the ways that those outcomes could be related to the characteristics of students measured in grades. The complex nature of the classroom and instruction, coupled with the roles of both academic abilities and behavior in determining outcomes for individual students and the class as a whole, are in many ways related to settings in which students will seek success in the future. These connections are most easily made regarding success in postsecondary education. While college classrooms are different than high school classrooms in important ways, they still require that students have a dedication to achieving educational success and behave in ways that allow instructors to provide
opportunities to learn. The characteristics of students measured by grades provide an approximation of those skills and normative orientations.

Many workplaces also require skills similar to those measured in high school grades, although the similarities between work and high school may not be as obvious as those between high school and college classrooms. Employees need to behave in sanctioned ways and work to achieve the goals of the organization. Their commitment or lack of commitment to organizational goals can impact their work performance and the ability of their fellow employees to achieve their own goals. In addition, employees will often need to be able to work independently. As illustrated in the analyses of the observations, grades can provide a direct or indirect accounting of the skills and abilities mentioned above.
CHAPTER 8:
RELATIONSHIPS BETWEEN GRADES AND TEST SCORES AND STUDENTS’ EDUCATIONAL AND OCCUPATIONAL OUTCOMES

This chapter presents the results of analyses that use grades to predict students’ future outcomes. The purpose of these analyses is to examine the predictive power of grades in light of the findings discussed in previous chapters. The current analyses will further clarify the potential uses of grades in research and educational policy.

The previous chapters have focused on developing an understanding of how grades are determined within a classroom context. The analyses indicated that grades are different than test scores in that they measure differing aspects of students’ educational experiences. Grades represent both social and academic aspects of students’ achievement while test scores primarily measure academic abilities and knowledge. For this reason, grades are expected to be related to outcomes, such as those analyzed in this chapter, that account for both social and academic factors.

The data for this chapter come from NELS:88. Descriptive statistics for all dependent variables are presented in Table 8.1. Because no teacher-reported variables were used in the analyses, the analyses contain the full sample of students from NELS:88 rather than the limited sample analyzed in previous chapters. As in previous chapters, data were imputed for all explanatory variables. Because no data were imputed for the dependent variables, the sample sizes vary for each outcome. The “micombine”
TABLE 8.1
DESCRIPTIVE STATISTICS FOR DEPENDENT VARIABLES IN THE ANALYSES

<table>
<thead>
<tr>
<th>DEPENDENT VARIABLES</th>
<th>Mean</th>
<th>Low</th>
<th>High</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Grade Average</td>
<td>7.085</td>
<td>1</td>
<td>13</td>
<td>12144</td>
</tr>
<tr>
<td>Math Grade Average</td>
<td>6.498</td>
<td>1</td>
<td>13</td>
<td>12144</td>
</tr>
<tr>
<td>English Test Score</td>
<td>32.801</td>
<td>16.97</td>
<td>51.16</td>
<td>12144</td>
</tr>
<tr>
<td>Math Test Score</td>
<td>47.971</td>
<td>14.80</td>
<td>78.1</td>
<td>12144</td>
</tr>
<tr>
<td>No College</td>
<td>.209</td>
<td>0</td>
<td>1</td>
<td>12135</td>
</tr>
<tr>
<td>Attended 2-year College</td>
<td>.253</td>
<td>0</td>
<td>1</td>
<td>12135</td>
</tr>
<tr>
<td>Attended 4-year College</td>
<td>.538</td>
<td>0</td>
<td>1</td>
<td>12135</td>
</tr>
<tr>
<td>Attained Bachelor's Degree</td>
<td>.338</td>
<td>0</td>
<td>1</td>
<td>12029</td>
</tr>
<tr>
<td>Attained Bachelor's Degree if Enrolled</td>
<td>.625</td>
<td>0</td>
<td>1</td>
<td>6477</td>
</tr>
<tr>
<td>Income in 1999</td>
<td>26628.90</td>
<td>1</td>
<td>50000</td>
<td>10441</td>
</tr>
<tr>
<td>Log of income in 1999</td>
<td>9.956</td>
<td>0</td>
<td>13.12</td>
<td>10441</td>
</tr>
<tr>
<td>Weeks worked in 1999</td>
<td>45.477</td>
<td>0</td>
<td>52</td>
<td>12144</td>
</tr>
<tr>
<td>Hours worked per week in 1999</td>
<td>40.861</td>
<td>0</td>
<td>80</td>
<td>12144</td>
</tr>
</tbody>
</table>

SOURCE: National Education Longitudinal Study of 1988

NOTE: Standard deviations for all non-dichotomous variables are presented in parentheses.

command was used for all regressions to estimate an average of the results for the five datasets created through the imputation process. However, “micombine” could not be used for the analyses that determined the goodness of fit of the models or to generate predicted probabilities. For that reason, those analyses were limited to one of the imputed datasets.

Four explanatory variables are used in the analyses. Two represent students’ English and math grades throughout high school. The National Center for Education Statistics generated the variables by averaging all English or math grades from students’ transcripts to create one measure for each subject. The other two explanatory variables are students’ 12th grade English and math test scores on the standardized tests.
administered as part of the NELS:88 survey. The control variables used in previous chapters – gender, race, and socioeconomic status – are also included in some analyses. These variables are included because they may lead to systematic variations in both the explanatory and dependent variables.

Three dependent variables are analyzed in this chapter. The first is whether or not a student enrolled in postsecondary education and the second is whether he or she successfully completed a four-year program and received a bachelor’s degree. The third dependent variable is respondents’ reports of income in 1999. It is broadly hypothesized that grades and standardized test scores will be related to outcomes for each dependent variable because success for each dependent variable is influenced by factors measured by both grades and test scores. Specific hypotheses for each dependent variable and the methods used for each group of analyses will be discussed below.

Although high school graduation is clearly an important outcome of interest, analyses are not presented predicting high school graduation due to the composition of the analytic sample in NELS:88. The variable representing high school graduation indicates that over 94% of the respondents graduated from high school or earned an equivalency degree by 2000. That outcome is not in keeping with current reports on graduation rates and illustrates the difficulty of tracking dropouts for follow-ups in educational surveys. Because of the biased nature of the sample, analyses predicting high school graduation would not be truly informative. It is also important to note that the sample for the rest of the analyses is biased toward students who successfully completed high school.
The first analyses in this chapter examine enrollment in postsecondary education. It is hypothesized that both grades and test scores will positively predict students’ likelihood of enrolling in postsecondary education as opposed to not enrolling. It is also predicted that students with higher grades or test scores will be more likely than peers with low grades or test scores to enroll in four-year programs rather than two-year programs. In terms of the relative importance of grades and test scores in predicting enrollment, both may play an important role. Grades indicate motivation and achievement orientation, so students with high grades should be more likely than others to pursue postsecondary education. However, institutional gatekeepers may focus more on test scores than grades because they provide a relatively simple method for comparing students. For this reason, students with high test scores may be more likely to be granted admission to postsecondary institutions. In addition, either or both measures may act as feedback for students and significant others when making decisions regarding postsecondary enrollment which would increase the ability of the measure to predict enrollment. Although survey participants were not informed of the results of their tests for NELS:88, the test scores are used in the analyses as a proxy for standardized test scores that students receive and might use as academic feedback in high school.

A multinomial logit model was used to predict the likelihood of enrollment in a postsecondary institution. The three categories are no postsecondary enrollment, enrollment in a two-year institution, and enrollment in a four-year institution. The initial step in the analyses was to determine the fit of various models. A model containing only grades as predictors of enrollment was estimated and another was run with only test scores. The two models were then compared to see which had a better “fit” – that is,
which can more adequately explain students’ likelihood of enrolling in postsecondary education (Long and Freese 2006). The model with test scores was judged to be stronger than the model with grades. An additional measure of fit was estimated to determine if the measures more accurately predicted college enrollment when paired than when used separately. A model that contained both grades and test scores was compared to the test score model. The model that included both measures was judged to be strong compared to the model with only test scores. This finding supports the possibility that combining both measures may be the most effective method for predicting postsecondary enrollment outcomes.

