

# The Influence of Selected Factors on Student Survey Participation and Mode of Completion

Tracy Hunt-White

National Center for Education Statistics  
1990 K Street, NW, Room 8121, Washington, D.C. 20006 / tracy.hunt-white@ed.gov

## Introduction

Surveys have become an integral part of modern society. The desire and need to measure opinions, behaviors, satisfaction, and trends have led to a growth in survey research. For example, news media conduct polls to test the political climate, businesses examine employee and customer satisfaction, government agencies investigate national household behaviors, and colleges want to know about the experiences and satisfaction of their students and alumni. Most organizations conduct extensively planned and designed studies to capture results that will help them learn how to better serve their constituents. However, to survey every member of a target population is too costly and time consuming; therefore, a representative sample is selected to reduce expense and effort. Since this sample stands for the entire target population, each person's participation is critical; hence, the higher the response rate, the more accurately the collected data represents the entire population.

Although survey procedures have become more efficient and the number of methods for collecting data has increased, people have become more reluctant to participate. Hence, all types of survey sponsors (government, academia, business, media, etc.) have been experiencing declining response rates in their studies (Couper, 2000; De Heer, 1999). This reduced participation directly impacts the usefulness of the data collected to describe the population from which the individuals were selected. In order to make the survey worth the expense and effort, researchers are seeking ways to gain the cooperation of sample members and/or to adjust the survey results to account for those who did not respond.

This study attempts to examine college student survey participation using a modified version of Groves and Couper's Conceptual Framework of Survey Participation by Householders. The factors used in this study include respondent characteristics, institutional characteristics (social environment and technological environment), and survey design features. This study examines how these factors are related to survey participation overall and by mode of completion.

## Statement of the Problem

Research organizations are finding it increasingly difficult to obtain sufficient survey response rates (Couper, 2000; Fowler, 2002). Researchers are expending more effort and resources to maintain representative response rates for surveys (De Heer, 1999). Hence, they are examining features of their survey design and trying different methods in order to gain better cooperation from individuals. These methods include sending out advance letters, offering incentives, and providing multiple ways to complete the survey. However, the research is mixed on just how successful these practices are in increasing participation. Groves, Singer, and Corning (2000) found that other attributes of the survey and the experiences and background of the individual participant might contribute to the mixed results. This study examines the characteristics of the participants and their environment for clues on who is likely to respond and which mode they will use. While researchers have no control over these characteristics, they can use this information to adjust the design features of their survey to increase participation.

## Research Needs

This study was conducted because of the specific population of interest, college students, and the use of multiple modes. Little research has been done looking at how established factors of survey participation impact college students and self-administered survey completion. This study provided an opportunity to look not only at student characteristics, but also at their college environment and their relationship to survey participation and mode of completion chosen.

## Conceptual Framework of Survey Participation

Researchers have a long-standing interest in understanding why people participate in surveys. This is especially true for organizations that conduct repeated surveys and have watched their efforts to obtain high response rates grow more costly and time consuming. In this study, the following factors are considered to impact survey participation: respondent

characteristics, the social environment in which the survey request takes place, the technological environment, and the design features of the survey as they relate to the mode used to complete the survey and when someone responds. These factors are briefly explained in terms of a modified version of Groves and Couper's Conceptual Framework of Survey Participation by Householders (1998). For this study, their conceptual framework was adapted to include constructs for Web-based surveys using a model of Web survey participation by Vehovar, Batagelj, Lozar Manfreda, and Zaletel (2002).

Groves and Couper's model for survey participation showed the ways in which four constructs influence a person's decision to participate in a survey. These constructs are the social environment, respondent characteristics, survey design features, and interviewer characteristics. Groves and Couper built a comprehensive model that would help researchers understand survey nonresponse in interviewer-administered household surveys. This information could then be used to increase survey participation or to make more accurate post-survey adjustments for nonresponse. Groves and Couper give a full explanation of their model in their work titled *Nonresponse in Household Interview Surveys* (1998).

