ABSTRACT

Title of Dissertation: NONPARTICIPATION OF THE 12TH GRADERS IN THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: UNDERSTANDING DETERMINANTS OF NONRESPONSE AND ASSESSING THE IMPACT ON NAEP ESTIMATES OF NONRESPONSE BIAS ACCORDING TO PROPENSITY MODELS

Young I. Chun, Doctor of Philosophy, 2009

Dissertation directed by: Professor Katharine Abraham
Joint Program in Survey Methodology

This dissertation examines nonparticipation of 12th graders in the year 2000 National Assessment of Educational Progress (NAEP), using a model of nonresponse developed by Groves and Couper (1998). NAEP is a continuing assessment of American student knowledge in various subject areas including mathematics and science, and the possibility that its results could be contaminated by a low response rate was taken as very serious. The dissertation evaluates the statistical impact of nonparticipation bias on estimates of educational performance in NAEP, by applying response propensity models to the NAEP mathematics and science survey data and the corresponding school administrative data from over 20,000 seniors in the 2000 High School Transcript Study (HSTS). When NAEP and HSTS are merged, one has measures of individual- and school-level characteristics for nonparticipants as well as participants in NAEP. Results indicate that nonresponse was not a serious contaminant, and applying response propensity based weights led to only about a 1-point difference out on average of 500 points in mathematics and of 300 points in science. The results support other recent research (e.g., Curtin, Press and Singer, 2000; Groves, 2006) showing minimal effects on nonresponse bias of lowered response rates.

NONPARTICIPATION OF THE 12TH GRADERS IN THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS: UNDERSTANDING DETERMINANTS OF NONRESPONSE AND ASSESSING THE IMPACT ON NAEP ESTIMATES OF NONRESPONSE BIAS ACCORDING TO PROPENSITY MODELS

by

Young I. Chun

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy
2009

Advisory Committee

Professor Katharine Abraham, Advisor and Chair
Professor John Robinson, Co-Chair
Professor Jeff Lucas
Professor Steve Martin
Dr. Andy Kolstad
Professor Sandy Hofferth

Dedication

I dedicate my dissertation work to my wife Myung-Joo Patricia Hong. This dissertation would not have been possible without her tireless support, tenacious encouragement, and unfailing love. She has spent numerous nights next to me to renew my efforts throughout the entire doctorate program. “Puffin,” you have been my best cheerleader and catalyst to fire my energy. You have shared many uncertainties, challenges, and sacrifices of mine during the completion of this dissertation.

I also dedicate this dissertation to the memory of my late mother Byung-Soon Lee, whose affection and love remains in me and has been vital for my efforts throughout the doctoral dissertation process. Mom, your love is ever stronger as I have completed a dissertation in your 30^{th} memorial year. I dedicate my dissertation also to my father Joseph Tok-Kyun Chun, who has devoted his whole life to me and my two brothers, Joe Young-Ho and David Young-Pok. Dad, your life-long diligence and confidence in me has lifted me to face challenges prudently and overcome fear bravely. I appreciate my youngest brother David for sparing many hours of wordsmithing despite his tight schedule with entrepreneurship. Young-Ho, I thank you for standing by me. My special thanks also go to my second mother Young-Hee Chun, who has closely supported my dad and her new family, allowing me to study all these years.

Finally, I dedicate this work and give special thanks to God. You are there for empowering me each day to take one step at a time. You have taught me patience and humility throughout the entire graduate program. I continue anchoring my life in you and discover joy as you guide me in new directions.

Acknowledgements

I would like thank all of those people who have helped this dissertation in the making for almost two decades. Above all, I wish to thank Professor Katharine Abraham, my advisor and committee chair, for all her encouragement, tenacious support, and intellectual empowerment. She has been a great inspiration as she gave me countless hours of reading, navigating, and most of all patience throughout the entire process. I could not have come this far without her caring, disciplining mentoring. I am so grateful to Professor Roger Tourangeau who connected me to Professor Abraham, who has been researching for years "nonresponse," the subject of my dissertation. I would like to thank my committee members, co-chair Professor John Robinson, Professor Jeff Lucas, Professor Steve Martin, and Professor Sandy Hofferth, for carefully reviewing my dissertation in their busy schedule and providing clear-sighted comments and suggestions for improvement. Co-chair Professor Robinson's invaluable comments made my dissertation a lot better, connecting findings to fair conclusions. My special gratitude goes to Dr. Andy Kolstad of the National Center for Education Statistics (NCES), a special member of the committee, for years of his support and insights. He was my mentor and colleague as well while I worked for National Assessment of Educational Progress (NAEP) for about a decade. I was able to continue studies of NAEP nonresponse for NCES, thanks to his sustaining support, while working at American Institutes for Research (AIR). NCES was where the seed of my dissertation about nonresponse in NAEP was planted. I also would like to thank Professor Robert M. Groves of the University of Michigan, currently the Director of the U.S. Bureau of the Census, for enlightening me to study nonresponse in the early years of my graduate training in Ann Arbor and encouraging me to finish my dissertation on nonresponse. My special thanks extend to Professor Joan Kahn for reviving my opportunity to complete this dissertation.

Additionally, I would like to acknowledge all of those people who provided encouragement, constructive criticism, or technical support with NCES restricted-use data: Associate Commissioner Dr. Peggy Carr and Dr. Steve Gorman, both at NCES, for ongoing encouragement to study NAEP including kind arrangement of their staff's steady support of my dissertation research; colleagues Shep Roey and Lloyd Hicks, both at Westat, for helping provide the High School Transcript Studies data to be merged with NAEP data; Professor Jeffrey Harring, Professor Gregory Hancock, Professor Robert Mislevy, Ms. Katrina Knudsen, Ms. Geraldine Todd, and Ms. Eppie Carroll, all at University of Maryland, for technical support to access restricted-use NCES data; and Dr. Janet Baldwin, Ms. Linda Hamilton, Dr. George Bohrnstedt, Dr. Bea Birman, and Dr. Mike Garret, all at AIR, for cheer up words and notes. I also owe a special thanks to all the staff at Assessment Division of NCES, or the NAEP family, who spurred me to finish my NAEP dissertation: Dr. Janis Brown, Dr. Arnie Goldstein, Dr. Andrew Malizio, Ms. Taslima Rahman, Dr. Emmanuel Sikali, Dr. Holly Spurlock, Mr. Richard Struense, Dr. Bill Tirre, Dr. Suzanne Triplett, Ms. Lisa Ward, and Dr. Sheida White. Also my gratitude goes to the faculty and administrative staff members of Department of Sociology at University of Maryland for their continued support on campus for almost two decades. I am particularly grateful to Professor Morris Rosenberg, who imprinted in my mind a love for sociology during my studies with him before he passed away.

Finally, I would like to thank my colleagues and mentors at NORC at the University of Chicago, Dr. Rachel Harter, Dr. Ken Copeland, Dr. Kirk Walter, and Dr. Fritz Scheuren, who provided me with fresh encouragement so that I can revise and complete the dissertation manuscript in time. It is my experience after experience that good things come out of NORC. Completing my dissertation is one of the many blessings I have received while I work with my wonderful colleagues at NORC.

Table of Contents

List of Tables

List of Figures

1. Introduction

Low participation rate of students in national survey assessments increases the potential for nonparticipation bias, a product of nonparticipation rate and the difference of characteristics between participating and nonparticipating students, and thus tends to lower data credibility (e.g., Smith, 1983). Nonparticipation (or nonresponse) bias has become more important in the National Assessment of Educational Progress (NAEP), where participation rates of the 12th graders' assessment have notably declined in recent years.

The final student participation rates over the last two decades at the 12th grade have been 10% to 35% lower than rates at grades 4 and 8. In 1990, the first year with the participation rate data available for Mathematics NAEP at grade 12, the overall participation rate was 66%, in comparison to 78% at grade 8 and 82% at grade 4 (See Figure A-1 in Appendix). The participation rate at grade 12 has continued to decline to 60% in 2000 while the participation rates at the two other lower grades during this period remained around the similar rate in 1990. The next five years, the participation rates at lower grades increased by about 10%; in contrast, the participation rate at grade 12 declined by about 5%. In 2005, the participation rate for the 12th grade Mathematics NAEP dropped to 56%, a decline of 10 percentage points from 1990. During the same period, the participation rate for the 12th grade Reading NAEP dropped by a similar magnitude to 55% (See Figure A-2 in Appendix). The further decline of participation rate among 12th graders may seriously affect validity of NAEP data. Statistical

Standards[1] enforced by National Center for Education Statistics (NCES) recommends, "In cases where prior experience suggests the potential for an *overall* unit nonresponse of less than 50 percent, the decision to proceed with data collection must be made in consultation with the Associate Commissioner, Chief Statistician, and Commissioner."

What pattern of twelfth graders' mathematics performance is observed from 1990 to 2005 when the participation rates had gradually declined? Because of changes in assessment content and administration, the results of NAEP student performance for 2005 could not be directly compared to those from previous years. Mathematics assessment at grade 12 was not carried out in 2003. Twelfth graders' performance data in Mathematics are thus comparable just for the period from 1990 to 2000. As the Nation's Report Card 2000 indicated (NCES, 2000), twelfth graders' performance showed overall gains from 1990 to 1996 by 10 points on the Mathematics scale score of 0 to 500. During this period, the participation rate was declined by 3 percentage points. In contrast, the average score for high school seniors was lower in 2000 (301) than in 1996 (304). During this period, the participation rate was further declined by 3 percentage points. It is hard to tell whether and the extent to which participation rate is associated with student performance as measured by Mathematics NAEP. Thus it requires a critical look into phenomena of nonparticipation in NAEP rather than parsimoniously attempting to relate it to student performance on the surface level.

The research issues facing this dissertation are multifold. How serious is nonparticipation in NAEP? What underlying process is behind student nonparticipation in education assessment? What micro- and macro-factors influence the nonparticipation

[1] See U.S. Department of Education Institute of Education Sciences (2003). NCES Statistical Standards, Standard 2-2-5, page 40.

of 12th graders in NAEP? Would social isolation theory be a useful navigator to help identify a set of micro- and macro-factors affecting nonparticipation? How closely is nonparticipation rate associated with a potential nonparticipation bias in NAEP estimates? How different are the characteristics of participants and nonparticipants in NAEP on key measures of assessment interest? What technical interventions would be feasible to reduce nonparticipation itself or adjust for potential nonparticipation bias?

The purpose of this dissertation is to investigate student- and school-factors affecting nonparticipation of 12th graders in NAEP by applying social isolation theory as guidelines and using a measurement and analysis model of nonresponse developed by Groves and Couper (1998). The dissertation research is also designed to evaluate the statistical impact of nonparticipation bias on estimates of educational performance in NAEP, using an approach used by Abraham, Maitland and Bianchi (2006). As Groves and Couper (2006) suggested, keen attention in this research is applied to investigating how strongly correlated the NAEP survey variables of interest are with (non)participation propensity, the likelihood of (non)participation. Such a research attention is justified by recent studies that demonstrated little empirical support to associate nonresponse rates with nonresponse bias (Merkele and Edelman 2002; Groves, 2006). The empirical findings might have *practical* implications about measures of interventions to adjust for nonparticipation bias and reduce nonparticipation itself in NAEP, by disclosing potential sources of nonparticipation.

What is nonparticipation? Nonparticipation in education assessments is a complex social process involving various individual and school factors as well as broad societal influences about schooling. In school assessment settings of NAEP in particular, it is essential to understand student- and school-level of influences on nonparticipation. Students themselves are the agents of decisions about nonparticipation in NAEP. Their background including their personal attitudes usually matters; school culture or school climate does affect student behavior like nonparticipation in NAEP. What is a theoretical navigator that helps understand a complex nonparticipation process in student assessment like NAEP?

Chapter 2 addresses a theoretical question by using a social isolation construct to explain student nonparticipation in NAEP. Chapter 2 examines studies of nonparticipation in various research realms, in an effort to develop a rich context and conceptual and analytical framework for the current study on 12^{th} grade nonparticipation in NAEP. Nonparticipation in this dissertation is used interchangeably with nonresponse, a term more frequently used in survey research literature. Chapter 2 reviews the origins and complex process of nonresponse, explains nonresponse bias by using deterministic and stochastic models, and explains approaches to assessing nonresponse bias in NAEP in particular. Chapter 3 turns to explaining research methods including data sources, sample design, and key variables and their relevance to social isolation construct used for the dissertation research. Chapter 3 presents an overview of a complex analysis plan from bivariate analysis to multiple logistic analysis to response propensity models for nonresponse adjustment. Chapter 3 is where one can envision the analytical value of merging NAEP student data with the school administrative data from the High School Transcript Studies (HSTS). Because the transcripts for the 2000 HSTS are collected from *all* students in the same NAEP sample of schools regardless of individual student's participation status in NAEP, the data merged between NAEP and HSTS include key correlates of nonresponse and makes robust assessment of nonresponse bias possible. In Chapter 4, I present findings from both bivariate analysis and multivariate logistic analysis to construct response propensity models. Chapter 4 demonstrates how alternative nonresponse weighing adjustment generated from response propensity models affect NAEP estimates in Mathematics and Science. I evaluate the re-weighted estimates in comparison with the current practice of NAEP adjustment for nonresponse which relies just on a few sampling frame variables. Chapter 4 is where I demonstrate applying the final response propensity model to estimate two related propensity models -- a contact model and a cooperation model conditional on contact -- and to investigate in turn the underlying mechanism of how social isolation variables of my choice would be robust enough to explain two sequential outcomes in NAEP: contact and cooperation. In Chapter 5, I conclude the dissertation by elaborating sociological implications for understanding individual and contextual factors affecting nonparticipation in NAEP, and unravel statistical impacts of nonresponse bias on NAEP estimates of educational performance. The implications of the findings are further discussed including interventions to improve adjustment for nonresponse bias and to reduce nonparticipation itself by tracing potential sources of nonparticipation in NAEP (e.g., student absenteeism, student refusal, and parental refusal on behalf of their children).

2. Theoretical Framework and Review of the Literature

I argue, according to social isolation theory of nonparticipation (e.g., Goyder, 1987; Groves, 1989, Groves and Couper, 1998), that students perceiving or experiencing "social isolation" (e.g., those feeling not supported in family, disengaged/not motivated in classrooms, or feeling insecure/unsafe in schools) are less likely to participate in an education survey assessment, “a temporary social event” where students are assessed about knowledge gained from established social institutions. For example, a student with little motivation in classrooms is more likely to skip a class. If students with less motivation or poor performance in classroom are also less likely to participate in a NAEP assessment, student acheivement in NAEP is likely to be overestimated. A student feeling insecure at schools troubled with gang activities is more likely to refuse participating in an assessment at school. At issue is how strongly correlated the assessment variables of interest are with nonparticipation propensity, the likelihood of nonparticipation in NAEP. I attempt to ground most key variables of interest in social isolation theory, as will be shown in the following sections.

Studies suggest that correlates of social isolation include demographic background factors, personal characteristics, and societal factors (Hortulanus, Machielse and Meeuwesen, 2006). Populations that are found to have high likelihoods of becoming social isolates include: the elderly, the sick and those with disability, people with lower incomes, lower educational levels, lower SES levels, and singles (e.g., Hess and Warning, 1978; Fisher, 1982). Personality characteristics that lead to becoming socially isolated include shyness, introversion, lack of social skills and the unwillingness to take social risks (Peplau and Perlman, 1982). Societal factors often associated with high social isolation include low participation in employment, club life, religious organizations, cultural activities, and volunteer work (House et al., 1982).

In order to test social isolation hypotheses in assessment survey context, I use strategies that provide us with data on key characteristics of respondents and non-respondents in NAEP by using the 2000 High School Transcript Study (HSTS) where characteristics of both participants and nonparticipants in the 2000 NAEP are contained. Because the transcripts for HSTS are collected from *all* students in the same sample of schools in which the NAEP 12th grade assessments are given, all students in NAEP assessment *including nonparticipating students* can be linked to the HSTS sample where characteristics of nonparticipants in NAEP can be studied along with that of participants from the social isolation perspective.

Participation in an assessment is an inherently tentative social process affected by personal and social factors as shown in Chart A-1 in Appendix. Thus I expect that full understanding of the process of assessment participation requires insight into key levels of influences simultaneously. I begin by exploring student- and school-level correlates of nonparticipation in NAEP by exploring variables that are justified by the construct of social isolation and evidenced in the literature. Next, I analyze the effects of key variables (i.e., student-level correlates and school-level correlates) on nonparticipation to evaluate their impacts in comparison with the current practice of NAEP merely involving some variables from the sampling frame.. Finally, I model them simultaneously across levels to understand the impact of a complete set of factors on nonparticipation in NAEP.

When I turn to assessing the impact of nonparticipation bias on NAEP estimates, I use the final multivariate model of nonparticipation propensity to adjust survey weights in order to account for differences in the probability of participation associated with student- and school-level correlates, which are grounded in social isolation construct. I apply an approach Abraham, Maitland and Bianchi (2006) used for nonresponse bias analysis, so I evaluate NAEP estimates calculated using weights that incorporate my own nonresponse adjustment based on a multivariate propensity model, in comparison with NAEP estimates calculated using NAEP final weights with a nonresponse adjustment.

2.1. Literature Review

Nonresponse or nonparticipation has long been recognized by social scientists as a key measure of survey quality, due to its potentially adverse effect on the ability to draw conclusions about a target population from a representative sample. The majority of nonresponse studies to date have been limited to bivariate analyses; consequently, theoretical frameworks that propose multiple influences have remained largely untested (Groves, Singer, and Corning, 2000). In order to develop methods to moderate the effects of nonresponse, recent studies have considered possible underlying micro and macro mechanisms driving nonparticipation and in turn attempted to evaluate nonresponse bias. This literature review examines studies of nonparticipation in various research realms, in an effort to develop a rich context and conceptual and analytical framework for the current study on the 12th grade nonparticipation in the National Assessment of Educational Progress (NAEP).

Nonresponse: Its Origins and Process

Nonresponse is the failure to obtain observations on some sample elements (Kish, 1965). Nonresponse rate, the percentage of the sample not observed, is often used mistakenly as a measure of quality of survey statistics, perhaps due to its easy documentation on many surveys. Nonresponse rates by themselves, however, take a number of different forms, depending on sources of nonresponse such as noncontact, refusal, and physical or mental incapacity (AAPOR, 2008; Groves, 1989; Kish, 1965). Understanding the origins of nonresponse is helpful for its control and reduction with proper intervention and for estimation of their distinctive effects on survey estimates of interest. Efforts to reduce noncontact can be distinguished from those to reduce refusals. When estimating nonresponse bias as referenced in the next section, knowledge about the underlying nonresponse mechanism helps to isolate factors that account for noncontact and refusal, respectively. In most telephone and face-to-face surveys, these three essential sources of nonresponse are readily distinguishable by interviewers. In mail and web surveys, however, they are generally indistinguishable from one another, as nonresponse is usually evidenced only by nonreturn of the questionnaire by mail or web.

Noncontact occurs when a sample person is not contacted by interviewers and hence never makes a decision about participation in a survey. Both refusals and the inability of the sample person to provide responses to the survey are generally viewed as requiring "contact" with the sample unit. Establishing contact with a sample unit is usually the first step in obtaining response. "Contactability" is a concept useful to understand the propensity for a sample unit to be contacted by an interviewer at any given moment in time (e.g., Groves and Couper, 1998; Stoop, 2005). In household surveys, Groves and Couper and Stoop both empirically tested and confirmed that contactability is a function of the three primary factors: physical impediments to accessing a sample unit, at-home patterns of a sample unit, and the timing and number of interviewer visits to the sample unit. In telephone surveys, Kish (1965) conceptualizes that not-at-homes depends on the respondent attributes (e.g., farmers are more available at home than urban workers, and housewives more than male employees) and the time of calls (e.g., daytime are bad for finding employed members of households, evenings and weekends being favorable interviewing hours). Empirical data Kish wished to support his argument related to at-home patterns have been steadily collected in a number of studies in subsequent years (e.g., Campanelli, Sturgis and Purdon, 1997; Groves, 1989; Groves and Couper, 1998; Stoop, 2005).