Similar analyses of fit were conducted with models that contained controls for student background characteristics along with the measures of educational outcomes. In these analyses, the model containing grades was judged to be strong compared to the model containing test scores. The change in the predictive power of grades compared to test scores across the models, based on the inclusion of background characteristics, indicates that test scores reflect aspects of students’ backgrounds that help to explain their likelihood of enrolling in postsecondary education. Therefore, when those characteristics are held constant within the model, grades are better able to predict enrollment. Similar to the model that did not utilize control variables, though, the model containing both measures was judged to be stronger than the models containing only grades or test scores. This again indicates that using both measures to predict enrollment may be the more appropriate choice than using only one. The following analyses will provide a more nuanced understanding of these findings and the differences between grades and test scores in terms of predicting college enrollment.
Tables 8.2 and 8.3 present multinomial logit models predicting college enrollment. Table 8.2 contains results for the unadjusted model which includes only grades and test scores as predictors of enrollment while Table 8.3 shows results for a model that also includes variables representing background characteristics. The results in Table 8.2 indicate that several of the measures of educational outcomes are powerful predictors of postsecondary enrollment. English grades and both English and math test scores are reliable predictors of enrollment, with English grades having the stronger consistent relationship with enrollment. Math grades do not predict enrollment except in the case of the choice between a two-year college or not attending a postsecondary institution. In that instance, math grades have an unanticipated negative relationship to the likelihood of postsecondary enrollment. The size of the effect is small but significant. It is unclear why math grades would be negatively related to postsecondary enrollment as they were expected to have a positive relationship with enrollment decisions. The effect may be the result of students basing postsecondary enrollment choices on factors that coincide with math grades so that the coefficient for math grades is representing an indirect effect rather than a direct negative effect of math grades on enrollment.

Table 8.3 illustrates why the model which includes grades and control variables was judged to be a better fit than the model containing test scores and control variables. In the model which includes background characteristics, English test scores lose significance as a predictor of enrollment and only remain moderately significant for predicting two-year versus four-year enrollment. This indicates that students’ results on standardized tests of English reflect background characteristics that are relevant in terms of postsecondary enrollment decisions. The finding provides evidence that standardized
TABLE 8.2
MULTINOMIAL LOGISTIC REGRESSION OF GRADES AND TEST SCORES ON
ENROLLMENT IN POSTSECONDARY EDUCATION

<table>
<thead>
<tr>
<th></th>
<th>2-year v. No PSE</th>
<th>4-year v. No PSE</th>
<th>4-year v. 2-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Grades</td>
<td>.127*** (.017)</td>
<td>.344*** (.017)</td>
<td>.217*** (.017)</td>
</tr>
<tr>
<td>Math Grades</td>
<td>-.045* (.018)</td>
<td>-.029 (.016)</td>
<td>.016 (.015)</td>
</tr>
<tr>
<td>English Test Score</td>
<td>.014* (.006)</td>
<td>.021*** (.006)</td>
<td>.008* (.004)</td>
</tr>
<tr>
<td>Math Test Score</td>
<td>.015** (.004)</td>
<td>.061*** (.004)</td>
<td>.046*** (.003)</td>
</tr>
<tr>
<td>N</td>
<td>12135</td>
<td>12135</td>
<td>12135</td>
</tr>
</tbody>
</table>

NOTE: Standard deviations for all coefficients are presented in parentheses.
NOTE: *p<.05; **p<.01; ***p<.001

Tests of English may lack validity in that they capture the influence of students’
backgrounds rather than providing an unbiased account of students’ academic abilities.
While previous chapters illustrated that all of the measures of educational outcomes are
influenced by students’ demographics, it appears that English test scores are particularly
prone to reflecting non-academic determinants of outcomes. The results for the other
measures included in the model are similar to those for the model which did not include
control variables. English grades and math test scores remain strong positive predictors
of enrollment although English grades have somewhat greater power to predict
enrollment when controlling for background characteristics as indicated by the goodness
of fit statistics.

The findings correspond to the prediction that students with high grades and test
scores are more likely than peers with lower grades and test scores to enroll in
postsecondary education, although the relationships between the measures of educational
TABLE 8.3
MULTINOMIAL LOGISTIC REGRESSION OF GRADES, TEST SCORES, AND CONTROL VARIABLES ON ENROLLMENT IN POSTSECONDARY EDUCATION

<table>
<thead>
<tr>
<th></th>
<th>2-year v. No PSE</th>
<th>4-year v. No PSE</th>
<th>4-year v. 2-year</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Grades</td>
<td>.102*** (.017)</td>
<td>.315*** (.019)</td>
<td>.213*** (.017)</td>
</tr>
<tr>
<td>Math Grades</td>
<td>-.037* (.018)</td>
<td>-.009 (.018)</td>
<td>.028 (.016)</td>
</tr>
<tr>
<td>English Test Score</td>
<td>.010 (.006)</td>
<td>.017** (.006)</td>
<td>.006 (.004)</td>
</tr>
<tr>
<td>Math Test Score</td>
<td>.016** (.005)</td>
<td>.050*** (.004)</td>
<td>.034*** (.004)</td>
</tr>
<tr>
<td>Female</td>
<td>.336*** (.061)</td>
<td>.201** (.071)</td>
<td>-.135* (.058)</td>
</tr>
<tr>
<td>Black</td>
<td>.357*** (.092)</td>
<td>.902*** (.100)</td>
<td>.544*** (.089)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>.509*** (.078)</td>
<td>.460*** (.089)</td>
<td>-.049 (.077)</td>
</tr>
<tr>
<td>Asian</td>
<td>.812*** (.160)</td>
<td>1.176*** (.159)</td>
<td>.363** (.107)</td>
</tr>
<tr>
<td>Native American</td>
<td>-.328 (.208)</td>
<td>-.352 (.264)</td>
<td>-.025 (.246)</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>.545*** (.048)</td>
<td>1.427*** (.049)</td>
<td>.882*** (.044)</td>
</tr>
</tbody>
</table>
N=12135                    | N=12135          | N=12135          |

NOTE: Standard deviations for all coefficients are presented in parentheses.
NOTE: *p<.05; **p<.01; ***p<.001

Outcomes and likelihood of enrollment are more complex than anticipated. The ability of measures of educational outcomes to predict enrollment differs by academic subject. For grades, English grades are more powerful than math grades. In terms of test scores, math test scores are superior to English test scores in explaining enrollment. The consistent differences across measures help to explain why the models which contained both grades and test scores were judged to be better fits than models that contained only one measure. Overall, it appears that using both grades and test scores to predict enrollment is most appropriate.

In order to understand the consistent differences in the viability of grades or test scores as predictors of enrollment across academic subjects, it is important first to take
into consideration the two facets of postsecondary enrollment – students’ decisions to enroll and gatekeepers’ decisions to accept students. Students make decisions regarding whether or not to continue education using various types of information including an understanding of their own motivation and abilities, academic feedback, and encouragement from significant others. Gatekeepers make decisions based primarily on academic information such as grades, test scores, and personal essays. While unrestricted admissions policies at many community colleges mean that few students are unable to gain admission to any postsecondary institution, the presence of gatekeepers may influence the type of institutions to which students will apply or be accepted.

The differences in the predictive abilities of measures of academic outcomes will be explored from both sides of the enrollment process. First, in terms of students’ decisions to pursue postsecondary education, while earlier analyses indicated few major differences in what predicts grades and test scores across academic subjects, the two measures may differ in how they work as feedback across subject. As was noted in Chapter Four, students are not always required to take four years of math courses. Therefore, by the time they reach their senior year or graduate from high school, their decisions regarding postsecondary education may be influenced more by their outcomes on recent standardized tests given in school or results on college admissions tests such as the SAT or ACT than by grades received in prior math courses. Significant others, such as parents and guidance counselors, also may base their recommendations to students on these results. For English, many schools require four years of courses which means that English grades could act as more salient feedback to seniors when making postsecondary education decisions.
In addition to the relevance of English grades due to their temporal proximity, it is also possible that students and those around them do not take results on the English sections of standardized tests as seriously as results for math sections. As was noted in Chapters Four and Five, analyses suggest that the content of standardized tests of English may not be tightly linked to the content of English courses. Students may become aware of these differences through taking tests and look instead to grades for feedback on their academic abilities. In addition, teachers and counselors may encourage students not to rely on these results when making decisions about their educational futures. In contrast, the content of math standardized tests appears to be more tightly linked to course content so that neither measure would be considered more representative of academic abilities than the other.