These constructs are widely cited in literature as influencing cooperation, and Groves and Couper's model is often the basis for other theories on survey participation. Researchers have expanded and modified the model to incorporate additional constructs relevant to different survey modes and survey populations. Vehovar et al. (2002) hypothesized that for Web-based surveys, a construct is needed for the technological environment in which the person completes the survey. They also expanded the definition of respondent characteristics to include a person's technical experience. However, with Web-based surveys, the interviewer construct can be eliminated since these surveys are completed directly by the respondents.

This study uses national data pertaining to college students and their institutions. This undertaking provides an opportunity to demonstrate whether these constructs—respondent characteristics, social environment, technological environment, and survey design—impact college students in the same way other populations were affected. Also, the use of multiple modes provides an opportunity to test how these factors are related to the mode a person selects to complete the survey. The factors are explained below.

**Respondent Characteristics.** All factors that influence the decision to cooperate in a study are filtered through the individual characteristics of the respondent. Each individual brings to fore a unique combination of background and experiences in which to interpret a request to participate in a survey. Researchers may have some information on the sociodemographics of the population they wish to study (e.g., gender, socioeconomic status, age, etc.). However, sociodemographics alone do not dictate a person's psychological disposition toward completing the survey. Other respondent characteristics may also influence survey participation. The individual variables comprising these other characteristics can be grouped into civic participation, educational background and experiences, employment, and technical experience.

**Social Environment.** The social environment consists of the global characteristics of any society in which the survey is conducted. There is a set of traits in any nation, state, or community that will have an impact on survey participation. These characteristics constrain the respondent and researcher by determining the context in which the survey request takes place. At the national level, issues such as economic policies and congressionally mandated federal studies impact survey participation. However, this study focuses on characteristics of the students' local community, i.e., the postsecondary institution.

**Technological Environment.** The technological environment consists of the global technological characteristics of the society in which the survey is conducted. These characteristics include the governmental or institutional telecommunication policies and the technological infrastructure existing at the time of the survey request that set parameters on both the respondent and researcher. This study focuses on the technology services and support available at the student's college.

**Survey Design.** The survey design factor deals with the decisions that researchers must make to conduct an effective study. This is the only construct in the model that is under the full control of the researcher. In addition to determining the survey topic, the target population, and the level of nonresponse they are willing to accept, researchers must also determine such issues as the modes used to collect the data, whether or not to offer incentives and the types of incentives, and the follow-up methods to be used. This study focuses on topic saliency (relevance of topic to individual), perceived legitimacy of the sponsor, modes of administration, and incentives.

Table 1 shows the four factors and their related variables used in this study. A subset of these variable was used in analyzing overall student interview participation since some variables were not available for student interview nonparticipants.

Table 1. List of constructs and variables analyzed

Construct and variables	
<p><b>Respondent characteristics</b></p> <p>Gender*</p> <p>Age*</p> <p>Race*</p> <p>Citizenship*</p> <p>Marital status*</p> <p>Disability</p> <p>Dependent student*</p> <p>Dependent children</p> <p>Civic duty (Military service)*</p> <p>Civic duty (Number of hours spent per month on community service)</p> <p>Cumulative college GPA</p> <p>Full-time attendance*</p> <p>Student level*</p> <p>Degree program*</p> <p>Highest level of education ever expected to complete</p> <p>Parents' highest degree</p> <p>Hours worked per week</p> <p>Job role</p> <p>Science or technology major</p> <p>Income</p> <p>Distance from home (miles)</p> <p>Taken distance education classes</p>	<p><b>Social environment</b></p> <p>Institution type*</p> <p>Urbanicity*</p> <p>Enrollment size*</p> <p>Tuition and required fees*</p> <p>Student-faculty ratio*</p> <p>Percentage of undergraduate living on-campus*</p> <p><b>Technological environment</b></p> <p>Institution offers distance education*</p> <p>Percentage of degrees awarded in science and technology*</p> <p>Number of computers available on campus*</p> <p>Students required to own computers*</p> <p>E-mail accounts for all students*</p> <p><b>Survey design</b></p> <p>Contact calls*</p> <p>Interest in topic (applied for student aid &amp; received aid)*</p> <p>Legitimacy of sponsor (received govt assistance)</p> <p>Legitimacy of sponsor (received federal student aid)*</p> <p>Incentives</p>

\*These variables are also available for student interview nonparticipants.