Refusals result from the direct denial given by the selected respondent, or from the denial of the interview by proxy (e.g., a mother refusing the interview to a selected teen child). Refusals are mostly considered permanent; Kish (1965) classified them as *unobtainable*, denoting a denial rather than a deferment of the observation, whether by interview, telephone, or mail. Kish's notion of "refusals" remains true today despite various causes of refusals by survey mode of data collection that now includes web and mixed modes. Groves (1989) insightfully distinguishes refusal nonresponse from other sources of nonresponse, especially noncontact nonresponse and the respondent's inability to answer the survey. In household survey Groves has studied, he finds that some sample persons in households are not measured because they cannot be contacted, because they are physically or mentally unable to respond, or because they refuse to cooperate with the interview request. Separating the effect on nonresponse of refusal from that of noncontact has guided research in subsequent decades. Historical trends indicate that the refusal component of nonresponse is increasing. (Brehm 1993; Groves and Couper 1998; de Leeuw and de Heer, 2002). In the re-analysis of nonresponse in the National Health Interveiw Survey (NHIS), a monthly cross-sectional personal interview survey, Groves (1989) reveals that the refusal rate is increasing although the other nonresponse categories are decreasing to reach a stable response rate from 1965 to 1985. He discovered that the proportion of the total nonresponse associated with refusals in the later 1960s is about 0.25, but it increases to the 0.60 range in the mid-1980s. The NHIS is a fine example of acquiring stable response rates despite losing to refusals those more typical of the full population. The Current Population Survey is another example that shows a steady increase of the refusal nonresponse component between 1965 and 1985 while the total nonresponse rate has been shown to be stable during this period (Groves, 1989). Breham (1993) reports that refusal rates for the National Election Study have climbed from well under 8 percent at inception in the early 1950s to the refusal rate near 25 percent in 1986. De Leeuw and de Heer (2002) further demonstrate that factors accounting for refusal are different from those for noncontact on the basis of analysis of nonresponse of time series for 16 countries and 10 various surveys.

Incapacity is when the physical or mental inability prevents the sample unit from providing answer to the survey. A respondent suffering from learning disability, illiterate, blind, or deaf would not be able to participate in a survey depending on mode of data collection. The survey capability is usually associated with the sample respondent’s age and health. In the National Election Studies, Brehm (1993) shows that capability declines with age such that sample person over 65 years old are the least likely to be capable of being interviewed. The elderly are generally more likely to vote in elections compared to the young; thus the election forecasting model, which is less sensitive to including the elderly, may mislead its biased estimates. Cohen and Duffy (2002) show in health surveys that the prevalence of common sources of ill-health in the over 75 population is likely to be underestimated as these old elderly are incapable of participating in health surveys. When the causes of incapability-based nonresponse are associated with survey estimates of interest, due attention is required to adjust for associated survey errors.

Groves and Couper (1998) were among the first researchers to demonstrate that nonresponse or nonparticipation is inherently a complex social process influenced by noncontact, refusal, and incapacity of the respondent. They investigated nonresponse in *household* interview surveys by analyzing several theorized influences on nonparticipation, including survey design, attributes of interviewers and participants, social interactions between interviewers and participants, and the social context in which the interview was initiated. Groves and Couper wove these constructs together to propose *social isolation* hypothesis, which maintains that those social isolates feeling out of touch with the mainstream culture of a society or those feeling cheated by larger society because of their membership in a group tend to ignore the norms of the larger society. Thus, those who are alienated or isolated from the broader society are less likely to comply with survey requests that represent such interests as “civic duty” of participating in voluntary surveys. Their tests of social isolation hypothesis relied on proxy indicators that are socio-demographic such as race/ethnicity, age, and gender. While acknowledging limitations of all these socio-demographic proxy measures of social isolation, Groves and Couper (1998) measured combined effects of these demographic variables along with proxy measures of social isolation at household level (e.g., single-person household, presence of children, household mobility, and type of housing structure) in multiple logistic regression models.

Relevance of social isolation or social integration to noncontact is somewhat elusive, considering that contactability is primarily a function of physical impediments to accessing a sample unit, at-home patterns of a sample unit, and the timing and number of calls or interviewer visits to the sample unit. Despite such a limitation, Lepkowski and Couper (2002) used the same set of social integration variables (e.g., marital status, summary measures of contacts with friends, relatives, and others) to model location (i.e. comparable to contact in panel surveys) and cooperation propensity in panel surveys. They showed various forms of social integration to be well associated with both location and cooperation in panel surveys of National Election Studies and Americans' Changing Lives survey. Other studies relate social isolation to individuals' living environment with a premise that the spatial environment can support or discourage social contacts (Hortulanus, Machielse and Meeuwesen, 2006). For example, a neighborhood in which most people are at work during the day offer few possibilities for social contact; chances of social contact are minimized when people no longer feel safe in their neighborhood due to high crime rates.

Groves, Cialdini, and Couper (1992) give attention primarily to the interaction between the respondent and the interviewer in their investigation of the participation process. The decision to participate in a survey in interviewer-administered surveys is primarily affected by the initial conversation between the interviewer and the respondent. They conceptualize that such an interaction process working towards compliance is influenced by several influences grounded on a peripheral persuasion approach. Whereas "central persuasion cues" refer to ideas and supporting data that bear directly upon the quality of the arguments in the message, "peripheral persuasion cues" include such factors as the attractiveness and expertise of the source, the mere number of the arguments presented, and the positive or negative stimuli that form the context within which the message was presented Those peripheral influences on compliance are 1) reciprocation that favor requests from those who have previously given something to you (e.g., survey incentives), 2) commitment and consistency that drives to behave in a similar way over situations that resemble one another (e.g., foot-in-the-door effect), 3) social validation that invokes behavior in ways similar to those like us (e.g., "all your neighbors participated in this survey.", 4) liking that complies with requests from attractive requestors (e.g., interviewers liked by respondents), 5) authority that invokes compliance with requests endorsed by those in positions of legitimate power (e.g., a survey sponsored by the federal government), and 6) scarcity that values rare opportunity (e.g., your reply representing hundreds of other samples).

Groves, Singer, and Corning (2000) extended the principles of survey compliance by Groves, Cialdini, and Couper (1992) and the framework of Groves and Couper (1998) by developing the "leverage-saliency theory" of survey participation. This theory postulates that the effect of any particular stimulus on a sample person's participation is a joint function of its centrality to the person (leverage) and its salience relative to other stimuli in the survey introduction. For example, in a survey about community issues, the survey questions have high leverage for a sample person with high community involvement. Such attributes of the survey topic can positively affect response propensity if the community aspect is made salient in the request to participate. On the other hand, making it salient would not increase response propensity by a sample person with low community involvement. Leverage-salience theory has been used to generate hypotheses about how survey design features such as mode of data collection, topic interest, and monetary incentives influence participation decision or response propensity.

Until recently, few studies have examined nonparticipation in the context of education research. Employing Groves and Couper's (1998) approach to analyzing nonparticipation using the social isolation theory as a model, Chun and Scott (2003) sought to illuminate nonparticipation behavior of teachers in Schools and Staffing Surveys, based on a similar theoretical framework. The authors emphasized that efforts to reduce nonresponse errors in teacher surveys require an understanding of the complex social situation in education surveys.

Recently, the educational research community has begun to focus on nonparticipation in education surveys and assessments. The 1999 International Conference on Survey Nonresponse included a few papers addressing nonresponse errors in school-oriented surveys. Furthermore, in 2004, the National Assessment Governing Board commissioned studies exploring motivation and nonparticipation of 12th graders in NAEP.

When examining student nonparticipation in education assessments, it is important to take into account both the broad social context in which this behavior takes place and the individual context where nonparticipation takes place. The current dissertation proposes that student nonparticipation in NAEP is influenced by school-level influences as well as student characteristics. Thus understanding the complex effects of all these factors across levels is as important as dissecting influence of each level of characteristics on nonparticipation. School factors may include: type of school (public versus private), urbanicity, school size, percentage of minorities, school region, percentage free or reduced lunch (in public schools), teacher-to-student ratio, and school climate. Student characteristics may include: grade, race/ethnicity, gender, SD/LEP, absenteeism, academic performance, level and quality of course-taking, and household variables. This framework for examining both individual and school influences is based on a growing body of research that demonstrates that macro- and micro-level factors influence the public's willingness to participate in surveys. (See Chart 1 in Appendix for a nonparticipation model I propose for NAEP.)

When further exploring studies that investigate the impact of broad societal influence on participation in survey, a few studies stand out. For example, Schleifer (1986) points out that public goodwill must be a priority for the survey research community because the success of survey research depends on the public's willingness to participate in its surveys. For this reason, Walker Research has conducted a biennial Industry Image Study since 1974, a study that examines the public's attitudes toward the survey research industry. Schleifer summarized the findings of the 1984 Industry Image Study, which measured "participation levels in survey, attitudes toward the interview experience, and feelings about the survey research industry." Chanley, Rudolph, and Rahn (2000) developed a measure of trust in the U.S. government from 1980 to 1997, and conducted the first multivariate time series appraisal of public trust in government. These results provided further evidence of the influence of public concern about crime, and provided new evidence of how declining levels of trust in government may influence elections and domestic policy-making. This information provides insight into the current study because parents of assessment participants are informed that NAEP is an assessment for the Department of Education, a well-known government agency. It is therefore likely that public trust in government may be an influence on the decision of whether parents allow their children to participate in NAEP.

Nonresponse Bias

Studies of nonresponse bias have been informed by studies of sources and process of nonresponse as discussed above. The magnitude of nonresponse bias is a function of both nonresponse rate and the extent to which nonrespondents are different from responondents on statistics of interest(Groves, 1989; Groves and Couper, 1998). That is, in cases of a sample of fixed size, the bias of the respondent mean is approximately:

$$B(Y_r) = nr/N\ (Y_r - Y_{nr})$$

Where

$B(Y_r)$ = Bias of respondent mean;

nr = Nonresponse size

N = Sample size

Y_r = Respondent population mean

Y_{nr} = Nonrespondent population mean

or

Bias (Respondent Mean) = (Nonresponse Rate in Population) x
(Difference in Respondent and Nonrespondent
Population Mean).

This formula indicates that the higher the nonresponse rate, the greater the bias of the respondent mean, and the greater the difference between nonrespondents and respondents, the larger the bias of the respondent mean. Best practices in surveys have been to reduce nonresponse rate in order to reduce nonresponse bias without paying due attention to the second essential component of nonresponse bias, the extent to which nonrespondents are different from respondents on statistics of interest. A traditional notion of linking high nonresponse rate to high response bias, however, has been recently challenged by several studies (Curtin, Presser, and Singer 2000; Keeter et al., 2000; Merkele and Edelman 2002) that individually demonstrated no strong relationship between nonresponse rates and nonrespnse bias. Groves (2006) further demonstrated by meta-analyzing 235 estimates from 30 studies that there is little empirical support to tie nonresponse rates to nonresponse bias. He persuasively showed that the central question is rather to investigate how strongly correlated the survey variable of interest is with response propensity, the likelihood of responding. With this perspective, the bias of the respondent mean approximates:

$B(Y_r) = Cov(Y, r)/ R$

Where

$B(Y_r)$ = Bias of respondent mean;

Y_r = Respondent population mean

r = Response propensity

R = Mean propensity in the target population

or

Bias (Respondent Mean) = (Covariance between survey variable, y, and response propensity, r) / (Mean propensity, R, in the target population)

Studies in the same special issue of Public Opinion Quarterly (Abraham, Matland, Bianchi, 2006; Groves et al., 2006) were motivated by the same concern about the presence of covariance between response propensity and the survey variables of interest. Furthermore, the study by Abraham, Matland, Bianchi (2006) demonstrated how a theory of “social integration” can guide selection of key variables in the logistic regression model that was eventually used for recalculating weights that account for differences in response propensities. This study, which elaborated a construct of social integration by Lepkowski and Couper (2002), stands out as most studies of nonresponse or nonparticipation are grounded on no theory as Goyder (1987) and Brehm (1993) called for. Groves et al. (2006) empirically discovered that the common influences on response propensity and the survey variable of interests are reactions to the survey sponsor, interest in the survey topic, and the use of incentives. Abraham, Helms, and Presser (forthcoming) demonstrated how the strong association between the causes of volunteering and the causes of survey participation was likely to overestimate hours of volunteering in the American Time Use Survey, thus showing the significant effect of the covariance term. Further in a meta-analysis of 959 estimates from 59 studies designed to estimate the magnitude of nonresponse bias, Groves and Peytcheva (2008) concluded that high response rates are not necessarily likely to reduce the risks of bias when the cause of participation is highly correlated with the survey variables. They strongly recommended exploring how each survey variable relates to causes of survey participation in order to predict what survey estimates are most susceptible to nonresponse bias.

Methods for assessing nonresponse bias are as diverse as causes and consequences of nonresponse (Groves, 1989; Groves and Couper, 1998; Groves and Peytcheva, 2008). I order these methods by the strength that they are reportedly valid and reliable when measuring nonresponse bias. The latest innovation to study nonresponse bias is to conduct experiments that attempt to produce variation in response rates across groups known to vary on statistics of interest (Groves, Presser, and Dipko, 2004; Groves et al., 2006). Experimental studies are most desirable to understand the specific conditions under which statistics of interest and response propensity are associated with each other. However, it is still difficult to separate the effect of nonresponse bias from measurement error as Groves et al. (2006) acknowledge. It is often not feasible to create experimental tests of various individual and social factors that affect nonresponse. It is premature to evaluate experimental benefits of nonresponse bias assessment due to mere lack of case studies at the moment.

A common approach to studying nonresponse, as Groves has repeatedly acknowledged over two decades (1989, 1998, and 2006), is the use of sampling frame or supplemental matched data available for both respondents and nonrespondents. In cases where records are available as a sampling frame or for matching (e.g., .Presser, 1981; Lin and Schaeffer, 1995; Abraham, Maitland, and Bianchi, 2006; Abraham, Helms, and Presser, forthcoming), estimates of nonresponse bias are constructed using frame or externally matched variables. The utility of this method is limited by the extent to which variables available in the frame or matched data are variables of key interest for a given survey. The accuracy of the data on the records is also subject to measurement errors, missing values, and other sources of survey errors.

Nonresponse follow-up studies are frequently conducted to compare estimates of respondents across key phases of data collection based on the assumption that reluctant respondents are proxies for final nonrespondents (e.g., Dunkelberg and Day, 1973; Smith, 1984; Groves and Wissoker, 1999). In a culture that values high response rate, nonresponse follow-ups have been routinized in most surveys; it is thus convenient and easy to identify studies of various interests. The value of this approach is, however, constrained by the empirically unconfirmed notion of a continuum of nonresponse ranging from the cooperative respondent through the reluctant respondent or the difficult-to-contact, to the hardcore nonrespondent. Studies (e.g., Curtin, Presser, and Singer, 2000; Guadagnoli and Cunningham, 1989) have failed to find evidence that converted nonrespondents substantially change survey estimates. The method does not address the characteristics of refusals.

Comparing response rates across key subgroups, usually derived from a sampling frame, is the most frequently used, yet the least valid method of nonresponse bias (e.g., Brick et al., 2003; Westat, 2003a and 2003b). Perhaps Groves (2006) listed it as the first method for assessing nonresponse bias to address the survey practitioner's attention to its misuse. It is easy to show the distribution of response rates across key background variables such as gender, race/ethnicity, age, socio-economic status, and census region as they are available from a sampling frame. However, it is premature to infer about nonresponse bias based on mere comparisons of response rates by subgroups. Subgroup variables used for such an analysis are not necessarily the only potential causes that affect both response propensity and survey variables of interests.

Each of the methods of nonresponse bias analysis has strengths and weaknesses; thus, using multiple methods simultaneously would complement each other as long as the focus is maintained on evaluating the covariance between response propensity and the survey variables of interests. Groves and Peytcheva (2008) is the latest comprehensive attempt to identify the circumstances that produce a relationship between nonresponse rates and nonresponse bias by combining most of methods as discussed above. The 59 studies, from which 959 estimates of nonresponse bias were extensively analyzed, appear to include only a single study conducted by National Center for Education Statistics. The database of nonresponse bias is concentrated in the biomedical field reportedly due to the availability of matched records.

Assessing Nonresponse Bias in NAEP

Nonparticipation in the National Assessment of Educational Progress is generally the consequence of: 1) refusal by a sample student to complete the assessment, 2) failure of the sample student to be present on the day of the assessment session (absence), or 3) other reasons including the sample student's incapability to take assessment due to disability or limited English proficiency. According to the NAEP disposition guidelines Assessment Administrators use on the day of NAEP assessment, there are over 30 disposition codes of participation outcomes (See Chart A-3 in Appendix). In NAEP, being assessed refers to those assessed in original or makeup session with usable data. Refusal occurs when 12th grader or their parents (on behalf of their children) refuse to participate in the assessment. 12th graders' absence in NAEP assessment happens for various reasons: temporary (less than two weeks) or long-term illness or disability, in-school suspension, and scheduling conflicts with a sporting event usually by athletics. Other reasons of nonparticipation, according to NAEP disposition codes, are usually tied to ineligibility such as withdrawal from school or disability.

In accordance with NCES Standards 4-4-1 and 4-4-2, NAEP carries out the nonresponse bias analysis, when response rates fail to meet the required NCES standard of 85%, by using base weights for each survey stage. The existing nonresponse bias method in NAEP relies on a few school-level variables in NAEP such as type of reporting group (public vs. private school), school location (urbanicity), census region, and school size measured by student enrollment.   The student-level variables selected for nonresponse bias are usually restricted to gender, age, race/ethnicity, and proxy measure of socio-economic status measured by student's eligibility for the national school lunch program. The NAEP method for assessing nonresponse bias minimally satisfies statistical standards of the National Center for Education Statistics (2003) as follows:

> "Any survey stage of data collection with a unit response rate less than 85 percent must be evaluated for the potential magnitude of nonresponse bias before the data or any analysis using the data may be released. Estimates of survey characteristics for nonrespondents and respondents are required to assess the potential nonresponse bias. The level of effort required is guided by the magnitude of the nonresponse."

There have been two types of nonresponse bias analysis conducted by NAEP: 1) comparison of respondents and nonrespondents across subgroups available from the sample frame, and 2) multivariate modeling to compare the proportional distribution of characteristics of respondents and nonrespondents to determine if nonresponse bias exists and, if so, to estimate the magnitude of the bias. The former approach is constrained by limited utility and number of frame variables which are not necessarily related to response propensity as well as variables of interest in NAEP. Asserting no evidence of nonresponse bias on the basis of similar distribution by subgroups is misleading. When this method finds certain variables associated with response, findings are reported without evaluating the direct impacts on NAEP estimates of potential nonresponse bias. The latter approach, while designed to identify the characteristics of samples least likely to respond, is limited by the extent to which predictors of interest exist only within NAEP sampling frame. For example, Westat (2003a) used limited NAEP sampling frame variables to conduct logistic regression analysis for predicting private school nonresponse for the grade 4 and grade 8 assessments in Reading and Mathematics in the 2003 NAEP. Westat (2003b) modeled in logistic regression analysis response outcome as a function of NAEP reporting group, type of school location, census region and size of school, which are all available from the sampling frame.

There have been no data available for evaluating the direct effect on NAEP achievement estimates of nonresponse bias. Nonresponse bias analysis reports prepared by NAEP have not conjectured as to the likely magnitude of any nonresponse bias in the NAEP student achievement results. Technical comments have been extremely limited in the widely used Nation's Report Cards on the perceived degree of success that has been attained in controlling NAEP nonresponse bias through the use of nonresponse adjustments. It is an untenable assumption that the sampling frame-based variables currently selected for assessing NAEP nonresponse bias are the only potential common causes affecting response propensity and NAEP statistics of interest.

3. Research Methodology

I begin this chapter by describing the data sources and characteristics of the sample I use for the study. I identify key variables selected for analysis and their relevance to social isolation construct applied to my research. I complete this chapter by detailing a plan of bivariate and multivariate analysis to identify correlates of nonparticipation in NAEP and of comparing alternative nonresponse weighting adjustments in NAEP to study their impacts on NAEP estimates.

3.1. Data Sources and Sample Design

The data I use for this dissertation come from two sources collected by the National Center for Education Statistics: 1) the 2000 NAEP survey assessment of 12th graders and survey of their teachers and principals, and 2) about 20,000 12th graders in the 2000 High School Transcript Study (HSTS) linked to the 2000 NAEP. Because the transcripts for the 2000 HSTS are collected from *all* students in the same NAEP sample of schools regardless of individual student's participation status in NAEP, rich analysis of correlates of nonresponse and robust assessment of nonresponse bias is possible. The subject areas that are assessed change across assessment cycle (See Chart A-2 in Appendix). In 2000, mathematics and science were assessed at all three grades (4, 8, and 12) for national main assessments of NAEP. As the 2000 HSTS collected transcript data just for mathematics and science, the joint NAEP and HSTS data used for nonresponse analysis focus on these two subjects.