The discussion above provides reasons why students may base educational decisions on different measures of academic outcomes. It is also important to evaluate why institutional gatekeepers may base admissions decisions on different measures. Several of the arguments made above also apply to gatekeepers and may help to explain the difference in the predictive nature of the two measures of outcomes. If students do not take four years of math courses, colleges may focus on test scores rather than grades as more contemporary indicators of students’ math skills. Admissions representatives may also believe that English grades and test scores represent different skill sets and choose to focus on grades which may offer a better accounting than test scores of valued abilities such as students’ critical thinking and writing skills.

The multinomial logit models in Tables 8.2 and 8.3 illustrate and help to clarify differences across academic subjects, but the coefficients are not intuitively interpreted.
For that reason, predicted probabilities were calculated to assess the relative ability of each measure to predict college enrollment. Rather than focusing on differences between English and math results, the predicted probabilities are used to approximate the effects of students having high or low GPAs or test scores. The results illustrate the overall effects of each measure when the results for subjects are combined rather than examined independently.

In order to determine which values of the explanatory variables to utilize for predictions, percentiles were estimated. To approximate high-achieving students, values at the 90th percentile are used. For low-achieving students, the 20th percentile is used. The 90th and 20th percentiles were chosen because they represent opposite ends of the achievement spectrum yet each point should present variations in students’ decisions and outcomes. Examining students at the extreme ends of the distribution would likely produce little variation in outcomes and, therefore, not be truly informative. Table 8.4 contains predicted probabilities for models that contain grades and test scores. Table 8.5 presents predicted probabilities calculated from a model that also contains control variables. In determining the probabilities, the value for grades or test scores was entered and every other variable in the model was held at the average value. The probabilities indicate how likely a student with those characteristics would be to enroll in postsecondary education.

An initial overview of the tables offers support for elements of the argument presented above regarding the varying roles of outcome measures as feedback. As mentioned in that discussion, few students are unable to gain admission to any postsecondary institution. Two-year postsecondary institutions, such as community
TABLE 8.3
PREDICTED PROBABILITIES FOR POSTSECONDARY ENROLLMENT BASED ON GRADES AND TEST SCORES

<table>
<thead>
<tr>
<th>Percentiles*</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades =</td>
<td>90th</td>
<td>20th</td>
<td>90th</td>
<td>20th</td>
<td>90th</td>
<td>20th</td>
</tr>
<tr>
<td>2-year</td>
<td>.16</td>
<td>.34</td>
<td>.15</td>
<td>.35</td>
<td>.07</td>
<td>.39</td>
</tr>
<tr>
<td>4-year</td>
<td>.77</td>
<td>.42</td>
<td>.80</td>
<td>.34</td>
<td>.91</td>
<td>.23</td>
</tr>
<tr>
<td>No College</td>
<td>.07</td>
<td>.24</td>
<td>.06</td>
<td>.31</td>
<td>.02</td>
<td>.39</td>
</tr>
</tbody>
</table>

NOTE: The sample size is 12135.
*For the analytic sample, the 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 23 for English and 33 for math.

TABLE 8.5
PREDICTED PROBABILITIES FOR POSTSECONDARY ENROLLMENT BASED ON GRADES, TEST SCORES, AND CONTROL VARIABLES

<table>
<thead>
<tr>
<th>Percentiles*</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests =</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades =</td>
<td>90th</td>
<td>20th</td>
<td>90th</td>
<td>20th</td>
<td>90th</td>
<td>20th</td>
</tr>
<tr>
<td>2-year</td>
<td>.16</td>
<td>.36</td>
<td>.18</td>
<td>.35</td>
<td>.09</td>
<td>.41</td>
</tr>
<tr>
<td>4-year</td>
<td>.78</td>
<td>.44</td>
<td>.76</td>
<td>.41</td>
<td>.89</td>
<td>.28</td>
</tr>
<tr>
<td>No College</td>
<td>.06</td>
<td>.20</td>
<td>.06</td>
<td>.24</td>
<td>.02</td>
<td>.31</td>
</tr>
</tbody>
</table>

NOTE: The sample size is 12135.
*For the analytic sample, the 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 23 for English and 33 for math.
colleges, often have open admissions policies and will accept all applicants who have a high school diploma or equivalent. This means that the effect of gatekeepers (whether through actual decisions not to admit students or students’ perceptions that they will not be admitted) should operate primarily through students’ enrollment in two- versus four-year colleges.

The predicted probabilities suggest that, overall, a minority of students at any achievement level chooses not to enroll in postsecondary education. This is partly due to the biased nature of the sample because most of the students are high school graduates. But it also indicates the low thresholds set for admission to some postsecondary institutions. Even students who did poorly on both measures of outcomes are less likely to opt out completely than to enroll in some form of postsecondary education – predominantly at two-year institutions. Students who do well on both measures of academic outcomes tend to enroll in four-year institutions. Those who do well on only one appear to more often be pushed out of four-year institutions and into two-year institutions rather than choosing not to enroll at all. These results highlight the power of grades and test scores in determining students’ educational futures – whether through their own choices or through institutional decisions.

Further investigation of the tables reveals additional information regarding each measure of outcomes. The predicted probabilities in Table 8.4, derived from the unadjusted model, support the results of the measures of fit which indicated that test scores are stronger predictors of college enrollment than grades. In all instances, students with high test scores are more likely to enroll in a four-year college than those with high grades.
In order to better understand the relative power of each measure, predicted probabilities were estimated to examine students who may perform well on both measures or on one measure but not on the other. For these probabilities, the scores for grades and test scores were entered while all the control variables were held at the average. The results are presented in the last four columns of Table 8.4. Students who are at the 90\textsuperscript{th} percentile for both grades and test scores have a 91\% likelihood of attending a four-year college. Students who perform at the 20\textsuperscript{th} percentile on both measures are considerably more likely to enroll in a two-year institution or not to enroll than to enter a four-year college.

When students are mismatched on the two measures, the overall probability of postsecondary enrollment drops compared to when they have high results on both measures. As was expected based on the goodness of fit results, test scores prove to be the stronger predictor of enrollment when compared to grades. Students who have mismatched outcomes with high test scores and low grades appear to resemble students who are high-achievers more than low-achieving students. This indicates that students who do well on standardized tests can overcome poor results on grades. They may do this by ignoring the negative feedback on the one measure and focusing on the positive feedback which encourages them to continue their education. Or they may be more likely to continue on because institutional gatekeepers weigh the positive outcome on tests more heavily than the negative outcome for grades. The high probability also may be due to the fact that English test scores appear to account for background characteristics that predict enrollment, as indicated by the goodness of fit statistics and regression results.
for the adjusted model. Table 8.5 presents predicted probabilities derived from the model that contained student background characteristics in order to explore this scenario.

In Table 8.5, students with high grades are somewhat more likely to enroll in a four-year college than a two-year college when compared to students with high test scores. The change results mainly from a drop in the predictive ability of test scores which occurs because of the inclusion of control variables representing background characteristics. This further indicates that test scores represent background factors that predict enrollment. Once again, students with results in the 90th percentile for both measures are likely to attend four-year colleges. Students who have mismatched results have a high probability of enrollment in a four-year college. They are equally likely to enroll whether they did poorly on grades or tests.