### Purpose of the Study

The four constructs of survey participation—respondent characteristics, social environment, technological environment, and survey design—constitute the framework by which this study views student interview participation and mode of completion chosen. Specifically, this study explores how various factors are related to a sample member's decision on whether and how to participate in a survey.

### Research Questions

This study attempted to answer two questions through the prism of the conceptual framework of survey participation by using data from a national study of postsecondary students and data pertaining to their institutions. An examination was conducted on how the factors in the framework impact whether a person participates in the survey and the mode of completion chosen (Web vs. telephone). The questions answered were:

- RQ1) How are respondent's characteristics, the institution's social and technological environments, and the survey's design features related to survey participation?
- RQ2) How are respondent's characteristics, the institution's social and technological environments, and the survey's design features related to mode of completion?

### Method

#### Data Sources

Three sources of data were used to identify the variables that comprise the factors associated with survey participation: the 2004 National Postsecondary Student Aid Study (NPSAS), the Integrated Postsecondary Education Data System (IPEDS), and the College Board.

**National Postsecondary Student Aid Study (NPSAS).** The National Postsecondary Student Aid Study (NPSAS) is a comprehensive study that examines how students and their families pay for postsecondary education. This is a sample study

conducted by the National Center for Education Statistics (NCES). It includes nationally representative samples of undergraduate and graduate students from all institution sectors. Students who receive student financial aid as well as those who do not receive aid participate in NPSAS. The study used student interviews and data from institutional administrative records and federal financial aid databases to provide details concerning a student's financial aid, background, enrollment pattern, educational experience, and employment status. NPSAS was conducted in academic years 1986-87, 1989-90, 1992-93, 1995-96, 1999-2000, and 2003-04.

Data from the 2003-04 NPSAS were used to identify the variables that form the respondent characteristics and the survey design factors. The data were also used to group sample members into the following outcome groups: Student interview participants versus student interview nonparticipants and Web respondents versus telephone respondents.

**Integrated Postsecondary Education Data System (IPEDS).** The Integrated Postsecondary Education Data System (IPEDS) was also a key source of data. IPEDS was established by NCES as a system of interrelated surveys designed to collect data from postsecondary education institutions. Data are collected at the institution-level for such topics as enrollment, degree completion, faculty, staff, finances, institutional characteristics, and academic libraries. Participation is mandatory for institutions that are eligible to participate in Title IV federal student assistance programs. Data from the 2002-03 IPEDS data collection cycle were used to identify variables that potentially comprise the social environment factor and part of the technological environment factor.

**College Board.** The College Board ([www.collegeboard.com](http://www.collegeboard.com)) collects information from colleges and universities on topics such as admissions, enrollment, campus programs, academic support services, and tuition and fees. Institutions voluntarily provide data to be published on the College Board's Web site and college guides. Variables regarding campus technology can be obtained from its Web site. However, because of the number of institutions in this study ( $n=1,400$ ), a data file was obtained containing the desired variables.

### Analytic Sample

The respondents for this research are a subset of study respondents from the 2004 National Postsecondary Student Aid Study (NPSAS). This sample is limited to approximately 69,000 NPSAS students who represented 18.3 million undergraduates enrolled in 2-year and 4-year public, private not-for-profit, and private for-profit postsecondary institutions in the U.S. in 2003-04. Study respondents for NPSAS not only included those who completed a Web-based or telephone interview, it also included students who were not interviewed but for whom key information could be obtained from other data sources (i.e., school records or federal financial aid databases) (Cominole et al., 2005). For this research, these NPSAS students were categorized as student interview participants and student interview nonparticipants. All analyses use data that have been weighted to represent the target population. Table 2 provides a breakout of the subjects by outcome group.