The National Assessment of Educational Progress

The National Assessment of Educational Progress (NAEP) is the only nationally representative and continuing assessment of what America's students know and can do in various subject areas. NAEP provides a common yardstick for measuring the progress of student performance for the nation at grades 4, 8 and 12, states currently at grades 4 and 8, and in some cases, selected urban districts. For national assessments including grade 12, students in public and private schools are assessed, but at the state level, assessment is carried out in public schools only currently just for grades 4 and 8. Assessments are conducted periodically in mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history. NAEP subjects change across assessment cycle as summarized in Chart A-2 in Appendix. NAEP is based on representative samples of students at grades 4, 8, and 12 for the *main* assessments every two years, or samples of students at ages 9, 13, or 17 years for the *long-term* trend assessments every four years that allows the performance of today's students to be compared with those from more than 30 years ago. These grades and ages were chosen because they represent critical junctures in academic achievement. For the 2000 NAEP linked to HSTS, the focus data of this dissertation research, the main assessments of mathematics and science were both conducted at grade 12.

NAEP provides results on subject-matter achievement on a scale of 0-300 or 0-500 points, instructional experiences, and school environment for populations of students (e.g., all 12th-graders) and groups within those populations (e.g., female students, Hispanic students, Black-White performance gap). NAEP can not provide scores for *individual students or schools* assessed. Because NAEP is a large-group assessment, each student takes only a small part of the overall assessment. In most schools, a small sample of the total grade enrollment is selected to take the assessment, and these students may not reliably or validly represent the total school population. Only when the student scores are aggregated at the state or national level are the data considered to be reliable and valid estimates of what students know and can do in the content area; consequently, school- or student-level results are never reported.

NAEP score scales are created via Item Response Theory (IRT, Lord, 1980) and scale score distributions are estimated for groups of students. IRT is a procedure of test analysis that assumes a mathematical model for the probability that an examinee will respond correctly to a specific test question, given the examinee's overall performance and characteristics of the questions on the test. NAEP score scales summarize student performance for the collection of assessment items representing the academic content specified in the NAEP frameworks specific to assessment subject. For each subject area (e.g., mathematics, science, reading), the framework determines the number of IRT scales. Each framework, developed by the National Assessment Governing Board, provides: the theoretical basis for the assessment, the direction for what types of items should be included in the assessment, how the items should be designed, and how the items should be scored. IRT models are used to describe the relationships between the item responses provided by students and the underlying score scales. IRT provides a common scale on which the performance of students receiving different blocks of items can be placed.

When the score scales are created, the parameters describing the item response characteristics are estimated (Mislevy and Bock, 1982; Muraki and Bock, 1997). NAEP does not produce individual test scores but does produce estimates of scale score distributions for groups of students classified by key background variables. The resulting scale score distributions describing student performance are transformed to a NAEP scale, and summary statistics of the scale scores are estimated. Statistical tests are used to make inferences about the comparisons of results for different groups of students or for different assessment years of NAEP.

Because NAEP scales are developed independently for each subject and each grade, scores should not be compared across subjects or grades. To provide a context for interpreting student performance, NAEP results are also reported as percentages of students performing below the *Basic* level, at or above the *Basic* and *Proficient* levels, and at the *Advanced* level.

The High School Transcript Study and its Linkage to NAEP

NCES has conducted a number of transcript studies of 12th graders since 1982 initially in conjunction with the first follow-up survey of the High School and Beyond Study. HSTS focuses on high school graduates' course-taking patterns, including the courses they took in different subject areas and the grades they received for those courses, whereas NAEP measures educational achievement in various subject areas for 12th-grade students. That is why the data linked at the student level are a rich source for examining the relationship between student course-taking patterns and educational achievement in select course subjects, as measured by NAEP. Beginning with the 1990 transcript study, HSTS has been conducted in conjunction with NAEP, including in 1994, 1998, 2000, and 2005. Among all these data files, the 2000 NAEP-HSTS linking data of mathematics and science are the one used for this dissertation research, and they contain transcript data including course-taking patterns linked to NAEP achievement data and various student and school variables. NCES provided about 80 disposition codes of student participation outcome for this 2000 linking data as shown in Chart A-3 in Appendix. Thus the 2000 joint NAEP-HSTS are a rich data resource that provides us with measures of individual characteristics including achievement factors and absenteeism and school characteristics for nonparticipating students as well as participating students in NAEP assessments.

Eligible schools participating in NAEP were informed about the HSTS 2000 when they received information about NAEP. Schools were provided with information about participating in the HSTS, including procedures that would be used to ensure confidentiality of the data, and the amount and nature of school staff time required for participating in the HSTS. For schools that agree to cooperate, students sampled for NAEP were all included in the HSTS sample. Transcripts were requested for all students who were assessed, and for sampled students who were absent or refused during NAEP assessment. In order for a transcript to become part of the "linked" database, both a completed NAEP assessment and a completed usable transcript from HSTS must be obtained for a student. This link enables one to identify the correlates of nonparticipation beyond NAEP variables, and assess the impact on NAEP estimates of nonresponse bias. It is noted that the linked database is to some extent limited by its own nonresponse. In the 2000 HSTS, there were 287 NAEP participating schools that were included in the HSTS study. Transcripts were collected from 261 NAEP schools. Thus the weighted school response rate equaled 93.3 percent while their weighted student response rate was 99.4 percent. The overall response rate for the 2000 transcript study’s NAEP students equaled 92.7 percent. As a result, there is about 7.3 percent of nonresponse among NAEP students in the 2000 HSTS sample I use for the study.

Sample Design

The HSTS 2000 is based on a sample of the schools and students included in the NAEP 2000. The 12th grade sample for the 2000 NAEP linked to the 2000 HSTS was a multistage probability-based sample of students. This was a national sample in which counties or groups of counties were the first-stage sampling units, and schools were the second stage units. The third stage of sampling consisted of the assignment of session type and sample type to sampled schools. The session type refers to the subject(s) being assessed (i.e., mathematics and science), while the sample type refers to the specific criteria for inclusion which were applied to the session. The fourth stage involved selection of students within schools and their assignment to session types. A total of 94 primary sampling units (PSUs) were included in the NAEP sample, and a sample of 248 schools actually linked to the 2000 HSTS (223 public schools and 25 private schools). Over 20,000 student data were linked between NAEP and HSTS including18,513 students from public schools and 1,034 students from private schools. The overall participation rate of 12th graders in 2000 NAEP assessments ranged from 62% to 64% depending on subjects (mathematics and science) and sample type related to provision of accommodations to students with disability.

3.2 Key Variables and Their Relevance to Social Isolation Theory

The outcome variables of NAEP interest are: 1) assessed, 2) absenteeism, 3) refusal that includes student and parental refusal on behalf of their children, and 4) other reasons of nonparticipation. I use over 80 official disposition codes of NAEP assessment to classify them into these major categories of participation outcome in close consultation with NCES and the NAEP participation guidelines (See Chart A-3 in Appendix). When there is a question about any classification, the NAEP experts of NCES are consulted and it is determined to make a reasonable classification together. For example, in case of a student left in the middle of the assessment, it is determined it is like a respondent refusing in the middle of a survey and incomplete data are not usable. It is not absence as the student showed up yet refused in the middle of the assessment session.

In NAEP, being assessed refers to those assessed in original or makeup session with usable data. 12th graders' absence in NAEP assessment happens for various reasons. Absence may be temporary (less than two weeks) or long-term depending on the nature of illness or disability. Students may be absent because of in-school suspension due to disruptive school behavior. Members of an athletic team may often be absent because of scheduling conflicts with a sporting event. Some teachers may not release students from their classes for whatever reason. According to the NAEP guidelines, refusal occurs when a 12^{th} grader refuses to participate in the assessment before being given a NAEP assessment booklet. Parents may refuse on behalf their children by notifying school of their unwillingness to allow their children to participate in NAEP assessment for whatever reason. Other reasons of nonparticipation, according to NAEP disposition codes, are usually related to ineligibility such as withdrawal from school. The initial sample size is 23,522 students who were included in the 2000 HSTS. The NAEP-linked HSTS sample is 20,549 after dropping 1,512 students not linked to NAEP and ineligibles. The eligible sample of 20,549 used for this study consists of the following:

15,220 students who participated, 3,320 students who were absent, and 2,009 students who refused or whose parents refused participating in NAEP assessment on behalf of their children. Thus the weighted participation rate only at student level for the NAEP-linked HSTS sample is 75.1 percent as it is shown later in Table 2. The reader is reminded the HSTS student sample is obtained from NAEP schools that agreed to cooperate. If the school-level response rate is accounted for, the overall school and student combined response rate for the 2000 NAEP-linked HSTS sample would be comparable to or somewhat lower than the overall response rates of 55-60 percent in 2000 as reported in Figures A-1 and A-2.

To test social isolation hypotheses about NAEP nonparticipation, the HSTS-NAEP linking variables I extract for analysis are proposed below including student correlates, school correlates, and social psychological school climate variables (See Table 1). All of these explanatory variables are proxy measures of social isolation except a couple of control variables such as student gender and census region of school location. I include a category with missing values, where applicable, to keep all cases for analysis. I continue to include them all in subsequent multivariate analysis so that I could eventually develop response propensity based weights for all individual valid cases and use all of them in turn to re-estimate NAEP scale scores.

Ideally I wish to include personality measures of social isolation, as Hortulanus, Machielse and Meeuwesen (2006) suggested, for making a close link between social isolation and participation in NAEP. Such a social psychological measure of social isolation could include the scale of shyness, introversion, and lack of social skills. It is also my wish to measure school-level factors of social isolation/integration by tapping students' involvement in study groups, after-school activities, religious organizations, and volunteer activities in order to associate the scope of these voluntary activities with participation in NAEP. However, the secondary analysis of the NAEP and HSTS data has constrained explanatory variables to the list presented below.

Table 1. Sample distribution by social isolation proxy predictors of participation outcome in NAEP Mathematics and Science: by student and school correlates, 2000

		Unweighted N	Unweighted Percent	Weighted Percent
	Overall	20549	100	100
Student Correlates				
Race/ethnicity	White	11382	55.4	67.7
	Black	3823	18.6	13.5
	Hispanic	3877	18.9	13.2
	Other race/ethnicity[1]	1467	7.1	5.6
Taking Advanced Mathematics	No	16949	82.5	83.6
	Yes	1577	7.7	8.5
	No records	2023	9.8	7.9
Taking Advanced Science	No	16955	82.5	84.0
	Yes	1571	7.6	8.1
	No records	2023	9..8	7.9
Carnegie Credits	<24	4543	22.1	21.6
	24-28	10344	50.3	51.8
	> 28	3599	17.5	18.5
	No records	2063	10.0	8.1
GPA	<=2.00	3006	14.6	12.4
	2.01 - 3.00	9373	45.6	43.7
	3.01-4.00	7947	38.7	42.5
	Others	223	1.1	1.4
Eligibility for National School Lunch Program	Ineligible	11760	57.2	61.2
	Eligible	4066	19.8	14.1
	Unknown	4723	23.0	24.7
School Correlates				
School Location	Urban	6868	33.4	25.7
	Suburban	9525	46.4	51.2
	Rural	4156	20.2	23.1
School Type	Public	19508	94.9	91.6
	Private	1041	5.1	8.4
School Enrollment	< 500	2095	10.2	15.1
	500 - 900	2583	12.6	13.6
	> 900	14166	68.9	63.4
	No records	1705	8.3	8.0

Continued - Table 1. Sample distribution by social isolation proxy predictors of participation outcome in NAEP Mathematics and Science: by student and school correlates, 2000

		Unweighted N	Unweighted Percent	Weighted Percent
Social psychological school climate variables				
Problem with gang activities	Serious or moderate	1639	8.0	4.4
	Minor or not a problem	17193	83.7	88.0
	No records	1717	8.4	7.6
Teacher absenteeism	Serious or moderate	4450	21.7	19.6
	Minor or not a problem	14161	68.9	71.8
	No records	1938	9.4	8.6
Parental support of student achievement	Very or somewhat positive	16908	82.3	84.7
	Somewhat or very negative	2028	9.9	8.6
	No records	1613	7.8	6.8
Other Control Variables				
Student gender	Male	9849	48.0	47.6
	Female	10663	52.0	52.4
Census Region of school	Northeast	11382	55.4	17.5
	Midwest	3823	18.6	24.1
	South	3877	18.9	38.9
	West	1467	7.1	19.5
NAEP Assessment Student Completed	Mathematics	9163	44.6	44.6
	Science	11386	55.4	55.4

NOTE: N is 20549. Totals are not 100 percent due to rounding. All correlates except student gender, census region of school and assessment subject are proxy measures of social isolation. Carnegie Credits refer to the number of credits a student received for a course taken every day, one period per day, for a full school year; a factor used to standardize all credits indicated on transcripts across the study. To compute GPA, points are assigned to each letter grade as follows: A=4 points; B=3 points; C=2 points; D= 1 point; F= 0 points. The points are weighted by the number of Carnegie credits earned, so that a course with 120 hours of instruction counts twice as much as one with 60 hours. The average of the points earned for all the courses taken is the grade point average. Courses in which a graduate did not receive a grade, such as pass/fail and audited courses, do not factor into the GPA calculation. 1. Other race/ethnicity includes Asian-Pacific American and American Indian.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, High School Transcript Study (HSTS), 2000

Student-level correlates of nonparticipation I plan to assess in the context of social isolation hypotheses are as follows: race/ethnicity, student eligibility for national school lunch program, takers of advanced mathematics or science courses (yes vs. no), GPA, Carnegie credits and other individual variables that are found to be significantly related to nonparticipation or student's academic performance as evidenced in literature.

Studies have shown that ethnic minority students such as Blacks and Hispanics tend to feel more isolated or insecure at schools. For example, there is an added psychological strain experienced by Black students (or Hispanic students) who enter a school environment dominated by White students (Roach, 2001). Roach continues that the alienation and estrangement felt by minority students can affect motivation, which in turn affects self-esteem and the sense of academic confidence required to do well at schools. Students of color are often shut out of more important networks, such as study groups. Isolation, whether it is intended or unintended, denies minority students access to the benefit of high achieving study group. As a result, some students of color tend to stay adrift either studying alone or not coming to schools at all. It is not a surprise to learn that minorities constitute the majority of high school students who failed to graduate (Swanson, 2006).

Twelfth graders who are eligible for free or reduced school lunch program usually come from socioeconomically disadvantaged families. The National School Lunch Program (NSLP) is a federally-assisted meal program that provides nutritionally balanced, low-cost or free lunches to children each school day. National income guidelines determine the eligibility of students based on their families' household size and income. Children from families with incomes at or below 130 percent of the Federal

Poverty Guidelines (FPG) are eligible for free meals. Those with incomes between 130 percent and 185 percent of the FPG are eligible for reduced-price meals. The majority of students eligible for NSLP are ethnic minorities. Such disadvantaged 12th graders with less socioeconomic support tend to feel isolated or academically disadvantaged at school. Studies support that children from families with low socioeconomic status make up a disproportionate number of those most at-risk for school failure (Knapp and Shields, 1990). Research consistently shows that living in concentrated poverty decreases schooling opportunity, academic achievement and quality of life (Lee and Smith, 2001). They confirm that students with higher SES were more academically engaged and successful than students with lower SES. Among high school students, low SES increases a psychological strain, resulting in further alienation and estrangement which in turn deflates self-esteem and damages academic motivation to compete with peers with high SES.

Students with poor academic performance, usually measured by GPA (or standardized Carnegie credits), are more likely to be intimidated or subject to bullying at schools where academic achievers tend to be liked by peers. Student engagement and achievement studies suggest a fine link between academic achievement and school engagement behavior (e.g., Mcevoy and Welker, 2000; Newman, 1992). Research consistently demonstrates that student engagement has a strong positive effect on academic performance (Appleton, Christenson, and Furlong, 2008). Student engagement refers to “students' willingness to participate in routine school activities, such as attending class, submitting required work, and following teachers' directions in class” (Chapman, 2003). That includes participating in the activities offered as part of the school program perhaps including participation in NAEP. The opposite of engagement is disaffection (Skinner, Belmont, 1993). Disaffected students are passive and give up easily in the face of challenges; they tend to withdraw from learning opportunities.

School-level correlates of nonparticipation I suggest include school location (urban, suburban, and rural), school type (public vs. private), and school enrollment size (0-500, 501-900, and 901 or more). From nonresponse research (e.g., House and Wolf, 1978; Brehm, 1993), it is well established that residents of inner-city areas of large metropolitan area exhibit the lowest level of cooperation, while those in rural areas have the highest rate of cooperation. As Groves and Couper (1998) pointed out, effects of urbanicity found in the literature may be explained in terms of greater population density, higher crime rates, and social disorganization that are often associated with life in large urban areas. I argue it is proper to apply this line of hypothesis to student nonparticipation in NAEP and expect that students in urban schools are more susceptible to negative facts discouraging their participation in NAEP. Students in public schools, in comparison with those in private schools, tend to come from economically less disadvantaged families and are likely to feel less engaged in schools.

To test social isolation hypotheses, I also use school-wide *social psychological* correlates of nonparticipation such as perception of problem activities at school, teacher absenteeism, and parental support of student achievement. As supported by the research of school climate, students feeling insecure at schools troubled with gang activities are less likely to attend schools or more likely to skip classes (Gottfredson, 1989; Howell and Lynch, 2000). Individual gang participation – and rates of gang participation in schools – is strongly associated with fear (or perceptions that the school environment is not safe), drug involvement, and other forms of deviant behavior. The association of perceptions that the school is unsafe with gang participation rate is especially strong.

Schools with a high proportion of teacher absenteeism, and poor parental support of student achievement are less likely to motivate students to engage in school activities (Miller,1980; Neuman et. al, 1995). Studies report a negative relationship between teacher absences and student achievement or other academic activities (Bayard, 2003; Cantrell, 2003;). When a regularly assigned teacher is absent, instructional intensity may be seriously reduced and regular routines of instruction may be disrupted. Low skill levels of substitute teachers may contribute to further reduction, not improvement, in instructional focus. Studies indicate that parental involvement is associated with higher student achievement outcomes (Epstein, 2001). These findings emerged consistently whether the outcome measures were grades, standardized test scores, or a variety of other academic measures including student engagement.

I consider other control variables such as student gender (male and female) and census region of school location (northeast, midwest, south, and west) as potential influences on participation in NAEP.

4. Analysis Plan

The current study of the12th grade nonparticipation in NAEP begins with the constructs of nonresponse Groves and Couper (1998) proposed in order to explore and isolate student- and school-level correlates of nonparticipation in NAEP. Next I conduct bivariate analysis to understand how strongly each identified variable is associated with participation outcome in NAEP. I turn to multivariate analysis to explore the extent to which a set of level-specific variables of social isolation affect nonparticipation. Multivariate analysis is further expanded so as to combine a model of school indicators of social isolation with student-level indicators of social isolation and understand which variables are more essential to understanding nonparticipation in NAEP.

The final nonparticipation propensity model I find most fitting to the data is what I use for evaluating the impact of nonresponse bias on NAEP estimates. Applying a model of nonresponse bias analysis by Abraham, Maitland and Bianchi (2006), I evaluate NAEP estimates reweighted using my new weights that incorporate nonresponse adjustment based on propensity model in comparison with NAEP estimates calculated using the current NAEP final weights with a nonresponse adjustment.

All analyses are performed using SAS and SPSS, and weights are properly accounted for the complex multi-stage clustered NAEP sample design. Re-estimating NAEP scale scores with alternative nonresponse adjustment is carried out by using WesVar that properly handles the complex NAEP sample design and variance estimation.

4.1 Identifying Determinants of Nonparticipation in NAEP

The analysis follows the approach by Abraham, Maitland, and Bianchi (2006) who examined correlates of different outcomes and the extent to which key correlates affect key survey estimates. In NAEP context, the analysis is focused on three major outcomes: 1) response referring to being assessed; 2) noncontact being parallel to absence; and 3) refusal referring to refusal by students or parents on behalf of their children to take NAEP assessments.

Bivariate Analysis

The test of social isolation theory begins with *bivariate analysis* at student- and school-level to understand the relationship between each variable and nonparticipation rate. The outcome variables of analysis are: 1) assessed, 2) absenteeism, and 3) refusal that includes student and parental refusal on behalf of their children.

At student level, the analysis of participation outcome is performed by proxy student-level measures of social isolation as proposed in literature (e.g., Groves and Couper, 1998), including student's race/ethnicity, eligibility for national school lunch program (NSLP), and student achievement. At school level, participation outcomes are analyzed by school variables including urbanicity of school location, school type (public vs. private), school enrollment size, and by school climate measures such as problem with gang activities, teacher absenteeism, and parental support of student achievement. Control variables are also included for analysis such as gender and census region of school location.