The three methods used above to analyze the ability of grades and test scores to predict postsecondary enrollment provide new insight into the use of measures of educational outcomes in research. The results were more complex than originally hypothesized and provide relevant information regarding two aspects of the measures’ usefulness in particular. The first consistent finding was that the exclusion or inclusion of students’ background characteristics in the analyses mediates the ability of grades or test scores to predict enrollment. The analyses revealed that the relationship between English test scores and enrollment was significantly reduced by the inclusion of control variables. Therefore, while English test results may appear to be strong predictors if used in isolation, they are actually reflecting other characteristics of students that affect their likelihood of enrollment. They are not acting as strong proxies for students’ academic
abilities or motivation and they do not appear to serve as feedback that students utilize when deciding whether or not to pursue postsecondary education.

The second primary finding regarding enrollment was that the measures of educational outcomes differed in how well they predicted enrollment across academic subject. English grades and math test scores were consistently strong predictors of enrollment while math grades and English test scores proved to be unreliable or unrelated to enrollment. The variations by subject are evidence that care should be exercised when grades or test scores are used to predict enrollment. Overall, using both measures may provide the strongest predictions, as suggested by the goodness of fit statistics. However, using an overall average of grades or test scores can confound relationships and mask differences in predictive powers that are relevant to research and the interpretation of results. Therefore, the best course of action, if both grades and test scores are used to predict enrollment, is to include separate variables representing outcomes by academic subjects.

The second dependent variable analyzed in this chapter is whether or not respondents graduated from college with a bachelor’s degree. Prior to discussing the analyses, clarification is needed regarding the perceived relevance of measures of high school educational outcomes for predicting college graduation in comparison to predicting postsecondary enrollment. While the two dependent variables both clearly indicate educational progress, they are conceived of as fundamentally different outcomes for the purpose of this research. Enrollment in postsecondary education represents students’ academic orientations, perceptions of their own abilities, and how they are perceived by outside arbiters based on their academic histories. Successfully graduating
from college requires different characteristics of students. Graduation depends on long-
term commitment to academic goals, a willingness to meet demands and behave in
accordance with institutional policies, and the capability to learn and complete assigned
work. The differences in what determines success for each dependent variable should
lead to differences in the predictive power of the explanatory variables.

For college graduation, it is hypothesized that grades will predict students’
likelihood of bachelor degree attainment more strongly than test scores. This hypothesis
is based on the findings in previous chapters coupled with the general requirements for
receiving a college degree. Earlier analyses indicated that grades offer accounts of
students’ achievement orientation and willingness and ability to work independently.
These characteristics are important for succeeding in college – a setting in which success
will be determined by similar standards to those used to assign grades in high school
classrooms. However, the knowledge and skills measured by test scores also will play an
important role in predicting college graduation. Overall, though, it is believed that
knowledgeable students who are unwilling to complete assigned work will be less likely
to graduate than hard-working students who may lack initial knowledge. Therefore, the
characteristics of students that are accounted for in grades should be better predictors of
college graduation than those captured by test scores.

Because not all respondents enroll in a four-year postsecondary institution, not all
respondents have an equal chance of graduating. In order to account for the differences
between the groups of students, two sets of models were run. One included all
respondents to predict the likelihood of graduation for all types of students. The second
set of analyses used a limited sample that included only students who reported enrolling
in a four-year institution to examine the predictive ability of grades and test scores for the more exclusive group of students, all of whom took the first official step to earning a degree.

Logit models were utilized to analyze the likelihood of college graduation.\textsuperscript{2} Similar analyses were conducted to those for postsecondary enrollment. First, the models were assessed for goodness of fit. The model with grades alone is stronger than the model with only test scores for both the full sample and the limited sample of students who enrolled in college. When compared to a model containing both grades and test scores, using both measures is judged to be a stronger fit than using grades alone. These results are the same for models that contain only the measure of educational outcomes and for models that also contain controls for students’ background characteristics. The durability of grades as predictors of college graduation was anticipated because the characteristics of students that grades were found to account for in previous chapters were hypothesized to benefit students throughout college. Test scores do not reflect these beneficial behaviors and attitudes and therefore cannot predict postsecondary success as well as grades. However, it is important to note that using both grades and test scores provides the strongest predictions, similar to the results for enrollment.

The results of the logit models for the full sample and the sample that is limited to students who enrolled in postsecondary education are presented in Tables 8.6. Both

\textsuperscript{2} A Heckman selection model (1979) could have been used to account for the characteristics that led students to enroll in college or not when predicting graduation. However, the model does not allow for the calculation of predicted probabilities. Because the benefits of using the selection model are limited and the results are generally similar to those obtained using a logit model, the logit model was used instead of the Heckman selection model because it allowed for predicted probabilities. The Heckman selection model that was calculated indicated that test scores did not predict college graduation so the logit models likely overestimate the positive relationship between test scores and college graduation.
TABLE 8.4
LOGISTIC REGRESSION OF GRADES, TEST SCORES, AND CONTROL VARIABLES ON GRADUATION FROM A FOUR-YEAR COLLEGE

<table>
<thead>
<tr>
<th></th>
<th>Enrollees*</th>
<th>Enrollees*</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Grades</td>
<td>0.300***</td>
<td>0.227***</td>
</tr>
<tr>
<td>Math Grade</td>
<td>0.058***</td>
<td>0.073***</td>
</tr>
<tr>
<td>English Test Score</td>
<td>-0.001</td>
<td>-0.006</td>
</tr>
<tr>
<td>Math Test Score</td>
<td>0.039***</td>
<td>0.022***</td>
</tr>
<tr>
<td>Female</td>
<td>0.197***</td>
<td>0.260***</td>
</tr>
<tr>
<td>Black</td>
<td>0.313**</td>
<td>0.071</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.268**</td>
<td>-0.336**</td>
</tr>
<tr>
<td>Asian</td>
<td>0.515***</td>
<td>0.325**</td>
</tr>
<tr>
<td>Native American</td>
<td>-0.819*</td>
<td>-0.689</td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>1.096***</td>
<td>0.699***</td>
</tr>
<tr>
<td>N=12029</td>
<td>N=6477</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Standard deviations for all coefficients are presented in parentheses.

NOTE: *p<.05; **p<.01; ***p<.001

*Sample size limited to students who enrolled in a four-year college.

Unadjusted and adjusted models were estimated but only the adjusted models are presented because the results were similar for both models. The main difference between the two is that math grades gain in significance from the unadjusted to the adjusted model, although math grades are significant predictors of college graduation in both models.

English grades and math test scores are strong predictors of college graduation for both the full sample and the sample limited to students who enrolled in a four-year college. This is similar to the enrollment model in which both of these measures were also found to be strong predictors. The model differs from the enrollment model in that math grades are now strong predictors of the outcome of interest. It was hypothesized in
the previous section that math grades do not predict enrollment because they do not function as a strong form of feedback for students. They may operate as reliable predictors of college graduation, however, because they measure academic and social skills that benefit students throughout college.

English test scores are not reliable predictors of college graduation in the adjusted or unadjusted models. This finding provides additional evidence that English tests do not operate as valid indicators of skills which are relevant for success in academic settings. Despite the lack of significance for English test scores, overall the results support the contention that both grades and standardized test scores can account for characteristics of students that aid them throughout college. In addition, they indicate that grades offer a more reliable accounting than test scores of academic skills and behaviors relevant to postsecondary success. Once again, the predictive capabilities of the measures vary by academic subject, suggesting that care should be taken when predicting success based on measures of educational outcomes.

Predicted probabilities were estimated to translate the logit results discussed above into more understandable terms. Table 8.7 presents predicted probabilities based on the regression models for the full sample of students. Table 8.8 contains probabilities for the model that includes only students who enrolled in a four-year college. Probabilities were calculated for the 90th and 20th percentiles of each measure. Both tables show that students with high grades are more likely to graduate from college than those with high test scores. This is consistent with the findings of the goodness of fit measures and the regression models. The overall probability of graduation is generally low in Table 8.7 because it includes the full sample of students – including those who
TABLE 8.7
PREDICTED PROBABILITIES FOR GRADUATION FROM A FOUR-YEAR COLLEGE BASED ON GRADES, TEST SCORES, AND CONTROL VARIABLES

<table>
<thead>
<tr>
<th>Test Percentiles*</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests = 90th</td>
<td>.56</td>
<td>.13</td>
<td>.41</td>
<td>.16</td>
<td>.73</td>
<td>.08</td>
</tr>
<tr>
<td>Grades = 90th</td>
<td>.44</td>
<td>.59</td>
<td>.84</td>
<td>.17</td>
<td>.92</td>
<td>.58</td>
</tr>
</tbody>
</table>

NOTE: The sample size is 12029.