Table 2  
Number and percentage distribution of subjects by outcome

Outcome	Unweighted		Weighted	
	Number	Percent	Number	Percent
Student interview participation	68,900	100.0	18,258,000	100.0
Student interview participant	48,000	69.7	12,734,000	69.7
Student interview nonparticipant	20,900	30.3	5,524,000	30.3
Mode of completion	48,000	100.0	12,734,000	100.0
Web	22,100	46.0	5,794,000	45.5
Telephone	25,900	54.0	6,940,000	54.5

Table 3 shows the detailed characteristics of the individuals included in this study. This table includes the background characteristics of the students and characteristics about the institutions they attended. Respondents were predominately White (64%), female (57%), and single (79%). Approximately 48% of the students were enrolled full-time. One-third (34%) of the respondents were first-year undergraduates and one-fourth (26%) were second-year undergraduates. Respondents' average age was 26. About one-half (46%) of participants attended a public 2-year institution, one-third (33%) attended a public 4-year institution, 14% attended a private-not-for-profit 4-year institution, and 6% attended a private for-profit institution. A large proportion of participants attended schools located in large urban areas (47%) while 17% attended in non-urban areas.

Table 3  
Percentage distribution of students for selected variables

Construct and variables	Total	Construct and variables	Total
<b>Respondent characteristics</b>	Percent	<b>Social environment</b>	Percent
Gender		Institution type	
Female	57.2	Public 2-year	46.4
Male	42.8	Public 4-year	33.2
Race		Private not-for-profit 4-year	14.4
White	64.4	Private for-profit 2-years or higher	5.6
African American	13.8	Private not-for-profit 2-year	0.5
Hispanic	11.5	Urbanicity	
Asian	5.5	Non-urban area	16.6
American Indian/Alaska Nat	1.5	Mid-sized urban area	34.5
Other/Multiple races	3.4	Large urban area	47.4
Citizenship		Unknown	1.5
Non-US citizen	7.0	Mean	
US citizen	93.0	Enrollment size	11,685
Dependent student		Student-faculty ratio	19.9
Independent	49.6	Percentage undergrad on campus	14.6
Dependent	50.4	<b>Technological environment</b>	Percent
Marital status		Offers distance education	
Married	23.1	Inst does not offer distance educ	17.5
Single	76.9	Inst offers distance education	82.5
Military service		Students required to own computers	
Never had military service	94.7	Computer not required	96.3
Had military service	5.3	Computer ownership required	3.7
Full-time attendance		E-mail accounts for all students	
Not Full time	52.5	Institution does not provide e-mail	33.1
Full time	47.5	Institution provides e-mail acct	66.9
Student level		Mean	
1 <sup>st</sup> year undergraduate	33.9	Number of computers on campus	955
2 <sup>nd</sup> year undergraduate	26.2	Pct degrees awarded in sci/tech	12.4
3 <sup>rd</sup> year undergraduate	15.0	<b>Survey design</b>	Percent
		Application for and receipt of student aid	
4 <sup>th</sup> or 5 <sup>th</sup> year undergraduate	17.4	Applied for and received aid	62.6
Unclassified undergraduates	7.5	Applied for & didn't receive aid	11.3
Degree program		Didn't apply & didn't receive aid	26.2
Less than bachelor's	41.5	Receipt of federal student aid	
Bachelor's degree	48.1	No federal student aid	54.6
No degree	10.4	Received federal student aid	45.4
Age	Mean	Mean	
	26	Contact calls <sup>1</sup>	14

<sup>1</sup> The number of calls made to sample members to request survey participation.

### Outcomes of Interest

Two dependent variables were used in the analysis of the research problems for this study: student interview participation status and mode of completion. A detailed description of the variables follows below.

**Student interview participation.** This variable is an indicator of whether or not sample members responded to the student interview. NPSAS collected data from students via a Web-based or telephone interview, but it also captured information from other data sources such as institutional records and national financial aid databases. Hence, some information was available for the nonparticipants of the student interview. This outcome allowed an examination of how well selected factors predict survey participation.