Multiple Logistic Analyses

Next I conduct logistic analysis with proxy measures of social isolation in NAEP at each of two levels: student-level, and school level including social psychological school climate variables. The dependent variable is participation versus nonparticipation that combines absence and refusal. Multivariate analysis is conducted to evaluate the effects of key variables while holding constant for other variables. For example, the effect of race/ethnicity on participation outcome in NAEP can be measured while holding constant for a set of other student- and school-level variables, thus increasing the explanatory power of a key variable in the model.

At student level, I have two models to test. In the initial model of the current NCES practice (Basic Model), I examine the impact of race/ethnicity, student gender, student eligibility for national school lunch program (i.e., a proxy measure of socio-economic status), school type (public vs. private), and Census classification of school location. I expand this model by adding student achievement variables such as Carnegie credits, GPA, taking advanced mathematics, and taking advanced science (Expanded Model) to evaluate the impact of student achievement variables on nonparticipation propensity beyond what an initial set of NAEP frame variables accounts for.

Finally in the fully expanded model (Full Model), I conduct analysis of multivariate logistic models that combine student- and school-level variables of social isolation and observe changes in statistical significance and size of logistic regression coefficients from one model to another. Acknowledging the limitations of measures of social isolation in explaining nonparticipation at each level of analysis, I combine in logistic regression model both student-level proxy measures of social isolation and school-level indicators in order to assess the extent to which major school variables explain away the impacts of student-level proxy indicators of social isolation and of student achievement variables on nonparticipation. In the final full model I add social psychological school climate measures of social isolation (e.g., school problems with gang activities, teacher absenteeism, and parental support of student achievement). It is a model that also includes proxy measures of school-level social isolation such as urbanicity of school location, and school enrollment size. This final full model is intended to assess the extent to which key school climate variables explain away the joint impacts on nonparticipation of student- and school-level indicators of social isolation. It is also a model intended to evaluate the effect on NAEP nonparticipation of student achievement variables (e.g., Carnegie credit and GPA), which are associated with NAEP performance measures, while controlling for a set of other student and school proxy measures of social isolation.

I develop the three models as described above initially with all 12^{th} graders who are sampled to participate in NAEP Mathematics or Science and linked to High School Transcript Study. I replicate the logistic analysis in turn to estimate the effects on NAEP nonparticipation of the same set of student- and school-level variables in NAEP Mathematics and Science, respectively. The eventual goal of logistic regression analysis is to create response propensity scores specific to each NAEP subject, develop alternative nonresponse weights also specific to each subject, and apply them to re-estimate NAEP scale scores in Mathematics and Science. Thus it is required to develop subject-specific logistic regression models. As Little and Vartivarian (2003) suggested, I have used unweighted rather than weighted logistic regression models as the basis for nonresponse weight adjustment. For consistency throughout this research, I have retained the unweighted coefficients in all logistic regression models.

Multivariate Analysis Accounting for Sources of Nonresponse

Using the final full model that would most comprehensively incorporate key explanatory variables at both student- and school-level including school climate variables, I estimate two related logistic regression models – a contact model and a cooperation model conditional on contact. The dependent variable in the contact model indicates that the sampled 12th grader was contacted in the pre-assessment phase and in turn explicitly refused or participated in the assessment phase, or was not contacted (absent). The dependent variable in the cooperation model indicates that the contacted student participated in or refused NAEP assessment, thus excluding non-contacted absent students from analysis. By comparing the extent to which a set of student- and school-level variables affects the outcome variable of contact or cooperation, I explore the underlying mechanism of how social isolation variables of my choice would be robust enough to explain nonparticipation phenomena in NAEP.

I develop both contact and cooperation models as described above initially with all 12th graders who are sampled to participate in NAEP Mathematics or Science and linked to High School Transcript Study. I replicate the logistic analysis in turn to estimate the effects on NAEP contact and cooperation of the same set of student- and school-level variables in NAEP Mathematics and Science, respectively. My analysis is designed such that it is possible to evaluate how the same set of student- and school-level measures of social isolation affects 12th graders' participation in NAEP globally and then their participation in subject-specific NAEP.

Marginal Effects on NAEP Participation

For all logistic regression models, I provide coefficients, standard errors, and odds ratio. Coefficients are useful to compare the effects on NAEP participation outcome of a set of explanatory variables at student- and school-level. A coefficient that is statistically significant and positive (negative) indicates that having the characteristics in question raises (lowers) the probability of the modeled NAEP participation outcome; however, it is difficult to interpret the size of the effect of each explanatory variable. To help interpret the impacts of explanatory variables on the modeled outcome, I add odds ratio to tables. An odds ratio that is statistically significant and greater than 1 indicates that the odds of the outcome variable (e.g., being assessed, contact, cooperation) increase multiplicatively by exponentialed coefficient for a target group as they are for a reference group; an odds ratio of less than 1 indicates that the odds of the NAEP outcome variable decrease by a factor of 1 minus the exponentialed coefficient estimate for a target group as they are for a reference group. For example, in a study of investigating heart attack (dependent variable), let's suppose gender, race/ethnicity, age, and history of family illness with heart attack are key predictors among others. The odds ratio of 1.07 for male, for example, would suggest that males are about 7 percent more likely to suffer from heart attack as compared to the reference group of females, while controlling for all other factors in a model.

To assist further in interpreting the logistic regression results, I calculate the implied change in the probability of the NAEP participation outcome associated with having versus not having each characteristic as referenced by each explanatory variable, evaluated at the average probability of observing the outcome for the sample as a whole. These marginal probability estimates are easy to understand as compared to logistic regression coefficients or odds ratio. The statistical significance of the marginal effects is determined based on the magnitude and standard errors of the corresponding logistic regression coefficients.

For actual calculation, I use Excel spread sheets such that each logistic regression coefficient associated with each given characteristic is converted to probability for student with the given characteristic. When this probability is subtracted from the average probability of observing the outcome for the NAEP sample as a whole, the result is change in probability associated with having each explanatory characteristic relative to student with average probability of the modeled NAEP participation outcome. I compute marginal effects on NAEP participation outcome such as being assessed, contact, and cooperation conditional contact, initially with the HSTS-linked NAEP sample for both Mathematics and Science. I repeat calculating marginal effects of each explanatory variable at student- and school-level for Mathematics and Science, respectively, by using sets of original logistic regression coefficients specific to each subject.

4.2. Comparison of Alternative Nonresponse Weighting Adjustments in NAEP

The next analysis is to assess whether and the extent to which reweighting the NAEP estimates of educational performance accounting for differences in participation propensities as modeled in this research makes any substantive change to key estimates of NAEP (e.g., scale scores overall and by key background variables, achievement gap by key variables such as gender and race/ethnicity). NAEP scale score results provide a numeric summary of what students know and can do in a particular subject and are presented for student groups such as gender, race/ethnicity, and school location by census region. Achievement gap describes student achievement in terms of gap, for example, between black and white students, between Hispanic and white students, and between male and female students. Evaluating achievement gap by key background variables is essence of the No Child Left Behind mandates. These reporting metrics by scale score and achievement gap greatly facilitate performance comparisons within a subject from one group of students to another in the same grade.

I use the participation propensities generated from the subject-specific final logistic regression model (Full Model) to calculate nonparticipation adjustment factors equal to the inverse of the estimated response propensity for each participating 12th grader in NAEP. Using the propensity-score-based weight adjustment, I recalculate NAEP estimates of scale score in the 2000 NAEP Mathematics and Science respectively, and compare recalculated NAEP estimates with estimates produced using the official NAEP estimation weights in the 2000 NAEP Mathematics and Science. For specific steps of nonresponse bias analysis, I apply the approach by Abraham, Maitland, and Bianchi (2006), who assess nonresponse bias by developing a theory-based propensity model that allowed them to better account for nonresponse in estimating time use in the American Time Use Survey. With due attention to the potential association between NAEP variables of educational performance and nonparticipation propensity, I use the final logistic model (Full model) that estimates logistic regression coefficients for each major predictor in NAEP. The dependent variable is participation versus nonparticipation that combines absence and refusal. Standard errors for the estimates from the regressions are estimated using the stratified Jacknife replication variance method (Krewski and Rao 1981), assuming two PSUs per stratum, which account for a complex sample design with multiple stages of sampling, unequal selection probabilities, and complex weighting procedures. Replicate weights are provided by NCES (Roey, S., et al., 2005).

The official HSTS-linked NAEP estimates reported by NCES are calculated using a set of eight weights that incorporate school and student nonresponse: NAEP-linked student base weight, school trimming adjustment factor, school nonrespsnse adjustment factor, school substitution adjustment factor, year-round school adjustment factor, student nonresponse adjustment factor, student trimming adjustment factor, and poststratification adjustment factor. Below are brief descriptions of each component of weights.

- **NAEP-linked student base weight** reflects a student's overall probability of being selected for the HSTS 2000.
- **School trimming adjustment factor** is a weighting adjustment procedure that involves detecting and reducing extremely large school weights. Unusually large weights can seriously inflate the variance of survey estimates such as weighted means.
- **School nonrespsnse adjustment factor** inflates the weights of schools that participated in the HSTS 2000 to account for eligible schools that did not participate.

School nonresponse leads to the loss of sample data that must be compensated for in the weights. Similar to the school trimming procedure, the purpose of the nonresponse adjustment procedure is to reduce the mean square error of survey estimates.

- **School substitution adjustment factor** adjusts for the difference in grade enrollment prior to sampling between the participating substitute school and its corresponding original school that it replaced.
- **Year-round school adjustment factor** applies only to students in year-round schools, where only a portion of the total student body was in school at any given point in time. The year-round adjustment factor inflated the weight to account for students who were on break at the time of student sampling.
- **Student nonresponse adjustment factor** inflates the weights of "responding" students to account for "nonresponding" eligible students.
- **Student trimming adjustment factor** is done to detect and trim extremely large weights at the student level. Large student weights generally resulted from compounding nonresponse adjustments at the school and student levels coupled with low to moderate probabilities of selection at the various stages of sampling. As with school trimming weights, the purpose of the trimming student weights was to reduce the effect of unusually large weights on survey estimates. Trimming may introduce a small bias but is designed to reduce the mean square error of sample estimates.
- **Post-stratification adjustment factor** is a weighting procedure that adjusts the weights of sample cases so that the weighted sample distribution is the same as some known population distribution, the Current Population Survey in case of the HSTS study.

My analysis includes all weighting components except the last two to make analysis comparable and less susceptible to errors. Student trimming adjustment, which is designed to reduce mean square error of sample estimates, in fact introduces a bias according to the HSTS Technical Report. In HSTS-linked NAEP, student trimming affects weights of just a few samples whose effect on mean square error is not necessarily positive. Poststratification adjustment is a procedure to adjust the weights of sample cases such that the weighted sample distribution is the same as some known population distribution. The control total is based on the Current Population Survey data. NCES has stopped using this adjustment as it is not possible to derive reliable counts of 12^{th} graders from the CPS data. Thus weights in my research are based on the remaining six weighting components. My weights substitute student nonresponse adjustment factor with the response propensity based weight that is derived from multivariate logistic regression models guided by constructs and proxy measures of social isolation theory.

I perform analysis with WesVar to properly account for the complex multi-stage clustered NAEP sample design, which cannot be handled by standard statistical packages such as SAS or SPSS. Chart A-4 in Appendix illustrates WesVar steps I have taken to re-estimate NAEP scale scores for the 2000 mathematics and science, respectively, with alternative nonresponse adjustment weights I develop based on response propensity. As described in Chapter on research methodology, WesVar allows us to generate NAEP scale scores based on Item Response Theory and scale score distributions are estimated for a group of students by key background variables such as gender and race/ethnicity.

5. Results

My analysis begins with bivariate analysis to understand the extent to which each social isolation variable is associated with nonparticipation in NAEP. I then explore the extent to which a set of variables of social isolation is likely to affect participation in NAEP in order to identify a model that is robust enough to predict participation outcomes in NAEP. Lastly I evaluate the impact on NAEP estimates of alternative nonresponse adjustment weighting that is developed from the final nonparticipation propensity model I find to be most fitting to the NAEP data.

5.1. Bivariate Analyses

Bivariate analyses are the first step to testing participation hypotheses based on social isolation theory. At student level, the analysis of participation outcome is performed by proxy student-level measures of social isolation as follows: student's race/ethnicity, eligibility for national school lunch program (NSLP), and student achievement variables as measured by experience of taking advanced mathematics or science courses, the number of Carnegie credits, and GPA.   At school level, participation outcomes are analyzed by school-level proxy measures of social isolation as follows: urbanicity of school location, school type (public vs. private), school size measured by school enrollment, and school climate measures that include problem with gang activities, teacher absenteeism, and parental support of student achievement. Control variables are also included for analysis such as gender, census region of school location, and NAEP assessment students complete.

In Table 2, I present participation outcome rates for the 2000 NAEP at grade 12. The table shows rates of being assessed, absence, and refusal by student and school variables. The table indicates that the overall rates of participation (i.e., being assessed) offer support confirming the social isolation hypothesis. Race/ethnicity is found to be a factor for determining the rate of participation in NAEP; black 12th graders are, overall, less likely to participate in NAEP. The most notable pattern is the higher student performance of 12th graders as measured by GPA and Carnegie credits, the more likely their participation in NAEP.

It is notable that students attending urban schools are about 19 percentage points less likely to participate in NAEP than students attending rural schools. Twelfth graders attending large (i.e., school enrollment > 900) public schools located in urban areas are less likely to participate in NAEP. Students attending private schools are about 20 percentage points more likely to participate than students attending public schools. Here are consistent patterns of participation difference evaluated by measures of school culture. Students attending schools suffering from problems with gang activities and teacher absenteeism are less likely to participate in NAEP.

Table 2. Weighted Proportion of student participation status in the 2000 NAEP grade 12 mathematics and science samples linked to the 2000 high school transcript study, by student and school variables

		Assessed	Refusal	Absent
	Overall	75.1%	9.6%	15.3%
Student Correlates				
Race/ethnicity	White	76.0%	9.5%	14.6%
	Black	69.8%	12.4%	17.8%
	Hispanic	75.9%	7.2%	16.9%
	Other race	75.9%	10.4%	13.7%
Taking Advanced Mathematics	No	76.3%	9.6%	14.1%
	Yes	81.7%	6.0%	12.2%
	No records	55.5%	13.3%	31.3%
Taking Advanced Science	No	76.5%	9.5%	13.9%
	Yes	79.6%	6.9%	13.5%
	No records	55.5%	13.3%	31.3%
Carnegie Credits	<24	68.3%	14.9%	16.8%
	24-28	78.6%	7.8%	13.6%
	> 28	81.9%	7.0%	11.1%
	No records	55.5%	13.4%	31.1%
GPA	<=2.00	60.1%	12.9%	27.0%
	2.01 - 3.00	74.7%	9.6%	15.7%
	3.01-4.00	79.6%	8.7%	11.7%
	Others	86.7%	7.2%	6.1%
Eligibility for National School Lunch Program	Ineligible	75.8%	10.1%	14.1%
	Eligible	76.5%	8.8%	14.7%
	Unknown	72.7%	8.9%	18.4%
School Correlates				
School Location	Urban	67.4%	13.6%	19.0%
	Suburban	74.1%	9.7%	16.3%
	Rural	86.1%	5.0%	8.9%
School Type	Public	73.4%	10.4%	16.2%
	Private	93.8%	.8%	5.4%
School Enrollment	< 500	88.5%	3.3%	8.3%
	500 - 900	89.0%	4.4%	6.5%
	> 900	70.7%	11.8%	17.5%
	No records	61.1%	13.4%	25.6%

Continued - Table 2. Weighted Proportion of student participation status in the 2000 NAEP grade 12 mathematics and science samples linked to the 2000 high school transcript study, by student and school variables

		Assessed	Refusal	Absent
Social psychological school climate variables				
Problem with gang activities	Serious or moderate	65.6%	14.5%	19.9%
	Minor or not a problem	76.9%	9.0%	14.2%
	No records	60.6%	14.1%	25.3%
Teacher absenteeism	Serious or moderate	72.0%	12.3%	15.8%
	Minor or not a problem	77.7%	8.5%	13.8%
	No records	61.2%	12.8%	26.0%
Parental support of student achievement	Very or somewhat positive	76.4%	9.0%	14.6%
	Somewhat or very negative	75.5%	10.5%	14.0%
	No records	58.8%	15.7%	25.5%
Other Control Variables				
Student gender	Male	74.5%	10.2%	15.3%
	Female	75.7%	9.1%	15.2%
Census Region of school	Northeast	77.5%	6.9%	15.6%
	Midwest	77.8%	8.6%	13.5%
	South	77.8%	8.4%	13.8%
	West	64.3%	15.7%	20.0%
NAEP Assessment Student Completed	Mathematics	76.6%	8.5%	15.0%
	Science	74.0%	10.5%	15.5%

NOTE: N is 20549. Except for rounding error, the numbers in each row sum to 100.0 percent. Other race/ethnicity includes Asian-Pacific American and American Indian.
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Transcript Study (HSTS), 2000; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Science Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment.

A notable exception is that the proportion at a school of students in National School Lunch Program, a proxy measure of socio-economic status, is found not to be associated with NAEP participation by 12th graders. Another notable exception is that the participation rate of Hispanic and Asian/Pacific students is comparable to that of White students. Gender does not appear to be associated with NAEP participation. Twelfth graders attending schools in the West region of census are least likely to participate in NAEP.

When one turns to the two important sequential components of nonparticipation in NAEP, he finds that contribution of student absence (i.e., noncontact) to nonparticipation is serious by a factor of about 3 to 2, compared to refusal (i.e., 15.3% vs. 9.6%). As explained in a previous section, 12th graders' absence in NAEP assessment happens for various reasons including temporary (less than two weeks) or long-term depending on the nature of illness; and refusal occurs when 12th grader refuse to participate in the assessment before being given a NAEP assessment booklet or when parents refuse on behalf their children by notifying school of their unwillingness to allow their children to participate in NAEP. Table 2 indicates that Black 12th graders are more likely to be absent and to refuse as well, compared to other race/ethnicity groups. The refusal rate by black students is about twice as high as that by Hispanic groups. Overall rates of absence and refusal provide evidence to support social isolation hypotheses. Students attending large public schools in urban area are more likely to be absent. It is notable that students attending urban schools are about twice as likely to be absent yet three times as likely to refuse, compared to students attending rural schools. When I turn to measures of school culture, it seems that the contribution of absence to nonparticipation stands out. The rates of absence are comparable among students when looking at school culture related to gang activities, teacher absenteeism, and parental support of student achievement. However, the rates of NAEP refusal among students attending schools troubled with gang activities, and teacher absenteeism are generally higher than those attending schools with relatively less school-wide problems.

5.2. Multivariate Logistic Analysis to Construct Response Propensity Models

An alternative nonresponse adjustment factor I develop for NAEP takes advantage of the final multivariate logistic regression model that allows us to incorporate a set of student- and school-predictors of response. Response propensity scores I use for developing alternative nonresponse adjustment are derived from the final multivariate logistic regression model I find robust and social isolation theory driven. The reader is reminded that multivariate analysis models are constructed to evaluate the effects of individual factors on participation outcomes in NAEP, while holding other factors constant.

<u>5.2.1. Multivariate Logistic Models</u>

I begin multivariate modeling to predict (non)response by using data from students sampled for NAEP mathematics and science in 2000, and continue estimating models for mathematics and science, respectively.

Table 5-2-1 summarizes the estimates from each of the logistic regression models with mathematics and science NAEP participation outcomes as dependent variable (i.e. assessed = 1; refusal/absent = 0), beginning with a basic model of the current NCES practice, an expanded model, and the full model. The overall model fits as measured by -2 log likelihood in the last row indicate that an expanded model is an improvement over the basic model and the final full model is also a significant improvement over the basic model as well as the expanded model.

A basic model including only NAEP frame variables including race/ethnicity and gender indicates that race/ethnicity is a strong predictor of NAEP participation. Compared to White students, the odds of Black students being a participant (by looking at the third column for the first model) are estimated to be about 28 percent lower, beyond and above what a set of variables in this model can account for including gender, eligibility for national school lunch program, school type, and census region of school location. For Hispanic students, the odds of NAEP participation are about 11 percent higher than for White students, again beyond and above what key predictors in this basic model can account for. Students attending private schools are extremely more likely to participate in NAEP than their counterparts attending public schools. Geographic location of schools 12th graders attend appears to be a significant predictor of participation in NAEP. Students attending schools in the West region are much less likely to participate than those attending schools in the Northeast region.