*For the analytic sample, the 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 23 for English and 33 for math.

TABLE 8.8
PREDICTED PROBABILITIES FOR GRADUATION FROM A FOUR-YEAR COLLEGE BASED ON GRADES, TEST SCORES, AND CONTROL VARIABLES FOR STUDENTS WHO ENROLLED IN A FOUR-YEAR PROGRAM

<table>
<thead>
<tr>
<th>Percentiles*</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
<th>20th</th>
<th>90th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tests = 90th</td>
<td>.82</td>
<td>.48</td>
<td>.71</td>
<td>.61</td>
<td>.86</td>
<td>.44</td>
</tr>
<tr>
<td>Grades = 90th</td>
<td>.18</td>
<td>.52</td>
<td>.29</td>
<td>.39</td>
<td>.14</td>
<td>.56</td>
</tr>
</tbody>
</table>

NOTE: The sample size is 6477.

*The percentiles are based on the full sample of students. The 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 23 for English and 33 for math.
never enrolled in college. When compared to the probabilities for enrollment, Table 8.7 indicates that many students who enter college do not receive a degree within eight years of high school graduation. Interestingly, while students with high test scores and low grades were as likely as or more likely to enroll than those with high grades and low test scores, the reverse is true regarding graduation. Students with high grades are considerably more likely to graduate, no matter what their test scores are, than those with high test scores. This finding highlights the different roles of grades and test scores as feedback in the decision making process and as indicators of student knowledge and behaviors that are applicable in a college setting. It appears that the factors that prompt students to enroll in college are not the same as those which lead them to be successful in college. Again, this disconnect suggests that careful thought should be given when using grades or test scores in research and policy formulation.

The probabilities discussed above represent both the likelihood of enrollment and graduation because the group that did not graduate consists of students who did not enroll in four-year colleges and students who did enroll but did not complete the requirements for a bachelor’s degree. While this type of analysis can be informative in terms of looking at the overall likelihood of college graduation, it is also important to look separately at the predictive capabilities of grades and test scores for the sample of students that enrolled in a four-year college. Focusing on the relationship between grades and test scores and the likelihood of success for this subgroup of students may reveal different effects than analyses that examine all students, regardless of enrollment status. The predicted probabilities for the limited sample of students are presented in Table 8.8.
The probabilities of graduation for the limited sample follow the same general trends as those in Table 8.7 in that students with high grades are more likely to graduate than those with high test scores and those with high results on both measures are more likely to graduate than those with strong outcomes on only one measure. The tables differ in that the overall probabilities are higher for all categories in Table 8.8 than in Table 8.7. This is due to the selective nature of the sample because all of the students took the initial step toward graduation by enrolling in a four-year college.

The probabilities indicate that once students have enrolled in college, they stand a good chance of graduating. Even students who scored in the 20th percentile on both measures of educational outcomes have a 44% chance of graduation if they enroll in a four-year college. This is compared to an 8% chance of graduation for the full sample of students with grades and test scores at that level. The high probability of graduation, given enrollment, may be due to selection bias. Academically undistinguished students who choose to enroll in four-year colleges may differ from their counterparts who do not enroll in a fundamental way. For instance, they may be more motivated, face more parental pressure, or have better finances to support their academic pursuits. The result also may be related to institutional supports provided to students who enroll. Students with lackluster academic records may have access to services such as tutoring or preparatory courses that help them to continue through to graduation. Whatever the explanation, grades and test scores differ in how they predict outcomes based on the sample that is being analyzed.

Taken together, the analyses of college graduation illustrate that grades are consistently strong predictors of graduation. Test scores provide a weaker method for
determining the likelihood of graduation. This is primarily due to the lack of a relationship between English test scores and graduation. While math test scores can be used to predict graduation, the overall strength of predictions based on test scores are weakened by the inclusion of English test scores. Both English and math grades provide strong predictions and, therefore, function as better predictors than test scores when the results for the academic subjects are used in conjunction. Similar to enrollment, though, grades and test scores are best able to predict graduation when both measures are included in models. The synergy of grades and test scores indicates that each measure represents somewhat different characteristics of students that are related to enrollment and graduation. While no statistical model would explain perfectly the likelihood of college graduation, a model that includes both grades and test scores accounts for many of the behavioral and academic characteristics of students that predict graduation.

The final dependent variable analyzed in this chapter is respondents’ self-reports of their income in 1999. Similar to college graduation, grades and test scores are expected to be related to respondents’ income because the measures of educational outcomes represent knowledge and skills that can benefit respondents in the workplace. It is hypothesized that respondents’ who earned high grades and/or standardized test scores will earn more income in 1999 than their peers who received lower grades or test scores. Grades are hypothesized to have a stronger relationship to income than test scores because they account for social and behavioral characteristics of students, as established in earlier chapters. These generalized soft skills will provide more benefit to students in most workplaces than the hard skills represented by test scores.
Part of the relationship between measures of educational outcomes and income is expected to be indirect in that the characteristics measured by grades and test scores are also related to respondents’ likelihood of earning a college degree. A bachelor’s degree benefits respondents in terms of income by increasing their likelihood of securing a job that pays more than if they did not have a degree. Therefore, including bachelor’s degree attainment in the models is expected to reduce or eliminate the relationships between grades and test scores and income in 1999.

It is important to remember that the respondents in these analyses have been out of school for only eight years. They are still relatively new to the occupational marketplace and their income patterns may not have stabilized. For instance, students who earned graduate degrees may just be beginning a career at this point or may still be in school. Also, respondents may not have had opportunities to earn promotions at work and secure additional income. For this reason, the results provide an idea of how grades and test scores could be used to predict income, but it is possible that they would function as more valid predictors if the reported income represented respondents who had graduated from high school more than eight years ago.

The income models are ordinary least square regressions because the dependent variable is continuous. The dependent variable is logged because it is not normally distributed – the distribution is skewed to the low end and the addition of each dollar of income does not have the same impact on income at each point in the distribution. Therefore, looking at the likelihood of a percentage increase in income is more appropriate than calculating actual dollar increases.
The models contain control variables to account for time worked in the previous year. These variables ensure that income indicates a regular wage rather than a yearly wage in case the respondent did not work a full year or had a part-time job. The two variables are respondents’ reports of how many weeks they worked in the year and how many hours they worked in the average week. Bachelor’s degree attainment is also included as a control in certain models.

Because ordinary least squares regression was used for these models, the variance explained by each model was evaluated using the adjusted R-squared value rather than the goodness of fit statistics used to evaluate the logistic regression models presented in this chapter. Adjusted R-squared was used as opposed to the regular R-squared value in order to account for the addition of independent variables. R-squared increases with the addition of each independent variable, thereby guaranteeing that a model with more independent variables will appear to explain more variance than models with fewer independent variables. Adjusted R-squared accounts for the addition of variables and offers an estimate of explained variance that does not automatically increase as independent variables are added.

The adjusted R-squared values for the models predicting income based on only grades or test scores indicate that the model with only grades explains less variance than the model with test scores. The value for the model with grades is .3729 while the value for the test score model is .3795. The model containing both measures has an adjusted R-squared of .3802, suggesting that combining both measures in this model provides the best method for predicting income.
In contrast to the unadjusted models, the adjusted R-squared for models which contain the controls for respondents’ background characteristics indicate that grades are stronger predictors of income than test scores, with respective R-squared values of .4007 and .3998. Similar to enrollment, it appears that test scores represent elements of respondents’ backgrounds that influence the outcome of interest. The adjusted R-squared value for the model containing grades, test scores, and control variables is .4018 which suggests that it is the strongest model in the group.