**Mode of completion.** This variable refers to the method used by the respondent to participate in the survey. Respondents either completed the survey via a self-administered Web-based survey or via the interviewer-administered telephone interview. This outcome allowed an examination of how selected factors predict the mode chosen for survey participation.

### **Predictor Variables**

This study uses a model that identifies the relationship among four key constructs that impact overall survey participation and the mode of completion chosen: respondent characteristics, social environment, technological environment, and survey design features.

**Respondent Characteristics.** In this study, the respondent characteristic factor was grouped into the following areas: civic duty, educational background and experience, sociodemographic characteristics, employment, and technical experience. The literature supports the importance of civic participation and sociodemographics to survey participation (Groves, Singer, and Corning, 2000). Civic duty was determined by whether the respondent participated in community service or served in the military. Educational background included variables on class level, attendance status, college grade point average, degree type, field of study, and highest level of education ever expected to complete. Sociodemographic variables not only included gender and race/ethnicity, but also disability status, marital status, dependency status, and parent's highest level of education. Employment included a variable on the number of hours a student worked per week and whether the student considered himself or herself to be primarily a student working to meet expenses, or primarily an employee attending college. Finally, technical experience was based on whether the respondent has a science or technology-related major or had ever taken a distance education class.

**Social Environment.** The social environment factor pertains to the characteristics associated with the student's institution. It consists of the following subfactors: general economic development, institutional characteristics, social cohesion, population density, and location. In this study, the general economic development of the institution was determined by its tuition and fees. The institutional characteristics include the institution's control and level (public 2-year, public 4-year, etc.). Social cohesion, the concept of how connected a student feels to the campus, was measured by the student/faculty ratio and the percentage of undergraduates living on-campus. The population density and location dimensions consisted of the enrollment size and urbanicity of the campus location. Research shows that the more urban the setting, and the less connected respondents feel toward their community, the less likely they are to respond to a survey (Groves & Couper, 1998).

**Technological Environment.** The technological environment factor was anticipated to consist of data from IPEDS and the College Board. This study used variables pertaining to whether or not the college offered distance education classes, had a large percentage of science and technology majors, provided e-mail accounts to its students, required students to own a computer, and the number of computers available on-campus. The assumption was that a campus that provided distance education, had a large percentage of science and technology majors, and encouraged computer ownership and communication via the Internet had the campus infrastructure necessary to allow students to complete Web-based surveys.

**Survey Design.** This study focused on the following survey design features: topic saliency, sponsorship legitimacy, follow-up contact calls, and incentives. Incentive amounts and the number of contact calls were directly captured from the NPSAS data set. However, the other variables required identifying data elements to serve as proxy indicators. Therefore, a student's interest in the topic was approximated by whether or not he or she receives student financial aid. Whether or not the student received federal student aid, or government support was used to determine the perceived legitimacy of the survey sponsor. In this case, the sponsor was the federal government.

### **Analytic Approach**

Multivariate analyses were used to address the research questions in this study. Logistic regression was used to test how respondent characteristics, the social and technological environments, survey design factors were related to survey participation and the mode of completion chosen.

Prior to logistic regression, chi-square analyses for contingency tables and correlation analyses were conducted in order to decide on which variables should be included in the logistic regression analyses. Only variables that were determined to have a significant association with the outcome at the significance level of 95 percent were included in the model. These tests showed that the following variables were not related to the survey participation outcome: citizenship status, tuition and fees, student-faculty ratio, whether the institution offers distance education, or requires computer ownership, percentage of degrees awarded in science and technology, and whether student received federal aid. For the mode of completion outcome, only urbanicity was not related to the outcome.

In addition to determining the relationship between the dependent variable and the predictors, intercorrelation analyses were performed to examine the relationship among the independent variables in order to eliminate variables with high correlation (multicollinearity). The results of the intercorrelation analyses show that dependent student status is highly correlated with age, tuition and fees is highly correlated with institution type, and receive any financial aid is correlated with received federal financial aid. The variables age, institution type, and received any aid were retained for the logistic regression models.