Table 5-2-1. Effects of Social Isolation indicators on Participation in NAEP Mathematics and Science (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			**Expanded Model with Student Achievement Variables**			**Full Model with School Culture and Control Variables**		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Intercept	-1.031*	.153	.357	-.728*	.157	.483	-.861*	.168	.423
Female	.019	.033	1.019	-.059	.034	.942	-.053	.034	.949
Race/ethnicity (ref=white)									
Black	-.326*	.045	.722	-.195*	.047	.823	.043	.050	1.044
Hispanic	.101*	.047	1.106	.236*	.049	1.266	.439*	.050	1.552
Others	.308*	.068	1.361	.333*	.069	1.396	.500*	.071	1.648
National school lunch program (ref=Ineligible)									
Eligible for school lunch	.077	.046	1.080	.145*	.047	1.156	.172*	.050	1.188
Unknown	-.394*	.041	.674	-.329*	.042	.720	-.309*	.044	.734
Private school	1.959*	.133	7.093	1.843*	.134	6.314	1.734*	.140	5.665
Census region (ref = NE)									
Midwest	.131*	.059	1.140	.112	.060	1.118	-.118	.063	.889
South	.420*	.051	1.522	.339*	.052	1.404	.215*	.054	1.240
West	-.250*	.054	.779	-.224*	.055	.799	-.238*	.057	.788
Took advanced courses in Math or Science				.013	.059	1.013	.181*	.060	1.198
Carnegie credits (ref = 24-28)									
Low # CC (16-23)				-.352*	.042	.703	-.223*	.043	.800
High # CC (>=29)				.208*	.052	1.232	.149*	.053	1.160
No CC records				-.750*	.086	.472	-.590*	.088	.554
GPA (2< ref <= 3)									
Low GPA < =2.0				-.180*	.074	.835	-.242*	.076	.785
High GPA > 3.01				.141*	.041	1.152	.130*	.042	1.139
GPA not reported				.529*	.201	1.698	.320	.206	1.377

Continued - Table 5-2-1. Effects of Social Isolation indicators on Participation in NAEP Mathematics and Science (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			Expanded Model with Student Achievement Variables			Full Model with School Culture and Control Variables		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Urbanicity of school location (ref = urban)									
Suburban							.162*	.038	1.176
Rural							.824*	.062	2.280
School enrollment (ref = large enrollment > 900)[1]									
Enrollment < = 500							.438*	.083	1.549
Enrollment (501-900)							.868*	.071	2.382
More problem with gang activities[1]							.067	.064	1.069
More problem with teacher absenteeism[1]							-.274*	.043	.760
Less parental support of student achievement[1]							-.178*	.057	.837
School-level information incomplete[2]							-.648*	.055	.523
Negative 2 Log Likelihood	22691.606			22187.609			21363.25		

Note: N is 20549. * significant at p < .05. 1. Samples with item missing data are added to the reference group of the majority. 2. School-level information incomplete is a dichotomous variable that attempts to capture pattern of item missing in the following four variables: school enrollment, problem with gang activities, problem with teacher absenteeism, and parental support of student achievement. The reference group (0) is where all four variables take valid data; the other group (1) is where any of four variables is missing.

In the expanded model, I add student-level proxy measures of social isolation such as student achievement variables (Carnegie credits, GPA, and taking advanced courses in mathematics or science), which I suspect be the covariates of both NAEP achievement estimates and participation propensity.

The most notable finding in the expanded model is that the higher student academic performance (as measured by the number of Carnegie credits and GPA), the more likely the student is to participate in NAEP assessment. Compared to students who earned Carnegie credits of 24 -28, 12th graders who have taken at least 29 Carnegie credits are about 23 percent more likely to participate in NAEP, and students who have taken less than 24 Carnegie credits are about 30 percent less likely to participate. GPA showed a comparable power of predictability. This effect on NAEP participation of academic measures of Carnegie and GPA holds even when such student characteristics as gender and race/ethnicity are controlled for. Race/ethnicity sustained its power of predicting NAEP participation yet at a level about 10 percentage points lower than it was in the basic model. When compared to students ineligible for NSLP, students eligible for national school lunch program are more likely to participate, and students whose eligibility is unknown are less likely to participate. The status of taking advanced courses in mathematics or science is not a useful predictor of participation outcome.

I have taken one more critical step of introducing into my model additional school-level proxy measures of social isolation, such as urbanicity of school location, enrollment size, and school culture as measured by social psychological perception of problem with gang activities, teacher absenteeism, parental support of student achievement. Census classification of school region is also added as a control variable.

This full model also includes school-level information incompleteness, a dichotomous variable that attempts to capture pattern of item missing in the following four variables: school enrollment, problem with gang activities, problem with teacher absenteeism, and parental support of student achievement. The reference group (0) is where all four variables take valid data; the other group (1) is where one or more variables take missing values. This additional variable is created to treat concerns with multicollinearity I have observed among the four variables and to keep all eligible cases in the study without resorting to list-wise deletion of cases with missing values.

Interestingly, estimates of the full model in Table 5-2-1 demonstrate that academic measures of Carnegie credits and GPA both sustain their power of predicting participation status beyond what a set of proxy measures of student and school-level social isolation accounts for participation outcome. School culture measures of teacher absenteeism and parental support both are likely predictors of NAEP participation. A social psychological measure of problem with gang activities at school is not found to be a useful predictor of participation outcome. When one turns to look into the effect of student-level variable, one finds that Hispanic ethnicity sustains its power of predicting NAEP participation outcome; being a Black student is no longer a factor for explaining the outcome. Eligibility for school lunch program continue its power of predicting participation outcome, although it is in the opposite direction of my hypothesis that students eligible for NSLP are less likely to participate in NAEP. In contrast, among 12th graders whose eligibility for NSLP is unknown, they are found to be less likely to participate in NAEP when compared to those ineligible for NSLP. Considering these mixed findings, one cannot properly infer how the NSLP variable affects participation outcome. Further difficulty of making inferences from these mixed findings is due to serious concerns over the years about the reliability of NSLP as a proxy measure of SES. Nevertheless, I keep this student-level variable as it is the only proxy measure of SES which is available despite its measurement problem.

When one examines the effect of school-level variables in the full model, urbanicity of school location is likely to account for participation outcome such that compared to students attending urban schools, students attending suburban schools are more likely to participate in NAEP, and students attending rural schools are much more likely to participate. It appears that the effect on participation of school size (as measured by 12th grade enrollment) is more complex, perhaps curvilinear. Students attending small schools (less than 500 enrollment) were about 50 percent more likely to participate compared to students attending large schools (i.e., enrollment of greater than 900). The coefficient for medium schools (i.e., enrollment of 501-900) indicates students attending moderate size of schools are much more likely to participate than students attending large schools.

It is notable that the indicator of school-level information incompleteness sustains its power of predicting participation outcome. The reader is reminded that it is a dichotomous variable that captures pattern of item missing in the following four variables: school enrollment, problem with gang activities, problem with teacher absenteeism, and parental support of student achievement. The reference group (0) is where all four variables take valid data; the other group (1) is where one or more variables take missing values. This composite variable is created to treat concerns with multicollinearity I have observed among the four variables and keep all eligible cases in the study without resorting to list-wise deletion of cases with missing values. This composite variable suggests that compared to students attending schools providing all valid information for the four school variables, students attending schools providing invalid data for one or more of school variables are about 48 percent less likely to participate in NAEP.

Census region of school location is found to be a significant control variable. Compared to students attending schools in the Northeast as classified by the national census, students attending schools in the West are about 20 percent less likely to participate. It reminds us of the similar pattern in bivariate analysis. In contrast, students in the South are more likely to participate than students in the Northeast.

Tables 5-2-2 and 5-2-3 summarize estimates from the logistic regression models, this time with student participation in mathematics and science, respectively, as a dependent variable. These models continue including a basic model with the current NCES practice, an expanded model, and the full model. When evaluating the final full model for explaining participation outcome in mathematics and science, respectively, one would find that most of the social isolation variables I have used (e.g., Carnegie credits, school culture measure, school size, urbanicity of school location, race/ethnicity) have statistically significant effects in each of these multivariate full models. Interestingly, the indicator of school-level information incompleteness sustains its power of predicting subject-specific participation outcome. A notable exception in the mathematics NAEP participation model is that Carnegie credit, not GPA, sustains its power of predicting mathematics participation outcome (i.e., the greater number of Carnegie credits 12th graders earned, the more likely their participation in mathematics NAEP would be).

In the multivariate final full model for explaining Science NAEP participation, there are a couple of exceptions to address: 1) students whose eligibility for NSLP are unknown are found to be less likely to participate than those ineligible for NSLP, and students eligible for NSLP are not significantly different from students ineligible regarding their participation in science NAEP, and 2) GPA sustains their power of predicting participation outcome in science with Carnegie credits holding its predicting power more among poor performing students.

Table 5-2-2. Effects of Proxy Measures of Social Isolation on Participation in NAEP Mathematics (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			Expanded Model with Student Achievement Variables			Full Model with School Culture and Control Variables		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Intercept	-1.027*	.230	.358	-.762*	.236	.467	-.825*	.254	.438
Female	-.026	.050	.975	-.093	.051	.911	-.097	.052	.908
Race/ethnicity (ref=white)									
Black	-.339*	.069	.712	-.224*	.071	.799	.032	.076	1.032
Hispanic	.122	.073	1.130	.259*	.075	1.296	.479*	.078	1.615
Others	.305*	.105	1.357	.336*	.107	1.399	.526*	.110	1.693
National school lunch program (ref=Ineligible)									
Eligible for school lunch	.151*	.072	1.163	.232*	.073	1.261	.297*	.077	1.345
Unknown	-.371*	.063	.690	-.316*	.064	.729	-.280*	.067	.756
Private school	2.070*	.198	7.922	1.963*	.200	7.123	1.789*	.211	5.986
Census region (ref = NE)									
Midwest	.106	.091	1.112	.078	.092	1.081	-.192*	.097	.826
South	.417*	.078	1.517	.328*	.081	1.388	.195*	.084	1.215
West	-.218*	.082	.804	-.192*	.085	.825	-.217*	.088	.805
Took advanced courses in Math or Science				.121	.093	1.129	.291*	.096	1.338
Carnegie credits (ref = 24-28)									
Low # CC (16-23)				-.308*	.064	.735	-.177*	.066	.838
High # CC (>=29)				.272*	.080	1.313	.176*	.083	1.192
No CC records				-.906*	.131	.404	-.739*	.134	.478
GPA (2< ref <= 3)									
Low GPA < =2.0				-.083	.113	.920	-.164	.115	.849
High GPA > 3.01				.122	.064	1.129	.117	.065	1.124
GPA not reported				.391	.310	1.479	.209	.317	1.232

Continued - Table 5-2-2. Effects of Proxy Measures of Social Isolation on Participation in NAEP Mathematics (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			Expanded Model with Student Achievement Variables			Full Model with School Culture and Control Variables		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Urbanicity of school location (ref = urban)									
Suburban							.260*	.059	1.297
Rural							.834*	.096	2.302
School enrollment (ref = large enrollment > 900)[1]									
Enrollment < = 500							.446*	.123	1.562
Enrollment (501-900)							1.045*	.117	2.844
More problem with gang activities[1]							.039	.097	1.040
More problem with teacher absenteeism[1]							-.374*	.065	.688
Less parental support of student achievement[1]							-.244*	.087	.783
School-level information incomplete[2]							-.711*	.083	.491
Negative 2 Log Likelihood	9806.326			9562.652			9152.814		

Note: N is 9,163. * significant at p < .05. 1. Samples with item missing data are added to the reference group of the majority. 2. School-level information incomplete is a dichotomous variable that attempts to capture pattern of item missing in the following four variables: school enrollment, problem with gang activities, problem with teacher absenteeism, and parental support of student achievement. The reference group (0) is where all four variables take valid data; the other group (1) is where any of four variables is missing.

Table 5-2-3. Effects of Proxy Measures of Social Isolation on Participation in NAEP Science (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			Expanded Model with Student Achievement Variables			Full Model with School Culture and Control Variables		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Intercept	-.998*	.205	.369	-.666*	.210	.514	-.855*	.223	.425
Female	.053	.043	1.054	-.036	.044	.965	-.025	.045	.975
Race/ethnicity (ref=white)									
Black	-.316*	.060	.729	-.174*	.062	.840	.051	.066	1.052
Hispanic	.084	.062	1.088	.216*	.064	1.242	.410*	.066	1.507
Others	.313*	.089	1.367	.337*	.091	1.401	.491*	.093	1.633
National school lunch program (ref=Ineligible)									
Eligible for school lunch	.023	.061	1.023	.087	.062	1.090	.086	.065	1.090
Unknown	-.414*	.054	.661	-.341*	.055	.711	-.332*	.058	.718
Private school	1.842*	.179	6.310	1.718*	.180	5.574	1.668*	.188	5.303
Census region (ref = NE)									
Midwest	.148	.079	1.159	.136	.080	1.146	-.067	.083	.935
South	.420*	.067	1.522	.344*	.068	1.411	.223*	.071	1.250
West	-.278*	.071	.757	-.253*	.073	.777	-.263*	.076	.769
Took advanced courses in Math or Science				-.058	.076	.943	.108	.078	1.115
Carnegie credits (ref = 24-28)									
Low # CC (16-23)				-.387*	.056	.679	-.258*	.057	.773
High # CC (>=29)				.164*	.067	1.179	.129	.069	1.138
No CC records				-.626*	.115	.535	-.467*	.118	.627
GPA (2< ref <= 3)									
Low GPA < =2.0				-.258*	.099	.773	-.307*	.101	.735
High GPA > 3.01				.158*	.055	1.171	.141*	.056	1.152
GPA not reported				.635*	.265	1.886	.410	.271	1.506

Continued - Table 5-2-3. Effects of Proxy Measures of Social Isolation on Participation in NAEP Science (assessed = 1; not assessed =0) - 2000 NAEP and High School Transcripts Study, Grade 12									
	Basic Model with Current Practice			Expanded Model with Student Achievement Variables			Full Model with School Culture and Control Variables		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Urbanicity of school location (ref = urban)									
Suburban							.090	.051	1.094
Rural							.833*	.083	2.299
School enrollment (ref = large enrollment > 900)[1]									
Enrollment < = 500							.410*	.112	1.506
Enrollment (501-900)							.755*	.091	2.127
More problem with gang activities[1]							.092	.084	1.097
More problem with teacher absenteeism[1]							-.198*	.056	.820
Less parental support of student achievement[1]							-.137	.077	.872
School-level information incomplete[2]							-.607*	.073	.545
Negative 2 Log Likelihood	12863.35			12593.83			12158.36		

Note: N is 11,386. * significant at p < .05. 1. Samples with item missing data are added to the reference group of the majority. 2. School-level information incomplete is a dichotomous variable that attempts to capture pattern of item missing in the following four variables: school enrollment, problem with gang activities, problem with teacher absenteeism, and parental support of student achievement. The reference group (0) is where all four variables take valid data; the other group (1) is where any of four variables is missing.

5.2.2. Multivariate Analysis of Components of Response: Contactability and Cooperation Rate

Tables 5-2-4a and 5-2-4b summarize two related models – a contact model and a cooperation model conditional on contact – using the final full model that incorporates key explanatory variables at both student- and school-level, including academic measures and school climate variables. The dependent variable in the contact model indicates that the sampled 12th grader was contacted in the pre-assessment phase and in turn explicitly refused or participated in the assessment phase, or was not contacted (absent). The dependent variable in the cooperation model indicates that the contacted student participated in or refused NAEP assessment, thus excluding non-contacted absent students from analysis. By comparing the extent to which a set of student- and school-level variables affects the outcome variable of contact or cooperation, I explore the underlying mechanism of how social isolation variables of my choice would be robust enough to explain participation behavior of 12th graders in NAEP.

Table 5-2-4a Full Model - Effects of Proxy Measures of Social Isolation on Contact and Cooperation in NAEP Mathematics and Science, Grade 12						
	Contact Model Contact (assessed and refusal) = 1			Cooperation, conditional on contact Model Assessed = 1		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Intercept	.256	.185	1.292	-1.061*	.410	.346
Female	-.111*	.040	.895	.035	.050	1.036
Race/ethnicity (ref=white)						
Black	.127*	.058	1.135	-.075	.071	.928
Hispanic	.165*	.057	1.179	.737*	.078	2.089
Others	.345*	.084	1.412	.615*	.104	1.850
National school lunch program (ref=Ineligible)						
Eligible for school lunch	.057	.058	1.059	.301*	.074	1.351
Unknown	-.480*	.050	.619	-.026	.065	.974
Private school	1.369*	.150	3.930	2.755*	.388	15.725
Census region (ref = NE)						
Midwest	-.197*	.072	.822	-.023	.097	.977
South	.144*	.063	1.154	.298*	.080	1.348
West	-.009	.066	.992	-.461*	.084	.631
Took advanced courses in Math or Science	.012	.071	1.012	.360*	.092	1.434
Carnegie credits (ref = 24-28)						
Low # CC (16-23)	-.040	.051	.960	-.413*	.062	.662
High # CC (>=29)	.130*	.064	1.139	.142	.081	1.153
No CC records	-.442*	.099	.643	-.617*	.130	.540
GPA (2< ref <= 3)						
Low GPA < =2.0	-.281*	.086	.755	-.179	.112	.836
High GPA > 3.01	.295*	.051	1.344	-.097	.062	.907
GPA not reported	.852*	.285	2.344	-.342	.271	.710

Continued - Table 5-2-4a Full Model - Effects of Proxy Measures of Social Isolation on Contact and Cooperation in NAEP Mathematics and Science, Grade 12

	Contact Model Contact (assessed and refusal) = 1			Cooperation, conditional on contact Model Assessed = 1		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Urbanicity of school location (ref = urban)						
Suburban	.080	.045	1.083	.246*	.055	1.279
Rural	.605*	.074	1.831	1.017*	.099	2.764
School enrollment (ref = large enrollment > 900)						
Enrollment < = 500	.293*	.094	1.340	.631*	.148	1.880
Enrollment (501-900)	.906*	.088	2.474	.707*	.108	2.028
More problem with gang activities	.099	.076	1.104	.042	.088	1.043
More problem with teacher absenteeism	-.159*	.050	.853	-.369*	.061	.691
Less parental support of student achievement	-.012	.069	.988	-.348*	.079	.706
School-level information incomplete	-.578*	.060	.561	-.542*	.081	.582
Negative 2 Log Likelihood	16939.4			11199.86		

Note: N is 20,549 for the contact model and 17,200 for the cooperation model.
* significant at p < .05.

In Table 5-2-4a, I summarize effects of social isolation variables on contact and cooperation conditional on contact, respectively. This table, where participation in mathematics or science is combined, is quite revealing in showing the consistent effect on both contact and cooperation of social isolation variables which include the following: race/ethnicity, academic indicators as measured by Carnegie credit, measures of school culture, school urbanicity and size, and incompleteness of school-level information.

Hispanics, not Blacks, are more likely to be contacted and cooperating than White students. Students attending private schools have higher contact rates and much higher cooperation rates than those attending public schools. Students attending rural schools have higher contact and cooperation rates than those attending urban schools. Students attending schools more troubled with teacher absenteeism have lower contact and cooperation rates than those less troubled with teacher absenteeism.

The two sequential outcomes of participation suggest that GPA is generally more useful in predicting contactability than cooperation rate, whereas Carnegie credit is more helpful for predicting contact rate among high performing students and for predicting cooperation rate among low performing students. School culture measures including perception of parental support of student academic achievement, and teacher absenteeism are all fine predictors of cooperation rates, whereas only teacher absenteeism is a fine predictor of contactability in NAEP. The indicator of school-level information incompleteness continues its power of predicting both contactability and cooperation rate among 12th graders sampled for NAEP.