When bachelor’s degree attainment is included in the model, the R-squared results are slightly different. Overall, the models that contain educational attainment explain more variance than those that do not. This result was expected because bachelor’s degree attainment was hypothesized to be a strong predictor of income. The adjusted R-squared values for the models containing bachelor’s degree attainment and either grades or test scores are .3888 for grades and .3925 for test scores. The value for the model containing both measures is .3925. These values indicate that test scores account for slightly more variance than grades and, for the first time, the model containing only test scores is as strong as the model containing both measures.

For the models containing control variables and bachelor’s degree attainment, both the grades and test scores models have an R-squared value of .4114. The value for the model containing both measures is slightly larger at .4117. The difference in adjusted R-squared values is so small that all three of the models appear to account for essentially the same amount of variance, suggesting that they could be used almost interchangeably to predict income when educational attainment is accounted for in the model.
The adjusted R-squared results do not confirm the hypothesis that grades operate as better predictors of income than test scores. While grades were stronger than test scores in the adjusted model, the inclusion of bachelor’s degree attainment in the model eliminated the difference between the predictive abilities of the two measures. The regression models presented below will help to clarify the differences between the two measures and explain their value as predictors of income.

Table 8.9 presents the coefficients for the unadjusted and adjusted regression models without a control for bachelor’s degree attainment. In the unadjusted model, grades have no relationship with income. Both test scores are related to income; although English test scores negatively predict income. The results are unexpected in a number of ways. Only math test scores perform as hypothesized and positively predict income. Because grades represent academic and social skills relevant to occupational success, it was expected that they would have a significant relationship with income. The lack of a relationship indicates that grades may have no relationship to income or their impact may be indirect and operate through other factors. The unexpected results for English test scores could be explained by previous analyses which illustrated that those scores tend to reflect aspects of students’ backgrounds that predict outcomes. It is possible that those characteristics lead to lower incomes and the regression results point to a spurious relationship between English test scores and income.

The adjusted results, presented in the second column of Table 8.9, provide support for the argument regarding English test scores. In that model, English test scores no longer have a significant relationship with grades because controls for background characteristics are included. This confirms that the scores are capturing background
TABLE 8.9

REGRESSION OF GRADES AND TEST SCORES ON SELF-REPORTED INCOME IN 1999

<table>
<thead>
<tr>
<th></th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>Coefficient</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Grades</td>
<td>-.005</td>
<td>(.004)</td>
<td>.006</td>
<td>(.004)</td>
</tr>
<tr>
<td>Math Grade</td>
<td>.002</td>
<td>(.004)</td>
<td>.001</td>
<td>(.004)</td>
</tr>
<tr>
<td>English Test Score</td>
<td>-.004***</td>
<td>(.001)</td>
<td>-.002</td>
<td>(.001)</td>
</tr>
<tr>
<td>Math Test Score</td>
<td>.006***</td>
<td>(.001)</td>
<td>.003**</td>
<td>(.001)</td>
</tr>
<tr>
<td>Female</td>
<td>-2.19***</td>
<td>(.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>-.053*</td>
<td>(.022)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>.011</td>
<td>(.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>.049*</td>
<td>(.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td>-.096</td>
<td>(.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Socioeconomic Status</td>
<td>.047***</td>
<td>(.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>.222***</td>
<td>(.016)</td>
<td>.208***</td>
<td>(.016)</td>
</tr>
<tr>
<td>Weeks Worked in 1999</td>
<td>.035***</td>
<td>(.001)</td>
<td>.034***</td>
<td>(.001)</td>
</tr>
<tr>
<td>Hours Worked per Week in 1999</td>
<td>.022***</td>
<td>(.001)</td>
<td>.021***</td>
<td>(.001)</td>
</tr>
</tbody>
</table>

N=10441

NOTE: Standard deviations for all coefficients are presented in parentheses.
NOTE: *p<.05; **p<.01; ***p<.001

factors that are likely to depress income. Math grades remain unrelated to income and English grades gain moderate significance and positively predict income. The results are somewhat more consistent with the hypotheses presented above and mirror the results for enrollment in which English grades and math test scores were found to be the most reliable predictors.

Table 8.10 presents results for the models with the addition of a control for bachelor’s degree attainment. As expected, the variable for bachelor’s degree is a significant positive predictor of income. The results for the measures of educational outcomes in the unadjusted model are similar to those for the unadjusted model that does
TABLE 8.10

REGRESSION OF GRADES, TEST SCORES, AND EDUCATIONAL ATTAINMENT ON SELF-REPORTED INCOME IN 1999

<table>
<thead>
<tr>
<th></th>
<th>English Grades</th>
<th>Math Grade</th>
<th>English Test Score</th>
<th>Math Test Score</th>
<th>Female</th>
<th>Black</th>
<th>Hispanic</th>
<th>Asian</th>
<th>Native American</th>
<th>Socioeconomic Status</th>
<th>Bachelor’s Degree</th>
<th>Weeks Worked in 1999</th>
<th>Hours Worked per Week in 1999</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-.005</td>
<td>.002</td>
<td>-.004***</td>
<td>-.006**</td>
<td>.006***</td>
<td>-.219***</td>
<td>-.053*</td>
<td>.011</td>
<td>-.096</td>
<td>.047***</td>
<td>.222***</td>
<td>.035***</td>
<td>.022***</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
<td>(.001)</td>
<td>(.001)</td>
<td>(.001)</td>
<td>(.014)</td>
<td>(.022)</td>
<td>(.019)</td>
<td>(.055)</td>
<td>(.055)</td>
<td>(.016)</td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
</tbody>
</table>

N=10441 N=10441

NOTE: Standard deviations for all coefficients are presented in parentheses.

NOTE: *p<.05; **p<.01; ***p<.001

not include bachelor’s degree attainment in Table 8.9. This indicates that the relationships between test scores and income are not mediated through the ability of test scores to predict educational attainment.

The adjusted model eliminates the negative relationship between English test scores and income, as predicted based on the previous adjusted model. In this model, grades are not significant predictors, suggesting that the positive effect of English grades on income in the previous model represented the greater likelihood that students with high grades would receive a bachelor’s degree which would improve their income potential. Math test scores remain moderate positive predictors of income. This suggests
that students with high math test scores have a propensity to enter high-paying occupations. The relationships may exist because students who do particularly well in math in high school are highly motivated or have gained knowledge that is valued in an occupational setting. These students may also be more likely to select college majors that lead to well-paid positions.

The results described above provide further proof that grades are not strong predictors of income seven years after high school graduation. As mentioned earlier, that may be because respondents’ incomes have not yet stabilized. It may also be because intervening factors have muted the effectiveness of grades as a measure of students’ abilities and behaviors. The elimination of the moderate relationship between English grades and income by the inclusion of bachelor’s degree attainment supports this conclusion. While high grades may represent characteristics of students that will benefit them throughout postsecondary education, those same characteristics may not be as important in the workplace. However, the characteristics that earn high grades are ultimately beneficial in terms of income because they increase educational attainment.

Tables 8.11 and 8.12 present predicted probabilities for the adjusted models. The results for the unadjusted models are not shown because the models were obscured by confounding background effects. The probabilities in Table 8.11 are based on a model that does not include bachelor’s degree attainment while those in Table 8.12 are derived from a model that does include a control for attainment. In both tables, the predicted incomes fall below the mean income for the sample (presented in Table 8.1). This indicates that factors not included in the tables have strong positive effects on income.
TABLE 8.11
PREDICTED PROBABILITIES FOR INCOME IN 1999
BASED ON GRADES, TEST SCORES, AND CONTROL VARIABLES

| Percentiles* | Tests = | | | | | | Grades = | | | | | | Income | |
|--------------|--------|---|---|---|---|---|---|--------|---|---|---|---|---|---|---|---|---|---|---|
|              | 90th   | 20th | 90th | 20th | 20th | 90th |
| Tests =      |        |      |      |      |      |      |
| Grades =     | 90th   | 20th |      |      |      |      |
| Income       | 22697.27 | 20148.80 | 22048.50 | 20351.30 | 23765.73 | 19463.57 | 21923.18 | 21088.86 |

NOTE: The sample size is 10441.
*For the analytic sample, the 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 24 for English and 34 for math.
TABLE 8.12
PREDICTED PROBABILITIES FOR INCOME IN 1999
BASED ON GRADES, TEST SCORES, CONTROL VARIABLES, AND EDUCATIONAL ATTAINMENT

<table>
<thead>
<tr>
<th>Tests =</th>
<th>Percentiles*</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
<th>90th</th>
<th>20th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades = 90th</td>
<td>21798.58</td>
<td>20634.09</td>
<td>21545.02</td>
<td>20706.44</td>
<td>22292.38</td>
<td>20280.19</td>
<td>21424.71</td>
</tr>
<tr>
<td>Grades = 20th</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: The sample size is 10441.