To test the accuracy of the logistic regression results, the sample was randomly divided into two groups prior to conducting the initial logistic regression analyses. Tabachnick and Fidell (2001) recommend cross validation for statistical regression to ensure that the results are generalizable (p. 135). Logistic regression results show how well a model describes the observed data; it seeks the model that best fits the data. However, it does not show how well the model predicts the data. Since the model already “knows” the relationship among the data, model-fit statistics may overestimate how well the model predicts the outcome. Therefore, cross validation is intended to eliminate errors in these estimates; that is, to ensure that the results of the analyses were not due to chance. The sample was randomly divided into two groups: 80 percent of the sample in the first group and 20 percent in the second group. The first group was used to generate the logistic regression model that best fits the data. The second group was used to test whether the model specified by the initial group is correct.

## Results

Analyses yielded results to answer to how multiple factors—respondent characteristics, social environment, technological environment, and survey design—are related to survey participation overall and by mode of response.

### Student Interview Participation Outcome

A logistic regression model was developed to answer how respondent’s characteristics, the institution’s social and technological environments, and the survey’s design features were related to survey participation. The full model [ $F(29, 1136)=35.86, p=0.00$ ] included all the variables determined to be important predictors of survey interview participation for each factor (see table 4).

**Respondent Characteristics.** Most of the respondent characteristics were shown to be significant predictors of survey participation. Age was the only variable that was not significant. The results showed that students who were male ( $OR=.76, p=0.00$ ) or single ( $OR=.72, p=0.00$ ) were less likely to participate in the student interview than females and married students (see Table 4). However, students with military service ( $OR=1.32, p=0.00$ ), and those who attended college full time ( $OR=1.57, p=0.00$ ) were more likely than their respective peers to complete the survey. In addition, Hispanics and Asians were more likely to participate in the survey than Whites when controlling for all other variables ( $OR=1.15, p=0.02$ ;  $OR=1.47, p=0.00$ , respectively), but American Indians were less likely to do so ( $OR=0.69, p=0.00$ ). Students in higher class levels were more likely than first-year students to participate in the survey (minimum  $OR=1.67, p=0.00$ ). Bachelor’s degree students were more likely than those pursuing a certificate or associate’s degree (i.e., less than a bachelor’s degree) to complete the survey ( $OR=1.44, p=0.00$ ).

**Social and Technological Environments.** The model showed that several variables pertaining to the institution were shown not to be significant predictors of the outcome when included in the model containing all variables. These variables include enrollment, percent of undergraduates living on campus, and whether an institution provided a campus e-mail account to all its students. However, institution type, urbanicity, and number of computers on campus were significant predictors in the model. Students attending public 4-year and private not-for-profit 4-year institutions were less likely to participate in the survey than students attending public 2-year institutions (each  $OR=.56, p=0.00$ ). Students in large urban areas were less likely to participate than students attending schools in non-urban areas ( $OR=.86, p=0.04$ ). The number of computers on campus had no substantial impact on the odds of completing a survey ( $OR=1.01, p=0.01$ ).

**Survey Design.** The only variable included in the model pertained to applying for and receiving aid. This variable was significant. Students who applied for aid and did not receive it were more likely to participate in the survey than those who received aid ( $OR=1.65, p=0.00$ ).

In general, the best-fitting overall model built to predict student interview participation contained mostly respondent characteristics variables, plus variables from the social and technological environments, and the survey design features. These factors worked together to impact survey participation.

Table 4  
 Logistic regression analysis: Student interview and all relevant predictors