Table 5-2-4b Full Model - Effects of Proxy Measures of Social Isolation on Contact, and Cooperation in NAEP Mathematics and Science (assessed = 1; refusal/absent =0) - 2000 NAEP and High School Transcripts Study, Grade 12

	Mathematics						**Science**					
	Contact Model			Cooperation, conditional on contact Model			Contact Model			Cooperation, conditional on contact Model		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Intercept	.263	.275	1.301	-2.080*	1.033	.125	.273	.251	1.314	-.686	.457	.503
Female	-.131*	.061	.877	-.029	.079	.971	-.100	.053	.905	.075	.065	1.078
Race/ethnicity (ref=white)												
Black	.192*	.089	1.212	-.186	.111	.830	.074	.077	1.077	.003	.093	1.003
Hispanic	.260*	.089	1.297	.710*	.123	2.034	.090	.075	1.094	.758*	.102	2.135
Others	.323*	.128	1.381	.710*	.168	2.033	.376*	.112	1.456	.557*	.132	1.746
National school lunch program (ref=Ineligible)												
Eligible for school lunch	.217*	.091	1.243	.347*	.116	1.415	-.054	.076	.947	.267*	.096	1.306
Unknown	-.453*	.076	.636	.006	.102	1.006	-.504*	.066	.604	-.045	.085	.956
Private school	1.346*	.218	3.843	3.996*	1.011	54.403	1.373*	.207	3.947	2.232*	.423	9.314
Census region (ref = NE)												
Midwest	-.305*	.108	.737	-.007	.158	.993	-.113	.096	.893	-.031	.124	.969
South	.124	.096	1.132	.292*	.128	1.339	.150	.083	1.162	.305*	.103	1.357
West	.027	.101	1.028	-.484*	.132	.616	-.040	.088	.960	-.457*	.108	.633
Took advanced courses in Math or Science	.139	.113	1.149	.436*	.149	1.546	-.072	.092	.931	.305*	.117	1.357
Carnegie credits (ref = 24-28)												
Low # CC (16-23)	-.029	.078	.972	-.357*	.096	.700	-.049	.068	.952	-.452*	.080	.637
High # CC (>=29)	.138	.098	1.148	.192	.131	1.211	.127	.084	1.136	.113	.103	1.120
No CC records	-.578*	.149	.561	-.771*	.207	.463	-.322*	.133	.725	-.515*	.169	.597
GPA (2< ref <= 3)												
Low GPA < =2.0	-.226	.130	.798	-.069	.179	.934	-.332*	.115	.717	-.257	.144	.774
High GPA > 3.01	.319*	.078	1.376	-.174	.097	.840	.278*	.067	1.320	-.047	.080	.955
GPA not reported	.440	.385	1.553	-.196	.484	.822	1.237*	.429	3.444	-.396	.331	.673

Continued - Table 5-2-4b Full Model - Effects of Proxy Measures of Social Isolation on Contact, and Cooperation in NAEP Mathematics and Science (assessed = 1; refusal/absent =0) - 2000 NAEP and High School Transcripts Study, Grade 12

	Mathematics						Science					
	Contact Model			Cooperation, conditional on contact Model			Contact Model			Cooperation, conditional on contact Model		
	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)	B	S.E.	Exp(B)
Urbanicity of school location (ref = urban)												
Suburban	.170*	.068	1.186	.335*	.087	1.397	.011	.059	1.011	.185*	.072	1.203
Rural	.610*	.111	1.840	1.087*	.160	2.966	.619*	.099	1.857	.977*	.126	2.657
School enrollment (ref = large enrollment > 900)												
Enrollment < = 500	.473*	.143	1.605	.241	.207	1.273	.131	.126	1.140	.951*	.215	2.589
Enrollment (501-900)	.876*	.133	2.400	1.288*	.219	3.627	.928*	.119	2.528	.462*	.126	1.588
More problem with gang activities	.154	.117	1.167	-.059	.134	.943	.065	.099	1.067	.121	.118	1.129
More problem with teacher absenteeism	-.222*	.076	.801	-.509*	.095	.601	-.108	.066	.897	-.272*	.080	.762
Less parental support of student achievement	-.146	.102	.864	-.302*	.125	.740	.094	.096	1.099	-.391*	.103	.676
School-level information incomplete	-.632*	.091	.532	-.599*	.125	.549	-.543*	.081	.581	-.500*	.107	.607
Negative 2 Log Likelihood	7351.69			4588.944			9552.459			6548.946		

Note: N is 9,163 for mathematics and 11,386 for science. * significant at p < .05.

As was true with the full model that included participation in either subject, I find in Table 5-2-4b that most “social isolation” variables tend to have statistically significant effects in predicting contact or cooperation conditional on contact in mathematics and science, respectively. In mathematics NAEP, the significant predictors of both contactability and cooperation on contact include the following: student achievement measured by Carnegie credits, school culture measure, race and ethnicity, eligibility for NSLP, school type, and school urbanicity. The sequential outcomes of response suggest that GPA is somewhat more useful in predicting contactability, whereas Carnegie credit is more helpful for predicting cooperation rate in mathematics. School culture measures including perception of parental support of student academic achievement, and teacher absenteeism are all fine predictors of cooperation rates, whereas only teacher absenteeism is a good predictor of contactability in mathematics NAEP.

In science NAEP, the significant predictors of contactability and cooperation include the following: race (Hispanics, not blacks, are more likely to be contacted and cooperating), GPA (the lower the GPA, the lower cooperation rate), school type (students attending private schools have higher contact and extremely higher cooperation rates than those attending public schools), urbanicity (students attending rural schools have higher contact and cooperation rates than those attending urban schools), school size (students attending small schools have higher cooperation rates), and school-level information incomplete (students attending schools providing more incomplete information have lower contact and cooperation rates). The sequential outcomes of participation in science suggest that Carnegie credit is more helpful for predicting cooperation rate, not contact rate, in science. School culture measure of teacher absenteeism and parental support has found its only utility for predicting cooperation rate in science; none of the school culture measures is found to be helpful for predicting contactability in science.

5.2.3. Marginal Effects on NAEP Participation

In Tables 5-3a and 5-3b, I present the marginal probability effects I have generated thus far from the multivariate logistic regressions with NAEP participation outcomes as dependent variables. Changes in predicted rates associated with having versus not having the indicated characteristics are evaluated at the overall rate for the full NAEP-HSTS sample of mathematics and science, together and then individually, based on the final full logistic models of response propensity. Estimates in the 4th column are implied probability of contact and cooperation. Bold-faced estimates are significant at p < .05. For example, the figure shown in the "Low # CC (16-23)" row of the "Assessed" column in Table 5-3a indicates that, evaluated at the mean probability of participation (being assessed), having earned only 16-23 Carnegie credits lowers the probability of participation by an estimated 4.5 percentage points. This estimate in bold is statistically significant. This estimate, which is derived from the multivariate logistic regression with NAEP participation outcome as a dependent variable, is quite close to the implied probability of contact and cooperation, negative 4.83.

The most striking result to emerge from the data in Table 5-3a is that social isolation variables like academic indicators of Carnegie credit and GPA and school culture measures significantly impact participation rate (being assessed) by 2 to 5 percentage points. Interestingly, school size and type, and school-level information incomplete affect the probability to be assessed by up to 20 percentage points. Other significant variables include race/ethnicity (Hispanics have higher response rate) and school urbanicity (students attending rural schools have higher response rates). These differences tend to be more affected by differences in cooperation rates, which is the similar pattern observed among 12th graders who are more troubled with teacher absenteeism, and lack of parental support of student achievement.

Table 5-3a. Marginal Effects on NAEP Participation (Being Assessed), Contact, Cooperation conditional on contact, and Comparison to Implied Probability: 2000 HSTS-NAEP, Grade 12

Predictor	Assessed	Contact	Cooperation, Conditional on contact	Implied Probability of Contact and Cooperation
(Mean of Probability)	74.08	83.85	87.61	73.46
Female	-1.03	**-1.57**	0.38	-1.06
Race/ethnicity (ref=white)				
Black	0.81	**1.64**	-0.83	0.73
Hispanic	**7.52**	**2.11**	**6.05**	7.05
Others	**8.41**	**4.15**	**5.29**	8.29
National school lunch program (ref=Ineligible)				
Eligible for school lunch	**3.16**	0.76	**2.91**	3.13
Unknown	**-6.36**	**-7.59**	-0.28	-6.86
Private school	**20.10**	**11.48**	**11.50**	21.02
Census region (ref = NE)				
Midwest	-2.33	**-2.84**	-0.25	-2.69
South	**3.91**	**1.85**	**2.89**	4.10
West	**-4.82**	-0.12	**-5.92**	-5.06
Took advanced courses in Math or Science	**3.32**	0.17	**3.41**	3.01
Carnegie credits (ref = 24-28)				
Low # CC (16-23)	**-4.50**	-0.55	**-5.22**	-4.83
High # CC (>=29)	**2.75**	**1.69**	1.46	2.73
No CC records	**-12.77**	**-6.91**	**-8.38**	-12.50
GPA (2< ref <= 3)				
Low GPA < =2.0	**-4.91**	**-4.18**	-2.08	-5.32
High GPA > 3.01	**2.42**	**3.61**	-1.09	2.21
GPA not reported	5.66	**8.56**	-4.21	3.61

Continued - Table 5-3a. Marginal Effects on NAEP Participation (Being Assessed), Contact, Cooperation conditional on contact, and Comparison to Implied Probability: 2000 HSTS-NAEP, Grade 12

Predictor	Assessed	Contact	Cooperation, Conditional on contact	Implied Probability of Contact and Cooperation
Urbanicity of school location (ref = urban)				
Suburban	**2.98**	1.05	**2.44**	2.99
Rural	**12.61**	**6.63**	**7.52**	12.62
School enrollment (ref = large enrollment > 900)				
Enrollment < = 500	**7.50**	**3.59**	**5.39**	7.86
Enrollment (501-900)	**13.11**	**8.93**	**5.87**	13.27
More problem with gang activities	1.26	1.29	0.45	1.51
More problem with teacher absenteeism	**-5.60**	**-2.27**	**-4.59**	-5.74
Less parental support of student achievement	**-3.55**	-0.17	**-4.30**	-3.74
School-level information incomplete	**-14.15**	**-9.40**	**-7.17**	-13.58

Note: N is 20,549. Bold-faced estimates are significant at $p < .05$. Changes in predicted rates associated with having versus not having the indicated characteristics are evaluated at the overall rate for the full NAEP-HSTS sample of mathematics and science, based on the final logistic models of response propensity including both subjects. Estimates in the 4th column are implied probability of contact and cooperation.

Table 5-3b. Marginal Effects on NAEP Participation (Being Assessed), Contact, Cooperation conditional on contact, and Comparison to Implied Probability: 2000 HSTS-NAEP by Subject, Grade 12

Predictor	Mathematics: Assessed	Mathematics: Contact	Mathematics: Cooperation, Conditional on contact	Mathematics: Implied Probability of Contact and Cooperation	Science: Assessed	Science: Contact	Science: Cooperation, Conditional on contact	Science: Implied Probability of Contact and Cooperation
(Mean of Probability)	75.49	84.33	88.75	74.84	72.93	83.45	86.67	72.33
Female	-1.84	**-1.81**	-0.30	-1.85	-0.50	-1.42	0.85	-0.54
Race/ethnicity (ref=white)								
Black	0.58	**2.37**	-1.99	0.38	1.00	1.00	0.03	0.89
Hispanic	**7.77**	**3.14**	**5.38**	7.50	**7.31**	1.20	**6.61**	6.64
Others	**8.42**	**3.81**	**5.38**	8.13	**8.55**	**4.56**	**5.24**	8.56
National school lunch program (ref=Ineligible)								
Eligible for school lunch	**5.07**	**2.66**	**3.03**	5.00	1.66	-0.76	**2.80**	1.65
Unknown	**-5.55**	**-6.95**	0.06	-6.12	**-7.02**	**-8.16**	-0.53	-7.47
Private school	**19.36**	**11.06**	**11.02**	20.33	**20.53**	**11.77**	**11.71**	21.34
Census region (ref = NE)								
Midwest	**-3.72**	**-4.46**	-0.07	-4.01	-1.35	-1.62	-0.36	-1.70
South	**3.42**	1.57	**2.60**	3.63	**4.18**	1.97	**3.15**	4.40
West	**-4.24**	0.36	**-5.81**	-4.60	**-5.50**	-0.57	**-6.22**	-5.65
Took advanced courses in Math or Science	**4.98**	1.74	**3.67**	4.71	2.09	-1.01	**3.15**	1.71
Carnegie credits (ref = 24-28)								
Low # CC (16-23)	**-3.41**	-0.38	**-4.08**	-3.76	**-5.38**	-0.69	**-6.13**	-5.68
High # CC (>=29)	**3.11**	1.73	1.78	3.07	2.48	**1.68**	1.26	2.53
No CC records	**-15.96**	**-9.22**	**-10.25**	-15.88	**-10.13**	-4.93	**-7.14**	-9.89
GPA (2< ref <= 3)								
Low GPA < =2.0	-3.16	-3.22	-0.71	-3.43	**-6.48**	**-5.11**	-3.25	-6.98
High GPA > 3.01	2.09	**3.78**	-1.86	1.71	**2.70**	**3.49**	-0.55	2.55
GPA not reported	3.65	4.98	-2.11	2.54	7.30	**11.11**	-5.27	4.63

Continued - Table 5-3b. Marginal Effects on NAEP Participation (Being Assessed), Contact, Cooperation conditional on contact, and Comparison to Implied Probability: 2000 HSTS-NAEP by Subject, Grade 12

	Mathematics				Science			
Predictor	Assessed	Contact	Cooperation, Conditional on contact	Implied Probability of Contact and Cooperation	Assessed	Contact	Cooperation, Conditional on contact	Implied Probability of Contact and Cooperation
Urbanicity of school location (ref = urban)								
Suburban	**4.49**	**2.12**	**2.93**	4.42	1.74	0.16	**2.00**	1.80
Rural	**12.15**	**6.50**	**7.15**	12.27	**13.17**	**6.90**	**7.86**	13.08
School enrollment (ref = large enrollment > 900)								
Enrollment < = 500	**7.30**	**5.29**	2.19	6.66	**7.30**	1.73	**7.72**	8.08
Enrollment (501-900)	**14.26**	**8.48**	**7.87**	14.84	**12.21**	**9.28**	**4.50**	12.21
More problem with gang activities	0.72	1.93	-0.60	1.20	1.78	0.88	1.34	1.88
More problem with teacher absenteeism	**-7.55**	**-3.16**	**-6.16**	-7.80	**-4.08**	-1.55	**-3.46**	-4.19
Less parental support of student achievement	**-4.80**	-2.03	**-3.38**	-4.58	-2.79	1.26	**-5.20**	-3.32
School-level information incomplete	**-15.29**	**-10.23**	**-7.50**	-14.63	**-13.44**	**-8.90**	**-6.89**	-12.86

Note: Note: N is 9,163 for mathematics and 11,386 for science. Bold-faced estimates are significant at $p < .05$. Changes in predicted rates associated with having versus not having the indicated characteristics are evaluated at the overall rate for the full NAEP-HSTS sample of mathematics and science, respectively, based on the final logistic model of response propensity for each subject. Estimates in the 4th column under each subject are implied probability of contact and cooperation.

As I turn to each subject (See Table 5-3b), I continue finding that most "social isolation" variables tend to have statistically significant effects in predicting contact or cooperation conditional on contact in mathematics and science, respectively. In both subjects, all else being the same, Hispanic 12th graders are more likely to be contacted and cooperating. As was true for the simple tabulations in bivariate analysis, participation rates are significantly higher for students attending private schools as compared to public school students, for students attending schools in rural areas as compared to urban schools, for students attending small schools (<500 and <900) as compared to large schools (> 901), for students attending schools with less problem with teacher absenteeism and with more parental support of student achievement, and for students attending schools providing more complete school information.

In mathematics NAEP alone, Carnegie credits are positively related to participation of being assessed but not significantly related to contactability. GPA shows some positive effect on raising contact rate, but not cooperation rate, among high GPA earners. School culture measures, when they are negative, all deflate cooperation rates.

In science NAEP alone, GPA tends to be a significant predictor of contact and cooperation rates, among both low and high GPA earners. Carnegie credits seem to have less predictive power of contact and cooperation rates among better performing students (> 29). School culture measure of teacher absenteeism and parental support has found its utility in predicting cooperation rate in science; none of the school culture measures was found to be helpful in predicting contactability in science.

5.3. Effect on NAEP Estimates of Alternative Nonresponse Weighting Adjustments

I expect that alternative NAEP estimates derived from logistic regression models of the response propensity are in general likely to be lower than official estimates of mathematics and science. As presented so far, I observe that students performing better, as measured by Carnegie credits or GPA, are found to be more likely to be participating in NAEP beyond and above what a number of key correlates of participation at student and school levels can account for. These correlates of proxy measure of social isolation I have conceptualized include the following: race/ethnicity, eligibility for school lunch (proxy measure of SES), school size/location/type, school-level information completeness, school characteristics as measured by school culture related to teacher absenteeism, parental support of student achievement, and problem with gang activities. I have carefully incorporated these factors into the alternative student nonresponse weight I have developed by applying logistic regression.

I also expect that alternative gap scores I re-estimate by key background variables such as race/ethnicity and school type, where I observe evidence of nonresponse bias so far, are likely to be wider. It is due to the pattern of participation in NAEP such that better performing students are found to be more likely to participate and poor performing students are less likely to participate, beyond and above what can be explained by a set of student factors (race/ethnicity, gender, eligibility for national school lunch) as well as school-level variables (school climate measures, school size, type, urbanicity, and location). The participation propensity scores I have incorporated into the alternative student nonresponse adjustment weighting factor reflect such a pattern of participation.

Thus I expect the NAEP achievement gap is likely to be wider in alternative weighting method, especially where background measures are found to be significant predictors of participation of 12^{th} graders in NAEP.

As described in the previous chapter, I calculate the estimated participation propensity for each NAEP participant based on the final full logistic regression coefficients. I compute the student nonresponse adjustment weight by taking the inverse of the estimated response propensity for each participating 12^{th} grader in NAEP. Using the propensity-score-based weight adjustment, I recalculate NAEP estimates of scale score in the 2000 NAEP Mathematics and Science, respectively. I perform analysis with WesVar to properly account for the complex multi-stage clustered NAEP sample design and to re-estimate NAEP scale scores with alternative nonresponse adjustment. I also adjust a set of replicate weights by a factor of alternative nonresponse weighting to produce proper standard errors of re-estimated NAEP scale scores.

Table 5-4. Effects of Weights on Estimates of Mean NAEP Scale Scores in Mathematics and Science, 2000 HSTS-linked NAEP at Grade 12

	Mathematics (0-500 scale)				Science (0-300 scale)			
	NAEP Final Weight		Own Final Weight with Alternative Nonresponse Adjustment		NAEP Final Weight		Own Final Weight with Alternative Nonresponse Adjustment	
	Score	SE	Score	SE	Score	SE	Score	SE
Overall Mean	303.1	1.1	302.3	1.0	146.6	1.0	145.4	1.0
Male	305.2	1.4	304.2	1.4	147.6	1.3	146.4	1.3
Female	301.3	1.1	300.6	1.0	145.6	1.1	144.5	1.1
(Male-Female)	4.0*	1.2	3.6*	1.2	2.0	1.3	1.9	1.4
White	309.1	1.1	308.2	1.1	152.9	1.1	152.1	1.1
Black	274.8	2.2	274.5	2.1	121.5	1.8	121.0	1.7
Hispanics	287.4	2.2	287.4	2.0	129.9	2.0	129.4	2.1
Others	320.2	3.7	318.5	4.0	150.5	3.7	148.8	3.1
(White - Black)	34.3*	2.2	33.7*	2.3	31.5*	2.0	31.1*	1.9
(White - Hispanics)	21.7*	2.2	20.9*	2.0	23.0*	1.9	22.7*	2.2
(White - Others)	-11.1*	3.4	-10.3*	3.8	2.4	3.7	3.3	3.0
Northeast	305.3	3.3	304.6	3.0	149.4	2.8	148.8	2.7
Midwest	308.6	1.7	308.0	1.8	150.0	1.7	149.3	1.8
South	298.2	1.9	298.0	1.7	142.4	1.3	141.6	1.2
West	303.0	2.2	300.6	2.6	147.4	2.9	143.8	2.7
(NE - Midwest)	-3.4	3.7	-3.4	3.5	-0.6	3.3	-0.5	3.3
(NE - South)	7.1	3.7	6.7	3.4	7.0*	3.1	7.1*	2.9
(NE - West)	2.2	3.9	4.0	3.8	2.1	4.0	4.9	3.8
Public	301.6	1.2	300.6	1.1	145.1	1.0	143.8	1.0
Private	318.5	2.7	318.2	3.0	163.5	1.5	163.3	1.6
(Private - Public)	17.0*	3.1	17.5*	3.2	18.4*	1.9	19.5*	1.9

Note: * significant at $p < .05$.