*For the analytic sample, the 90th percentile grades are A- for English and B+ for math and the 20th percentile grades are C- for English and D+ for math. The 90th percentile test scores are 46 for English and 68 for math and the 20th percentile are 24 for English and 34 for math.
The predicted incomes in Table 8.11 follow a standard pattern in that respondents with high grades or test scores earn more than those with low grades or test scores. Students with high grades earn slightly more than those with high test scores. These results are in keeping with the adjusted R-squared results for this model which indicated that grades were stronger predictors of income than test scores. The positive effect of grades is most likely due to the absence of bachelor’s degree attainment in the model, though, as that factor was found to mediate the effect of grades.

The results in Table 8.12, which are based on a model including bachelor’s degree attainment, reveal differences between predicted incomes based on grades and test scores that are generally smaller than for the model without educational attainment. This compression in income differences makes sense because holding educational attainment constant reduces the relationship between income and grades or test scores so changing the values of the measures of educational outcomes exerts less influence on income than in the model that does not contain bachelor’s degree attainment. The variations in the probabilities show that the positive effects of grades compared to test scores are reduced. However, students with high grades still earn slightly more than those with high test scores. This is surprising because in the regression models, grades are not significantly related to income while math test scores are positively related to income. The small difference in outcomes may be a result of the negative, although insignificant, relationship between English test scores and income. That relationship may cancel out a portion of the positive effect of math test scores on income and lead to stronger predictions of income based on grades. Again, this suggests that when grades and test scores are utilized in research, they should be analyzed separately by academic subject.
Ultimately, though, the differences are small and the adjusted R-squared values for the grade and test score models are the same indicating that they provide equal, although weak, predictions of income.

The results presented in this chapter provide a strong foundation for assessing how best to utilize measures of educational outcomes as predictors of later life outcomes. It is important to clarify that it is not believed that grades or test scores themselves are directly benefiting respondents several years after high school. For instance, employers do not pay respondents more because they received an A in 12th grade English. In certain circumstances, receiving good grades may benefit respondents by providing initial access to college or the workplace. However, that access to entry is not providing enduring benefits because respondents still have to prove they are capable in order to maintain their positions. The behaviors and abilities that earned the grades are causing the continuing effects on outcomes such as college graduation.

However, grades were not shown to be strong predictors in all instances. Their viability varied by the outcome being assessed and the subject for which the grade was assigned. English grades were strongly related to postsecondary enrollment and graduation and weakly related to income. Math grades were strongly related to the likelihood of graduation but no other outcome. The varying connections between grades and the outcomes of interest often did not correspond with the presented hypotheses. The same is true of the results for test scores which also proved to be more complex than expected. Further research is warranted to continue to develop an understanding of how measures of educational outcomes function as predictors.
Despite remaining questions, the analyses in this chapter do provide findings which can benefit researchers and policymakers who are currently assessing educational outcomes. Combined with the findings from previous chapters, which clarify what grades measure in terms of both academic and social factors, these results are particularly helpful because they provide a holistic evaluation of grades as a measure of educational outcomes. The earlier analyses contributed to this goal by presenting findings that helped to clarify what academic and behavioral factors are measured in grades, both in terms of direct and indirect effects. This chapter complements the earlier analyses by examining how grades work as predictors of later outcomes, based on the characteristics for which they were found to act as proxies. Overall, the findings indicate that grades can offer a compelling alternative or complement to standardized test scores. To ensure the strongest results, neither measure should be averaged across academic subject because predictive power varies both by measure and by subject. Finally, combining both measures in research that examines students’ later outcomes is likely to result in the strongest predictions.
CHAPTER 9:
CONCLUSION

The preceding chapters analyze grades as a measure of educational outcomes in relation to standardized test scores and assess their potential for use in research. While each chapter stands on its own as a contribution to developing an understanding of grades, it is also necessary to synthesize the information and assess the overall contribution of the research. The following discussion will consider the information presented in earlier chapters, address the relevance of the findings in terms of theory, research, and policy, and discuss areas of opportunity and avenues for further exploration.

One of the principal findings of the research is that grades are a fundamentally different indicator of academic outcomes than standardized test scores. The primary difference between the two measures is that grades account for behavior, attitudes, and achievement orientation while test scores primarily provide an account of students’ knowledge gains and do not offer a strong assessment of the characteristics captured in grades. The social aspect of grades indicates that they evaluate elements of students’ academic outcomes that are of interest to sociologists of education. The analyses also establish that engaging in approved behaviors and illustrating dedication to educational goals promotes academic achievement for individual students and may facilitate opportunities to learn for all students within a classroom. These findings provide an
essential starting point for developing insight into the appropriate use of measures of academic outcomes and for understanding the contribution of the analyses to larger conversations regarding education.

The research makes theoretical contributions in that it utilizes and combines established sociological theories to develop an understanding of grades. By incorporating theories that address the two principal mandates of education – instruction and socialization – the theoretical framework creates a comprehensive foundation that acknowledges the importance of both aspects of achievement. Examining both goals simultaneously reveals the connections between socialization and instruction, demonstrating that well-socialized students behave in ways that allow them to receive more benefits from instruction than their less committed peers, thereby improving their academic outcomes.

With this understanding in mind, the theories are utilized to understand how measures of educational outcomes account for instruction and socialization. While test scores primarily measure instruction, grades are theorized to capture elements of both educational goals. Because the theory establishes the importance of both goals, it implicitly demonstrates the value of using a measure of educational outcomes that accounts for instruction and socialization. In addition, the framework illustrates the importance of feedback in the academic process, indicating that regular and understandable feedback could affect later educational outcomes. Grades are a strong form of feedback to students and significant others and that feedback can play a role in achieving academic goals by reinforcing or changing behaviors. Applying the existing theoretical elements to understanding the value of what grades measure and the potential
effects of grades on students serves the dual purpose of strengthening the existing
theories and improving the understanding of educational outcomes.

The research also expands on the existing theories through examinations of
classroom interactions. The analyses of classroom observations provide a supplement by
offering insight into the effects of student behavior and teacher-student interactions on
opportunities to learn. They demonstrate that students’ behaviors matter beyond the
potential impacts on the individual student’s academic achievement. Students’ behaviors
in the classroom, as primarily determined by their levels of socialization, mediate
instruction and opportunities to learn and, therefore, impact academic outcomes for both
the individual student and other students in the classroom. How teachers address and
react to the behaviors also plays a role in determining the impacts of the behavior on
other students – highlighting the multifaceted role teachers play in determining students’
academic outcomes. These findings directly address the interaction between socialization
and instruction and reinforce the validity of accounting for student behavior in grades.
The theoretical framework was subsequently expanded to include a more complete
understanding of the importance of enforcing behavioral standards and pursuing
socialization as an educational goal.

The theoretical framework, with the addition of the insights provided by the
classroom observations, highlights the previously underestimated relationship between
grades and standardized test scores. While grades and test scores measure different
aspects of educational outcomes, they are intimately connected. Teachers use grades as
tools not only to assess students’ academic outcomes, but also to enforce standards that
will improve those outcomes which, in turn, can impact students’ results on standardized tests.