	<u>B</u>	<u>SE</u>	Odds Ratio	<u>t</u>
<b>Respondent characteristics</b>				
Male	-.274	.027	0.76	-10.25 *
Age	.002	.002	1.00	0.66
Race (reference=White)				
African American	-.054	.052	0.95	-1.03
Hispanic	.142	.060	1.15	2.35 *
Asian	.386	.078	1.47	4.96 *
American Indian/Alaska Native	-.376	.119	0.69	-3.17 *
Other/multiple races	1.144	.124	3.14	9.23 *
Single	-.330	.056	0.72	-5.89 *
Military service	.278	.071	1.32	3.9 *
Full-time attendance	.454	.035	1.57	12.82 *
Undergraduate level (reference=1 <sup>st</sup> year undergraduate)				
2nd year undergraduate	.511	.051	1.67	10.01 *
3rd year undergraduate	.693	.047	2.00	14.82 *
4th or 5th year undergraduate	.522	.048	1.69	10.76 *
Unclassified undergraduates	.105	.173	1.11	0.60
Degree program (reference=Less than bachelor's)				
Bachelor's degree	.364	.089	1.44	4.06 *
No degree	.531	.080	1.70	6.62 *
<b>Social environment</b>				
Institution type (reference=Public 2-year)				
Public 4-year	-.586	.098	0.56	-5.99 *
Private not-for-profit 4-year	-.576	.109	0.56	-5.29 *
Private for-profit 2-years or higher	-.277	.080	0.76	-3.46 *
Private not-for-profit 2-year	-.276	.143	0.76	-1.93
Urbanicity (reference=Non-urban area)				
Mid-sized urban area	-.080	.067	0.92	-1.20
Large urban area	-.152	.073	0.86	-2.09 *
Unknown	.095	.098	1.10	0.97
Enrollment (per 1000)	.003	.003	1.00	1.01
Percent living on campus (per 5 percent)	.006	.007	1.01	0.90
<b>Technological environment</b>				
Number of computers on campus (per 100)	.005	.002	1.01	2.61 *
Campus e-mail account provided	.087	.079	1.09	1.10
<b>Survey design</b>				
Application for and receipt of student aid (reference= Applied for and received aid)				
Applied for & didn't receive aid	.501	.067	1.65	7.48 *
Didn't apply & didn't receive aid	-.301	.046	0.74	-6.58 *

\*p < .05

**Crossvalidation.** In order to verify the predictive ability of the final model, classification tables were generated for the initial analysis group, upon which the logistic regression analyses were performed, and for the holdout group. The classification table showed the fit between the actual and predicted group membership. The logistic regression model was able to correctly classify 70.3% of the analysis group. The model was equally accurate when applied to the holdout group, correctly classifying 71.7%. The model did an excellent job of correctly classifying student interview participants, but a poor job of classifying nonparticipants. The model was able to correctly predict 98% of the student interview participants for both samples but around 5% of the student interview nonparticipants. Cases were over classified into the largest group. Hosmer and Lemeshow (1989) stated, "Classification is sensitive to the relative sizes of the two component groups and will always favor classification into the larger group, a fact that is independent of the fit of the model" (p. 147). However, since the goal

of the classification table for this study was to determine if the model would classify the holdout sample in the same way as the analysis sample, the classification tables have proven that they do. Hence, no further cross validation testing is needed.

In general, the overall model built to predict student interview participation did an adequate job of predicting participation, Most of the respondent characteristic variables were significant predictors of the outcome, and at least one variable for the other factors were significant predictors

### **Mode of Completion Outcome**

An overall logistic regression model was developed to answer how respondent's characteristics, the institution's social and technological environments, and the survey's design features related to mode of completion (Web vs. telephone). The full model [F(52, 983)=86.78, p=0.00] contained all the predictors determined to be important predictors of mode of completion (see Table 5).

**Respondent Characteristics.** Most of the respondent characteristics were shown to be significant predictors of mode of completion. However, the following variables were not significant: disability status, military service, full-time attendance status, and income. The results showed that being a U.S. citizen (OR=.66, p=0.00), male (OR=.85, p=0.00), single (OR=.88, p=0.04), and having dependent children (OR=.75, p=0.00) were negatively related to completing the interview via the Web. African American (OR=.50, p=0.00) and Hispanic (OR=.80, p=0.01) students were less likely to complete the survey via the Web than White students. However, Asian students (OR=1.29, p=0.0) were more likely to complete the Web instrument than White students. In addition, students with higher GPAs (OR=1.27, p=0.00), who were a science/technology major (OR=1.20, p=0.01), or who had taken distance education classes (OR=1.72, p=0.00) were more likely to use the Web mode. Students working on a bachelor's degree were more likely to complete the survey via the Web compared to those pursuing a certificate or associate's degree (OR=1.33, p=0.00). Seniors were more likely to do the Web survey than freshmen (min OR=1.22, p=0.00). In addition, students whose parents had some college or higher were more likely than their peers whose parents highest education was high school or less were more likely to use the Web survey (min OR=1.14, p=0.046).