Table 5-4 summarizes re-estimated NAEP scale scores by subject in comparison with the official NAEP estimates produced, using the current NAEP weights developed for each the 2000 NAEP Mathematics and Science. Estimates in the table include NAEP scale scores overall and by key background variables, and achievement gap by key variables such as gender, race/ethnicity, and school type. Standard errors of estimates are included in the second column under each set of data. NAEP scale score results are a numeric summary of what students know and can do in a particular subject. Mathematics are on a scale of 0 to 500; Science on a scale of 0 to 300. Achievement gap describes student achievement in terms of the gap, for example, between Black and White students, between Hispanic and White students, and between male and female students. Evaluating achievement gap by key background variables is the essence of the "No Child Left Behind" mandates. Key education policies at the federal level are guided by their impacts on reducing such an achievement gap.

The most notable pattern in this table appears to be about how closely NAEP scale scores lie between estimation methods using NAEP final weight and my own alternative weight within each subject. Reweighting in mathematics lowers the NAEP mean estimates by 0.8 point on a scale score of 0-500. The gender gap is lowered by a mere 0.3 point. The mathematics achievement gap between White and Black 12th graders is narrowed by 0.4 point score. Reweighting widened the mathematic achievement gap between students attending private and public schools, by a mee 0.5 point score. The regional difference, in particular between schools in the Northeast and the West, gets about twice wider due to reweighting (2.2 points vs. 4.0 points).

Reweighting in science appears to lower scale scores overall and gap scores are found to be a little wider by key background variables including race/ethnicity, school type, and census region. The overall mean scale score in science declines by 1.2-points on a scale score of 0-300. Reweighting lowers science scores for both male and female students, thus not affecting the gender gap much. The science scores by race/ethnicity are generally lower than official estimates of NAEP. Thus the achievement gap between White and other races is not affected. The only exception is the achievement gap widened between White and others including Asian-Pacific American and American Indian students. Reweighting appeared to widen the achievement gap between students in private and public schools, with an increase of over 1-point. As was seen in mathematics, reweighting widened the regional gap of science scores, in particular between schools in the Northeast and the West, getting more than twice wider (2.1 points vs. 4.9 points). The reader is cautioned that given the size of associated standard errors, the observed change may be small.

6. Discussion and Conclusions

I began this dissertation research motivated by the relatively low response rate of NAEP at 12th grade (i.e., about 10% to 35% lower than rates at grades 4 and 8). I was concerned about the potential for nonresponse bias in NAEP estimates due to the difference between participants and nonparticipants in NAEP or the extent of covariance between NAEP variables of interest and response propensity, as Groves and Couper (2006) theorized. I explored from this research empirical implications in response propensity models of identifying student- and school-level factors affecting nonparticipation of 12th graders in NAEP. I examined NAEP estimates for 12th graders by applying the approach used by Abraham, Maitland, and Bianchi (2006) to evaluate the impact of nonresponse bias on NAEP estimates.

The analysis provides evidence on the origins and the implications NAEP nonparticipation associated with this broad context of nonresponse research I began. First, I have investigated nonresponse bias, using a concept of social isolation (or social integration) to identify a set of variables applied to developing response propensity models. I have analyzed to the NAEP 2000 data a social isolation construct which Groves and Couper (1998) applied initially in household surveys. It can be seen as a social integration approach to building nonresponse models proposed by Lepkowski and Couper (2002) and by Abraham, Maitland, and Bianchi (2006). The social isolation framework has been applied to investigate how a set of factors determining 12th graders' participation in NAEP might be useful to evaluate their effects on sequential process of participation involving contactability and cooperation. The contactability model takes into account absence (i.e., noncontact); the cooperation model, refusal either by students or by parents on behalf of their children. I have documented that the contribution of absence to NAEP nonparticipation is about 50% higher than for refusal by students and their parents. The utility of the HSTS-linked NAEP data is demonstrated by testing the social isolation hypotheses and designing approaches to improve nonresponse bias analysis. It should be noted that this research, constrained by lack of direct measures of social isolation, could include such a social psychological measure of social isolation, using scales of shyness, introversion, and lack of social skills. It is also desirable to measure school-level factors of social isolation/integration by tapping students' involvement in study groups, after-school activities, religious organizations, and volunteer activities in order to associate the scope of these voluntary activities with participation in NAEP.

Second, I find evidence of significant relationships between participation and a number of student- and school-level variables, but no evidence that reweighting the data in the fashion as suggested by alternative response propensity models has affected the NAEP estimates. I have found evidence confirming the covariance between NAEP variables of interest and response propensity. Namely I observed a significant relationship of response propensity with measures of academic achievement (e.g., Carnegie credit and GPA) and contextual measures of school culture (e.g., perception of problem with teacher absenteeism and parental support of student achievement), respectively. In the mathematics NAEP, I observed higher participation rates for 12th graders whose academic achievement suggests better academic performance at school as measured by Carnegie credits, even after controlling for student characteristics -- such as sex, race/ethnicity and school-level variables such as school type, urbanicity and school size. In the science NAEP, I observe student achievement as measured by GPA plays an essential role in predicting participation rates in the context of controlling for a set of student- and school-level variables as used for science NAEP.

However, when the response propensity models derived from multivariate logistic regressions are applied to re-estimating NAEP scale scores, there is no evidence that reweighing the data has a significant or meaningful effect on the NAEP estimates in both mathematics and science. That is not a ground to rule out nonresponse bias in NAEP estimates, since other subject-specific student- or school-level variables could account for the differences between participants and nonparticipants. Reweighting with my own alternative nonresponse adjustment has lowered the mathematics mean estimates by a mere 1-point on a score scale of 0-500 and the science mean estimates by approximately 1-point on a scale of 0-300. When comparing NAEP estimates calculated from the official NAEP weight and my own alternative weight, the achievement gap in NAEP mathematics appears to be pretty close to each other by gender and race/ethnicity. The mathematics achievement gap gets a little wider when comparing the private-public achievement gap, and it gets notably wider when evaluating regional differences, in particular between schools in the Northeast and West. I observe a similar pattern in science NAEP when evaluating the impact on estimates of nonresponse bias in science with a re-weighted factor.

This research extends the findings by Curtin, Presser and Singer (2000) in demonstrating minimal damage of nonresponse bias. A traditional notion of linking high nonresponse rate to high response bias has been also challenged by Keeter et al. (2000) and Merkle and Edelman (2002) who showed no strong relationship between nonresponse rates and nonrespnse bias. Groves (2006) further demonstrated this by meta-analyzing 235 estimates from 30 studies that there is little empirical support to associate nonresponse rates to nonresponse bias. Findings from the current research with NAEP data strengthen such an argument.

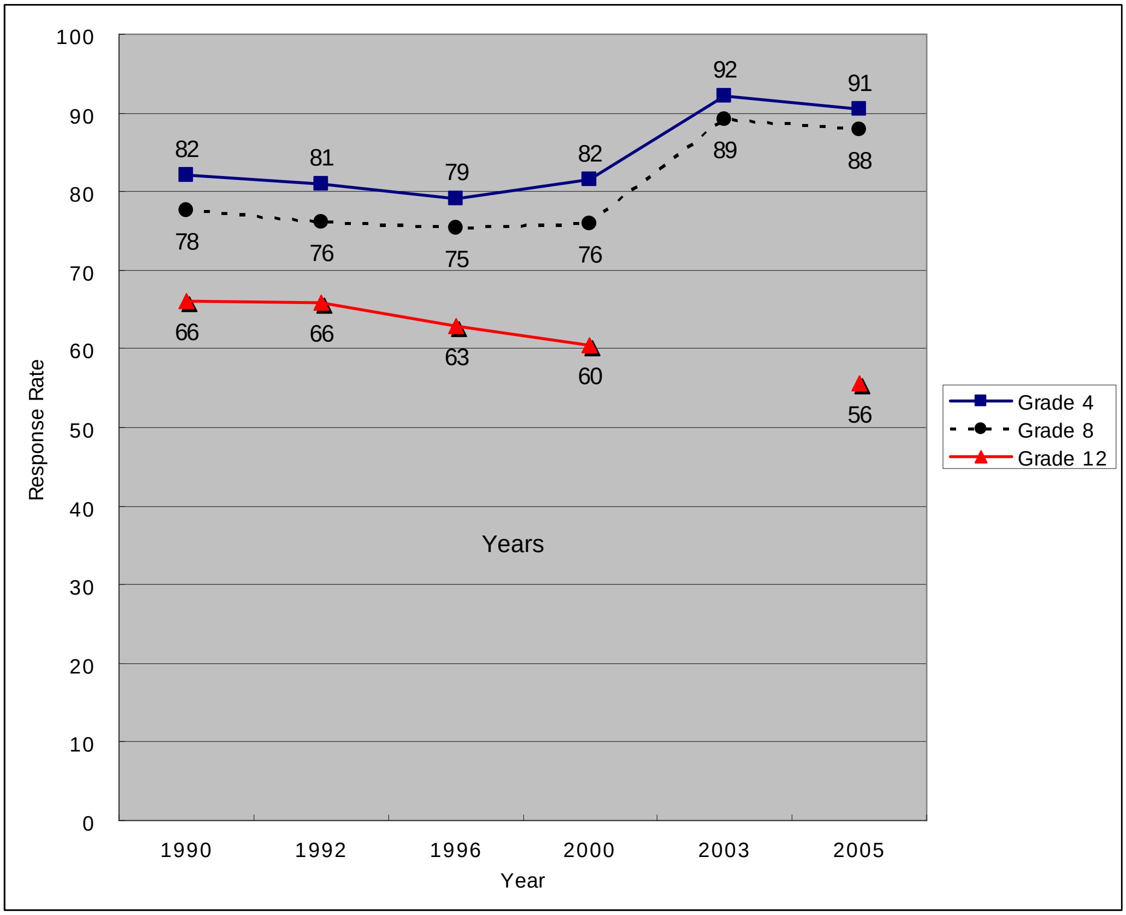
NAEP scores in 2000 mathematics and science reweighted with my response propensity model would not affect most of statistical inferences made about achievement gap by key variables in the year 2000, as the net effects on NAEP scores appear not to be large. Previous NAEP publications in mathematics and science indicate that even one-point of scale score can on occasion make a difference especially when it is about the achievement gap by such key variables as gender, race/ethnicity, and eligibility for national school lunch program (a proxy measure of poverty).

Third, it might be useful to develop in the future a nonparticipation index, an indicator of participation difficulty. This indicator may be constructed on the basis of a response propensity model of student- and school-level variables. Such a nonparticipation index may be linked specifically to the origins of nonparticipation -- student refusal, parental refusal, and student absence -- so that corresponding conversion strategies can be effectively developed in the NAEP field of data collection. NCES recently reported that the response rate of NAEP at grade 12 has been increased in the 2007 Writing Assessment, speculating it was perhaps due to design changes, best practice guidelines that recently began (e.g., offering more make-up sessions of NAEP assessment at school), or demographic shifts in the student population. However, it is not empirically possible to confirm which of design changes or best practice has contributed to increasing the response rate at grade 12. No experimental studies have been carried out to test the impact of individual NAEP features on increasing response rate.

The 2009 (January to March) round of NAEP Mathematics and Science at grade 12 will not be officially released until 2010 to detect changes in response rates and performance scores in Mathematics and Science at grade 12. Despite this uncertainty and lack of any experimental studies of intervention, it would be desirable to continue offering more make-up sessions of NAEP assessment at school. Twelfth graders' absence in NAEP assessment happens for various reasons. Empirical findings support that Black 12th graders attending large public schools in urban areas are more likely to be absent, compared to peers in other race/ethnicity groups. If students in this school setting are more encouraged for participation by additional make-up sessions, it should reduce a potential bias due to noncontact in particular.

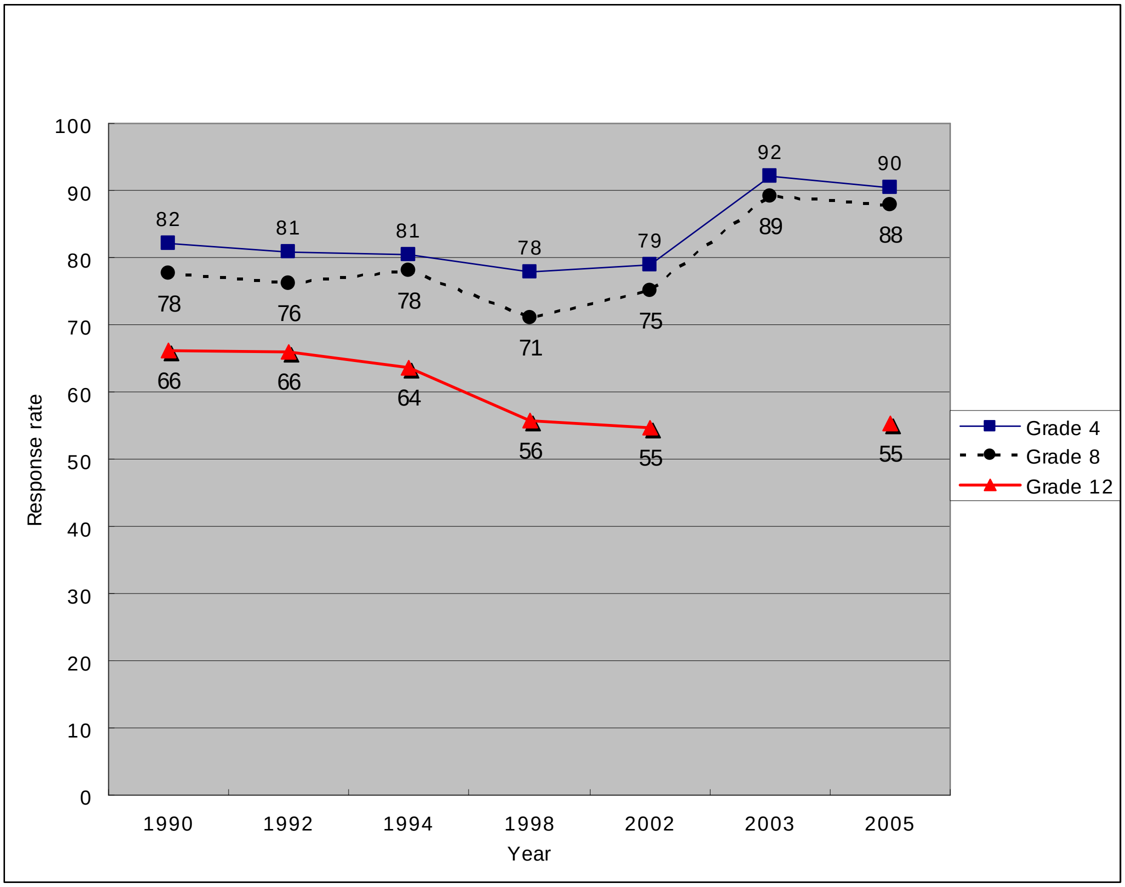
Appendix

Figure A-1. Overall school and student response rates before substitution, NAEP Mathematics, by grade: Various years, 1990 to 2005



Note: The 2003 NAEP Mathematics Assessment did not include grade 12.
Source: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1998, 2002, 2003, and 2005 Mathematics Assessments.

Figure A-2. Overall school and student response rates before substitution, NAEP Reading, by grade: Various years, 1990-2005



Note: The 2003 NAEP Reading Assessment did not include grade 12.
Source: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 1990, 1992, 1998, 2002, 2003, and 2005 Reading Assessments.

Chart A-1. NAEP Nonparticipation Model

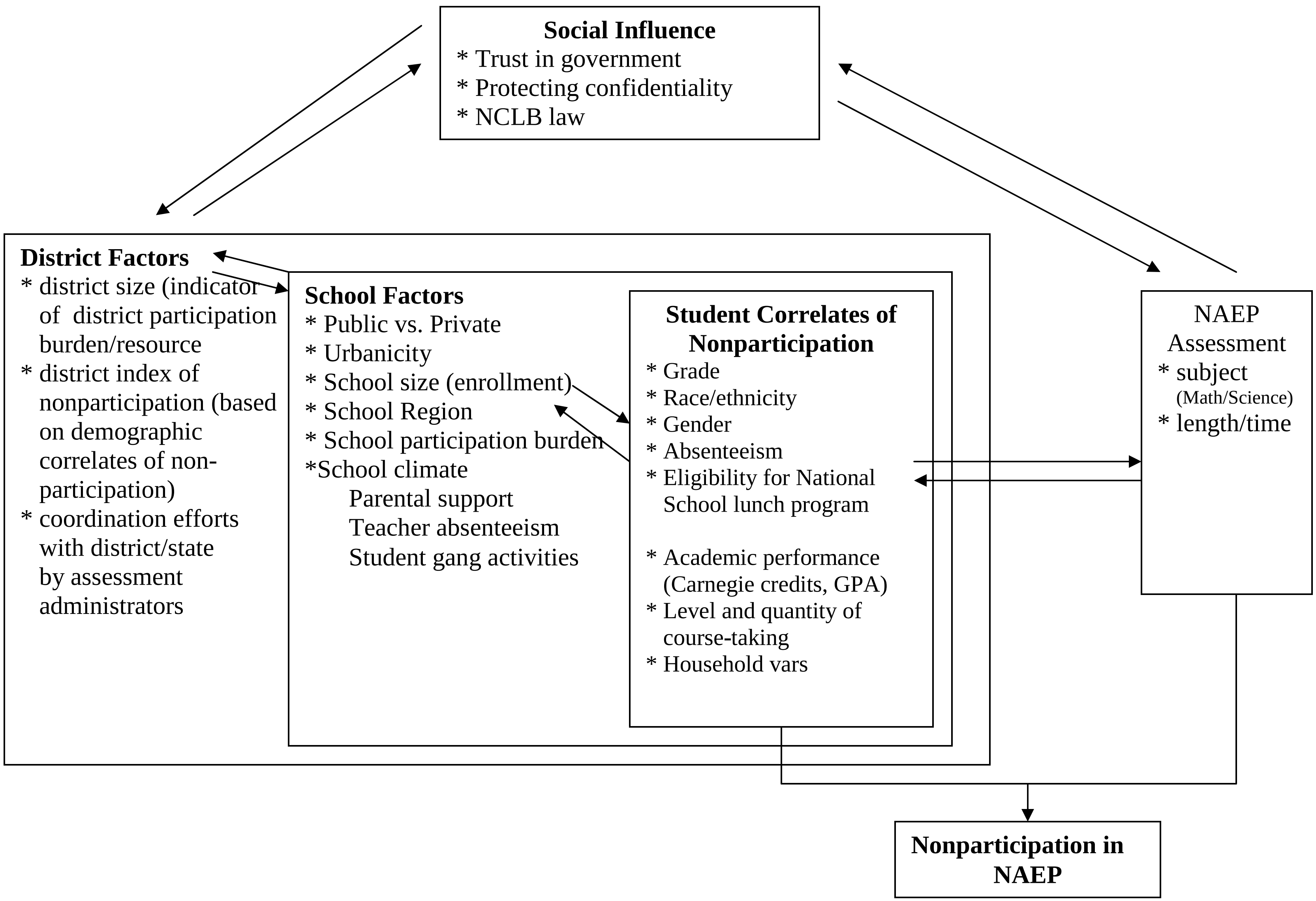


Chart A-2. Subject areas assessed, by assessment type: Various years, 1969–2001

Assessment year	Subject areas assessed in NAEP national main assessments	Subject areas assessed in NAEP long-term trend assessments	Subject areas assessed in NAEP state assessments
1969–70	Citizenship Science Writing	Science[1]	§
1970–71	Literature Reading	Reading[1]	§
1971–72	Music Social studies	†	§
1972–73	Mathematics Science	Mathematics[1] Science[1]	§
1973–74	Career and occupational development writing	†	§
1974–75	The arts Index of basic skills Reading	Reading[1]	§
1975–76	Citizenship/social studies Mathematics[2]	Citizenship/social studies[1]	§
1976–77	Basic life skills[2] Science	Science[1]	§
1977–78	Consumer skills[2] Mathematics	Mathematics[1]	§
1978–79	The arts Music Writing	†	§
1979–80	The arts Literature Reading	Reading[1]	§
1981–82[3]	Citizenship Mathematics Science Social studies	Mathematics[1] Science[1]	§
1984	Reading Writing	Reading Writing	§
1986	Computer competence Literature[2] Mathematics Reading Science U.S. history[2]	Mathematics Reading[4] Science	§
1988	Civics Document literacy[2] Geography[2] U.S. history Reading Writing	Civics[1] Mathematics Reading Science Writing	§
1990	Mathematics Reading Science	Mathematics Reading Science Writing	Mathematics[5] (gr 8 only)
1992	Mathematics Reading Writing	Mathematics Reading Science Writing	Mathematics[5] (gr 4 and 8) Reading[5] (gr 4 only)
1994	Geography Reading U.S. history	Mathematics Reading Science Writing	Reading[5] (gr 4 only)
1996	Mathematics Science	Mathematics Reading Science	Mathematics (gr 4 and 8) Science (gr 8 only)

Chart A-2. Subject areas assessed, by assessment type: Various years, 1969–2001

Assessment year	Subject areas assessed in NAEP national main assessments	Subject areas assessed in NAEP long-term trend assessments	Subject areas assessed in NAEP state assessments
		Writing	
1997	The arts (grade 8 only)	†	†
1998	Civics Reading Writing	†	Reading (gr 4 and 8) Writing (gr 8 only)
1999	†	Mathematics Reading Science	†
2000	Mathematics Reading (grade 4 only) Science	†	Mathematics (gr 4 and 8) Science (gr 4 and 8)
2001	Geography U.S. history	†	†

§ State assessments began in 1990.
† Not applicable; no subjects were assessed.
[1] This assessment appears in reports as part of long-term trend. Note that the civics assessment in 1988 is the third point in trend with citizenship/social studies in 1981–82 and in 1975–76. There are no points on the trend line for writing before 1984.
[2] This was a small, special study administered to limited national samples at specific grades or ages and was not part of a larger national main assessment. Note that this table includes only assessments administered to in-school samples; not shown are several special NAEP assessments of adults.
[3] Explanation of format for year column: Before 1984, the national main NAEP assessments were administered in the fall of one year through the spring of the next. Beginning with 1984, the national main assessment was administered after the new year in winter, although the assessments to measure long-term trend continued with their traditional administration in fall, winter, and spring. Because the national main assessment is the largest component of NAEP, beginning with 1984 its administration year is listed (rather than the two years over which trend continued to be administered.) Note also that the state assessment is administered at essentially the same time as the national main assessment.
[4] The 1986 long-term trend reading assessment is not included on the trend line in reports because the results for this assessment were unusual. Further information on this reading anomaly is available in Beaton and Zwick (1990).
[5] State assessments in 1990–94 were referred to as Trial State Assessments.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP).