Empirically, the study helps to clarify whether grades or standardized test scores are more appropriate for use in research. The two measures are often used as proxies for educational achievement; however, “achievement” may not be clearly defined. By understanding the aspects of students’ academic achievement gauged by each measure, researchers can strengthen research designs by making informed decisions regarding which measure of educational outcomes to utilize.

The best option, when using grades or standardized test scores as independent variables, may be to use both measures. Each indicator is a measure of achievement that taps the various aspects of educational outcomes. Including both can strengthen the predictive value of analyses by accounting for more aspects of students’ achievement than either measure alone. However, each measure has specific strengths in terms of predicting outcomes. If only one is used, the appropriate choice will vary based on the outcome of interest. The primary consideration is to carefully evaluate and understand what each measure represents so that it is clear what is being accounted for in the analyses.

When used as dependent variables, grades and test scores are two distinct measures. Different factors and characteristics will predict each measure of educational outcomes. For instance, analyses may indicate that a school characteristic or educational intervention does not directly influence test scores. However, the same variable may impact grades through modifications in student behavior. The effect of the variable at issue would be viewed differently based on the outcome used as the dependent variable.
Given the disparities between the two measures, research using grades instead of standardized test scores as a measure of academic outcomes may reflect differing aspects of education and is most likely not comparable to research that utilizes test scores.

Finally, the research speaks strongly to aspects of educational policy. Standardized test scores are the primary method for evaluating student progress and achievement under current educational policies. The analyses presented here indicate that this could present a limited view of student achievement because the tests primarily measure knowledge gains. Grades offer a more comprehensive understanding of student achievement by also accounting for behavior and achievement orientation. As the analyses in Chapter Eight show, this orientation and the accompanying behaviors can have important impacts on future outcomes. In addition, using both measures may be the most effective method for evaluating educational progress. Policies that evaluate students’ outcomes using both measures, with a clear understanding of what the measures mean and how they differ, may provide nuanced insights into educational quality and students’ abilities that test scores alone cannot contribute.

As noted in Chapter One, many policymakers and researchers may have reservations about using grades to assess student progress because of the assumed subjective and contextual characteristics of grades. This study indicates, though, that grades may not be as subjective as previously thought. Accounting for behavior and acceptance of educational norms does not make grades subjective and unreliable. As the analyses show, those factors are important elements of academic success. If the types of behaviors and attitudes included in grades vary across teachers, courses, and subjects, then reliability becomes a concern. However, the results discussed in Chapter Four
which include school-level variables and interactions between ability groups and students' behaviors provide an initial indication that teachers factor similar behaviors into grades across schools, subjects, and ability levels. These similarities may be due to teacher training and the type of peer-based socialization to the profession described by Lortie (1975). While further research is needed to clarify and confirm these assertions (particularly regarding teachers’ interpretations of student behaviors), preliminary findings indicate that teachers are relatively consistent across locations in accounting for students’ behaviors and acceptance of educational norms.

The analyses do show that grades are affected by the context in which they are assigned – particularly by ability group placement. While behavioral standards are similar across groups, the academic component of grades appears to vary based on the achievement level of students’ classmates. This finding indicates that research using grades should account for the academic environment in which teachers assign the grades. While this consideration heightens the complexity of using grades as measures of educational outcomes, it is not prohibitive. Compared to the time and effort that goes into designing and analyzing standardized test scores, it represents a minor effort. And, once again, this characteristic of grades points to the value of using both measures to analyze academic outcomes. Test scores provide a baseline for academic competence and progress while grades offer an estimate of students’ dedication to educational success and course-specific abilities that are not measured on standardized tests.

A pervasive finding that relates to educational policy is that teachers matter for determining students’ academic outcomes. While policy focuses on teacher competence and subject-matter qualification – reflecting the concentration on improving knowledge
gains—it often overlooks the important role of teachers in formulating evaluations of students. Teachers are the interpreters and final arbiters of students’ educational success in terms of grades. The research demonstrates that this role is doubly significant because it is also a factor in determining students’ success on standardized tests. A broadened conception of teachers’ primary importance in determining students’ academic success could inform policy decisions in a way that would encourage further training, collaboration, and assessments of reliability in grading standards across teachers.

The current research, while providing a strong base for understanding grades, also points to several areas of further research and exploration. First, there are some limitations based on the data used for the analyses. A larger number of cases in the analyses would have strengthened the findings. Also, having grades and test scores available for entire classes or schools would have provided an opportunity for a more in-depth quantitative examination of how academic context affects grades. The analyses presented in Chapter Six also would benefit from more closely matched reports of student behavior from teachers and students in order to clarify the role of teachers in interpreting students’ behaviors and directly accounting for them in grades. The qualitative analyses could be strengthened through the collection of more data representing additional academic subjects, ability groups, and schools.

The research is also necessarily limited in scope as it is an initial exploration of a large topic. Some clear areas of opportunity for continued research are present. Perhaps most importantly, the theory and findings indicate that contextual effects are important so examining peers in the classroom could be particularly revealing. The qualitative observations begin an exploration of that area but the research could be furthered with
more qualitative analyses and comprehensive quantitative data. In addition, the theory and research did not examine the roles of parents and friends in shaping students’ results on grades through impacts on behaviors or attitudes. Addressing this larger context of students’ lives may provide further insights into the determinants students’ grades.

The control variables representing students’ background and demographic characteristics revealed persistent differences in grades across groups, despite accounting for test scores and behaviors. These inequalities were not discussed because they were beyond the scope of the research. However, they provide evidence for entrenched inequalities in grades that are similar to disparities in test scores across demographic groups. The presence of consistent gaps is of particular interest to researchers because grades account for distinct aspects of students’ academic achievement that test scores do not measure and, therefore, inequalities in grades may result from different mechanisms than test score gaps. Examining the causes and consequences of disparities in grades across groups represents an opportunity to further sociologists’ understanding of educational inequalities.

Because the research examines grades as supplements or alternatives to standardized test scores, it would also be valuable to evaluate additional measures of achievement. One of the most commonly used alternatives is students’ self-reported grades. Although transcript-indicated grades and self-reported grades are highly correlated (Kelly 2008), differences do exist between the two measures and they may produce different results in analyses. Class rank is another alternative measure of achievement that is clearly context-dependent and may yield interesting findings regarding determinants of rank and the usefulness of the ratings in predicting outcomes.
Another clear area of opportunity for further research is to examine grades and their meanings for students who are not in high school. Grades assigned to elementary and postsecondary students may account for different types of behaviors and academic abilities than high school grades. For instance, elementary school grades may be more likely to explicitly include a behavioral component than high school grades. Elementary classrooms are also likely to be more heterogeneous in terms of ability levels than high school classrooms because the schools may not practice ability grouping. The differences in terms of assessment considerations and context may reveal variations in the meaning and effects of grades for elementary students and their parents.

Postsecondary grades offer similar opportunities in that behavioral standards in college may not be the same as those in high school due to class size, frequency of student-teacher interactions, and greater teacher autonomy. In addition, the academic context will be different because students self-selected to attend and had to meet minimum requirements to gain admittance. The characteristics of elementary and postsecondary schools that distinguish them from high schools offer a promising opportunity to explore differences in grading practices and the usefulness of grades as measures of educational outcomes at varying student age levels.

Ultimately, the research presented in the preceding chapters represents the initiation of a research agenda that has the potential to improve assessments of academic achievement and educational outcomes. While the current measures used for evaluation – standardized test scores – provide important information regarding educational success, they do not describe the educational process fully. Grades measure a more complex educational field in which behavior, effort, dedication, and ability can matter.
independently of achievement. Therefore, they come closer to capturing the reality of schools and classrooms. The process of learning, negotiating, and succeeding in classrooms is more similar to the environments in which students must seek success in the future than is the relatively non-social environment that is reflected in test scores. For the reasons noted above, students, teachers, researchers and policymakers will benefit from the addition of a well-conceptualized and reality-based measure of academic achievement to their understandings of educational outcomes.