Chart A-3. Reclassification of NAEP Disposition Codes into Own Sources of Nonparticipation: 2000 National Assessment of Educational Progress

NAEP Definition	NAEP Disposition Codes	Own Types of Nonparticipation	Number of Cases
Total			23522
Assessed students – original session			
In session full time	10	P	13540
No responses in booklet while student was in session full time.	11	R	10
In session part time. Student left the session in the middle of assessment	12	R	322
Original session incomplete due to interruption like fire drill or incompletion of a hands-on science booklet	13	ON	32
Other (e.g. a page missing from an assessment booklet)	14	ON	24
Assessed students – makeup session			
In session full time	20	P	1467
No responses in booklet while student was in session full time.	21	R	44
In session part time. Student left the session in the middle of assessment	22	R	68
Original session incomplete due to interruption like fire drill or incompletion of a hands-on science booklet	23	ON	11
Other (e.g. a page missing from an assessment booklet)	24	ON	5
Absent student			
Temporarily not in school (less than 2 weeks) due to illness or disability	40	NCA	3320
Long-term not in school (more than 2 weeks) due to illness or disability	41	I	38
Chronic truant. Student attends school occasionally, if ever.	42	I	38
Suspended or expelled including in-school suspension	43	I	26
In school yet did not attend session (e.g., student was known to be in school yet not released by teacher)	44	R	248
Disruptive behavior. Student in school yet not notified of assessment because of disruptive behavior	45	I	13
Parent refusal. Parent officially notified school of not allowing student to participate in assessment	46	R	283
Student refusal. Student refused to participate before being given an assessment booklet.	47	R	715
Other absence (e.g., student came to session too late.)	48	R	319
Given wrong booklet	49	I	31

Continued - Chart A-3. Reclassification of NAEP Disposition Codes into Own Sources of Nonparticipation: 2000 National Assessment of Educational Progress

NAEP Definition	NAEP Disposition Codes	Own Types of Nonparticipation	Number of Cases
Not linked to NAEP Mathematics or Science: Ineligible due to withdrawal, home schooled, not in sample yet assessed at school convenience	51-56	I	1512
Excluded due to extreme disability or so limited English proficiency	60-66	I	1243
Assessed with accommodations provided for students with moderate disability or language proficiency	70-79	P	213

NOTE. The abbreviations below are used for own types of nonparticipation. Participation outcomes are reclassified in consultation with NCES. P = Participation in assessment. ON = Other nonparticipation. NCA = Noncontact absence. R = Refusal. I = Ineligible
SOURCE: U.S. Department of Education, National Center for Education Statistics, High School Transcript Study (HSTS), 2000; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Science Assessment; U.S. Department of Education, National Center for Education Statistics, National Assessment of Educational Progress (NAEP) 2000 Mathematics Assessment.

Chart A-4. WesVar Procedure of Analysis of NAEP Scale Scores

Step 1. I prepare WesVar data files for mathematics and science from an SPSS data file that includes the following: 1) weights of choice (official HSTS-linked NAEP weight, and alternative final weight I have developed; 2) the variables that identify case ID, final sampling weight, strata, primary sampling units; 3) variables of interests including gender, race/ethnicity, census region, public vs. private, and five sets of plausible values in the restricted-use data to estimate NAEP score distribution by key background variables; and 4) 62 sets of replicate weights adjusted by alternative weighting factor for each assessment subject.

Step 2. I import 62 sets of replicate weights of HSTS-linked NAEP as provided by NCES (Roey, S., et al., 2005) and another 62 sets of replicate weights adjusted by alternative weighting factor by assessment subject. WesVar uses one of five replication methods to calculate variance of survey estimates. I select a replication method of my choice, Jackknife 2, for proper NAEP analysis (NCES, 2005)

Step 3. I create a WesVar workbook and specified my analysis to generate NAEP scale score distribution by key background variables with two weights, respectively, official HSTS-linked final weight, and alternative final weight of my development based on social isolation theory.

References

American Association for Public Opinion Research (2008), www.AAPOR.org. Obtained on October 1, 2008

Abraham, K.G., Helms, S.E., and Presser, S. (Forthcoming). How social processes distort measurement: the impact of survey nonresponse on estimates of volunteer work in the United States. *American Journal of Sociology*.

Abraham, K.G., Maitland, A., and Bianchi, S.M. (2006). Non-response in the American Time Use Survey: Who is Missing from the Data and How Much Does It Matter? Public Opinion Quarterly, 70, 676-703.

Appleton, J., Christenson, S.L., and Furlong, M.J. (2008), "Student engagement with school: Critical conceptual and methodological issues of the construct," Psychology in the Schools, Vol. 45, Issue 5, 369-386.

Bayard, S. R. (2003). A study of the relationship between teacher absenteeism, teacher attributes, school schedule and student achievement. Unpublished Ed.D, Florida Atlantic University.

Beaton, A.E., and Zwick, R. (1990). The effect of changes in the National Assessment: Disentangling the NAEP 1985-86 reading anomaly. (No. 17–TR–21) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Brehm, J. (1993). The Phantom Respondents; Opinion Surveys and Political Representation. Ann Arbor: University f Michigan Press.

Brick J., Michael J., Ferraro, D., Strickler, T. and Rauch, C. (2003). No. 8: The 2002 NSAF Response Rates. Urban Institute, Washington, DC.

Burgess, D., Haney, B., Snyder, M., Sullivan, J.L., and Transue, J.E. (2000). Rocking the vote: Using personalized messages to motivate voting among young adults. *Public Opinion Quarterly*, 64, 29- 52.

Campanelli, P., Sturgis, P. and Purdon, S. (1997). "Can you hear me knocking: an investigation into the impact of interviewers on survey response rates." London: the Survey Methods Centre at SCPR.

Cantrell, S. (2003). *Pay and performance: the utility of teacher experience, education, credentials, and attendance as predictors of student achievement at elementary schools in LAUSD*. Los Angeles, CA: Los Angeles Unified School District, Program Evaluation and Research Branch.

Carton, A. and Loosveldt, G. (1999). How the initial contact can be a determinant for the final response rate in face-to-face surveys. Department of Sociology, K.U. Leuven. Retrieved Dec 13, 2005 from <http://www.jpsm.umd.edu/icsn/papers/carton.htm>.

Chanley, V.A., Rudolph, T.J., and Rahn, W.M. (2000). The origins and consequences of public trust in government. *Public Opinion Quarterly*, 64, 239-256.

Chapman, E. (2003) "Assessing student engagement rates," ERIC Clearinghouse on Assessment and Evaluation. ERIC identifier: ED482269.

Chun, Y. and Scott, L. (2003). Would Social Isolation Theory Explain Teacher Participation in Surveys? Paper presented at the 57th annual conference of the American Association for Public Opinion Research, New Orleans, LA.

Chun, Y.., and Robertson, K.W. (1996). Reducing nonresponse in business surveys, Bureau of Labor Statistics. Retrieved Dec 13, 2005 from <http://www.amstat.org/sections/SRMS/Proceedings/papers/1996_170.pdf>.

Cohen, G. and Duffy, J.C. (2002). Are nonrespondents to health surveys less healthy than respondents?. *Journal of Official Statistics,* Vol. 18, No. 1, 13-23

Curtin, R., Presser, S. and Singer, E. (2000). "The Effects of Response Rate Changes on the Index of Consumer Sentiment." Public Opinion Quarterly, 64:413-428.

Davern, M., Rockwood, T.H., Sherrod, R., and Campbell, S. (2003). Prepaid monetary incentives and data quality in face-to-face interviews. *Public Opinion Quarterly,* 67, 139-147.

de Leeuw, E. and de Heer, W. (2002). “Trends in household survey nonresponse: a longitudinal and international comparison” in Groves, R. M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2002), Survey Nonresponse. New York: John Wiley and Sons, Inc.

Dunkelberg, W., and Day, G.S. (1973). “Nonresponse bias and callbacks in sample surveys,” Journal of Marketing Research, 10 (May): 160-168.

Epstein, J. (2001). *School, family, and community partnerships*. Boulder: Westview Press.

Fischer, C. (1982). To dwell among friends: Personal networks in town and city. Chicago: University of Chicago Press.

Gallagher, P.M., Fowler, F.J., and Stringfellow, V.L. (1999). “The Nature of nonresponse in a Medicaid survey: Causes and consequences.” Paper presented at the International Conference on Survey Nonresponse, Portland, Oregon.

Goldstein, K.M. and Jennings, M.K. (2002). The effect of advance letters on cooperation in a list sample telephone survey. *Public Opinion Quarterly*, 66, 608-617.

Gottfredson,D. (1989).Developing effective organizations to reduce school disorder. In O.Moles (Ed.), *Strategies to reduce student misbehavior* (pp. 87–104).Washington, DC: Office of Educational Research and Improvement.

Goyder, J. (1986). "Surveys on Surveys: Limitations and Potentialities," Public Opinion Quarterly, 50, 1, 27-41.

Goyder, J. (1987). The Silent Minority: Nonrespondents on Sample Surveys. Boulder, CO: Westview Press

Groves, R.M. (1989). Survey errors and survey costs. New York: John Wiley and Sons, Inc.

Groves, R. M. (2006), Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, 70, 646-675.

Groves, R.M., Cialdini, R.B., and Couper, M.P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56, 475-495.

Groves, R.M., and Couper, M.P. (1998). *Nonresponse in household interview surveys.* New York: John Wiley and Sons, Inc.

Groves, R.M., Couper, M.P., Presser, S, Singer, E., Tourangeau, R., Acosta, G.P., and Nelson, L. (2006). Experiments in Producing Nonresponse Bias, *Public Opinion Quarterly*, 70, 720-736.

Groves, R.M., Dillman, D.A., Eltinge, J.L., and Little, R.J.A. (2002). Survey Nonresponse, New York: John Wiley and Sons, Inc.

Groves, R. M. and Peytcheva, E. (2008), The impact of nonresponse rates on nonresponse bias, *Public Opinion Quarterly*, 72, 167-189.

Groves, R. M., Presser, S and Dipko, S. (2004). "The Role of Topic Interest in Survey Participation Decisions." *Public Opinion Quarterly*, 68 (1): 2-31.

Groves, R.M., Singer, E., and Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly*, 64, 299-308.

Groves, R. and Wissoker, D. (1999). Methodology Report No. 7: Early Nonresponse Studies of the 1997.National Survey of America's Families. Washington, DC: Urban Institute.

Guadagnoli, E. and Cunningham, S. (1989). “The Effects of Nonresponse and Late Response on a Survey of Physician Attitudes.” *Evaluation and the Health Professional* 12:318-328.

Hess, B.B. and Waring, J.M. (1978) Changing patterns of aging and family bonds in later life. The Family Coordinator, 27.

Hortulanus, R, Machielse, A, and Meeuwesen, L. (2006). Social isolation in modern society. NY: Routledge.

House, J.S., Robbins, C. and Metzner, H.L (1982). The association of social relationships and activities with morality: Prospective evidence from the Tecumseh community health study. American Journal of Epidemiology, 116. 123-140.

House, J.S., and Wolf, S. (1978). “Effects of Urban Residence on Interpersonal Trust and Helping Behavior.” Journal of Personality and Social Psychology. 36, 9, 1029-1043.

Howell, J. C., and Lynch, J. (2000). *Juvenile Justice Bulletin*. U.S. Department of Justice.

Washington, D.C.

Keeter, S., Kohut, A., Miller, A., Groves, R., and Presser, S. (2000). Consequences of reducing non-response in a large national telephone survey. Public Opinion Quarterly, 64(2), 125-48.

Kish, L. (1965). Survey Sampling. NY: John Wiley.

Knapp, M. and Shields, P. (1990). Reconceiving academic instruction for the children of poverty. *Phi Delta Kappan*, 71: 753-58.

Krewski, D. and Rao, J. (1981). Inference from stratified samples: properties of linearization, jackknife and balanced repeated replication. *Annals of Statistics*, 9, 1010-1019.

Kulka, R.A. (1994). "The use of incentives to survey 'hard-to-reach' respondents: A brief review of empirical research and current practice." Paper prepared for Seminar on New Directions in Statistical Methodology, Bethesda, MD.

Kulka, R.A., and Singer, E. (May 26, 2000). Paying respondents for survey participation, Draft Paper.

Lee, V, and Smith, J.B. (2001). Restructuring high schools for equity and excellence. New York: Teachers College Press, 2001.

Lepkowski, J.M. (1988). "The treatment of wave nonresponse in panel surveys." In Kasprzyk, D., Ducan, G., and Singh, M.P. (eds.), Panel Survey Design and Analysis. NY: Wiley.

Lepkowski, J.M. and Couper, M. (2002). "Nonresponse in the second wave of longitudinal household surveys," in Groves, R., Dillman, D., Eltinge J., and Little R. (eds), Survey Nonresponse, 259-274, New York: Wiley.

Lin, I. F., and Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. *Public Opinion Quarterly*, 59:236-258.

Link, M.W., Malizio, A.G., and Curtin, T.R. (2001), Use of targeted monetary incentives to reduce nonresponse in longitudinal surveys. American Association of Public Opinion Research.

Little, R., and Vartivarian, S (2003). On weighting the rates in nonresponse weights. Statistics in Medicine 22:1589-99.

Lord, F.M. (1980). Applications of Item Response Theory to Practical Testing Problems. Hillsdale, NJ: Lawrence Erlbaum Associates.

Merkle, D. and Edelman, M. (2002). "Nonresponse in exit polls: a comprehensive analysis," in Groves, R., Dillman, D., Eltinge J., and Little R. (eds), Survey Nonresponse, pp 243-258, New York: Wiley.

Mcevoy, A., and Welker, R. (2000). Antisocial behavior, academic failure, and school climate: a critical review. Journal of Emotional and Behavioral Disorders

Miller, D. (1980). Family maladaptation reflected in drug abuse and delinquency. In M. Sugar (Ed.), Responding to adolescent needs. NY: S Medical and Scientific Books.

Mislevy, R.J., and Bock, R.D. (1982). BILOG: Item Analysis and Test Scoring with Binary Logistic Models [Computer program]. Chicago: Scientific Software.

Muraki, E., and Bock, R.D. (1997). PARSCALE: IRT Item Analysis and Test Scoring for Rating-Scale Data. Chicago, IL: Scientific Software International.

National Center for Education Statistics (2000). The Nation's Report Card: Mathematics Highlights 2000, Washington DC.

National Center for Education Statistics (2003). NCES Statistical Standards, Institute of Education, U.S. Department of Education, NCES 2003-601. Washington, DC.

National Research Council. 2003. *Engaging Schools: Fostering High School Students' Motivation to Learn*. Committee on Increasing High School Students' Engagement and Motivation to Learn, National Research Council. Washington, DC: National Academy Press.

Newmann, Fred M. 1992. *Student Engagement and Achievement in American Secondary Schools*. New York: Teachers College Press.

Newman, Richard Stuart, and L. Goldin. 1990. "Children's Reluctance to Seek Help with School Work." *Journal of Educational Psychology*, 82: 92-100.

Newmann, Fred M., Gary Wehlage, and Susie D. Lamborn. 1992. "The Significance and Sources of Student Engagement." Pages 11-39 in Fred M. Newmann (ed.), *Student Engagement and Achievement in American Secondary Schools.* New York: Teachers College Press.

Neuman, S., Hagedorn, T., Celano, D., and Daly, P. (1995). Toward a collaborative approach to parent involvement in early education: a study of teenage mothers in an African community. American Educational Research Journal, 32 (4), 801-827.

Peplau, L. A. and Perlman, D. (1982). Loneliness: a sourcebook of current theory, research and therapy. NY: John Wiley.

Presser, S. (1981). "A comparison of Harris sample composition with census, IRS, and NORC," In E. Martin, D. McDuffee, and S. Presser (eds.), Source book of Harris National Surveys: Questions 1963-76. Chapel Hill, North Carolina; Institute for Research in Social Science.

Roach, R. (2001). Where are the Black men on campus? *Black Issues in Higher Education*, 18(6), 18-21.

Roey, S., Caldwell, N., Rust, K., Hicks, L., Lee, J., Perkins, R., Blumstein, E., and Brown, J. (2005). The 2000 High School Transcript Study User's Guide and Technical Report (NCES 2005–483). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

Schleifer, S. (1986). Trends in attitudes toward and participation in survey research. *Public Opinion Quarterly*, 50, 17-26.

Singer, E., Groves, R.M., and Corning, A.D. (1999). Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation. *Public Opinion Quarterly*, 63, 251-260.

Singer, E., Hoewyk, J.V., and Maher, M.P. (2000). Experiments with incentives in telephone surveys. *Public Opinion Quarterly*, 64, 171-188.

Skinner, E.A., and Belmont, M.J. (1993). "Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year." *Journal of Educational Psychology*, *85*(4).

Smith, T. (1983). "The Hidden 25 percent: an analysis of nonresponse on the 1980 General Social Survey," Public Opinion Quarterly, 47, 1.

Smith, T. (1984). "Estimating Nonresponse bias with temporary refusals." Sociological Perspectives, 27, 4, 473-489.

Stoop, I. (2005). The hunt for the last respondent: nonresponse in sample surveys, The Hague: Social and Cultural Planning Office of the Netherlands

Swanson, C.B. (2006). High school graduation in Texas: Independent research to understand and combat the graduation crisis. Editorial Projects in Education Research Center, Bethesda, MD.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2000). The High School Transcript Study. Washington, D.C.

Vaillancourt, P.M. (1973). Stability of children's survey responses. *Public Opinion Quarterly*, 37, 373-387.

Warriner, K., Goyder, J., Gjertsen, H., Hohner, P., and McSpurren, K. (1996). Charities, no; Lotteries, no; Cash, yes. *Public Opinion Quarterly*, 60, 542-562.

Westat. (2003a). Analysis of private school nonresponse bias in the 2003 NAEP.

Westat. (2003b). Logistic regression analysis for predicting response status of private schools in the 2003 NAEP.

Whitworth, E.M. (1999). "Can the use of a second mailing increase mail response rates in historically hard-to-enumerate Census tracts?" U.S. Census Bureau. Paper presented at the International Conference on Survey Nonresponse, Portland, Oregon.

Willimack, D.K., Schuman, H., Pennell, B, and Lepkowski, J. (1995). Effects of a prepaid nonmonetary incentive on response rates and response quality in a face-to-face survey. *Public Opinion Quarterly*, 59, 78-92.