**Social and Technological Environments.** The model showed that of the variables pertaining to the social environment, the student-faculty ratio and institution type were not significant predictors. More specifically, students attending public 2-year institutions were just as likely as students at public and private not-for-profit 4-year institutions to use the Web mode. While enrollment size and percent living on campus were significant predictors, they had no substantial impact on the odds of completing the survey via the Web (min OR=1.02, p=0.04). Hence, the social environment had no real impact on completing the survey via the Web.

For the technological environment, the model showed that whether or not an institution offered distance education classes or required computer ownership were not significant predictors of Web completion. However, students attending institutions that provided campus e-mail accounts (OR=1.2, p=0.00) were more likely to complete the Web survey. The number of computers on campus and the percentage of science & technology degrees awarded by the institution had no substantial impact on Web completion.

**Survey Design.** Several variables pertaining to survey design features were included in the model. The most overwhelming predictor of completing the survey via the Web was receiving an incentive. The odds of completing the Web survey were 20 times higher for those who received an incentive than for those who did not get an incentive (p=0.00). This result was expected because of the design of the survey. Incentives were offered for early response and for refusal conversion. Sample members received an early response incentive if they completed the survey via the Web within four weeks of the initial survey request. If these students did not respond by the end of the fourth week, follow-up calls were made to obtain responses via a telephone interview. Respondents did not get an incentive during this period. Near the end of NPSAS, another incentive was offered in an attempt to convert nonrespondents (i.e., those who were hard to reach or initially refused to participate). The results also showed that students who applied for, but didn't receive aid were less likely than those who received aid to participate in the study via the Web (OR=.85). However, government assistance was not found to be a significant predictor of mode of completion.

Table 5  
 Logistic regression analysis: Mode of completion and all relevant predictors

	<u>B</u>	<u>SE</u>	Odds Ratio	<u>t</u>
<b>Respondent characteristics</b>				
Male	-.162	.039	0.85	-4.18 *
Age	-.007	.003	0.99	-1.99
Race (reference=White)				
African American	-.688	.062	0.50	-11.15 *
Hispanic	-.219	.078	0.80	-2.80 *
Asian	.251	.086	1.29	2.91 *
American Indian/Alaska Native	.016	.203	1.02	0.08
Other/multiple races	-.156	.103	0.86	-1.51
Single	-.132	.065	0.88	-2.04 *
Disability	.089	.068	1.09	1.31
Dependent Children	-.291	.067	0.75	-4.38 *
Military service	-.040	.131	0.96	-0.31
Community service hours (per 5 hours)	-.029	.007	0.97	-4.02 *
College GPA	.237	.033	1.27	7.23 *
Full-time attendance	-.012	.043	0.99	-0.29
Undergraduate level (reference=1 <sup>st</sup> year undergraduate)				
2nd year undergraduate	.089	.054	1.09	1.64
3rd year undergraduate	.038	.065	1.04	0.58
4th or 5th year undergraduate	.197	.066	1.22	2.98 *
Unclassified undergraduates	.405	.155	1.50	2.60 *
Degree program (reference=Less than bachelor's)				
Bachelor's degree	.282	.010	1.33	2.83 *
No degree	-.167	.093	0.85	-1.80
Highest level of education ever expected to complete (reference=Less than bachelor's)				
Bachelor's degree	-.119	.081	0.89	-1.47
Master's degree	-.109	.089	0.90	-1.22
Doctorate/first-professional degree	.232	.094	1.26	2.45 *
Parents' highest degree (reference=High school or less)				
Some college	.471	.054	1.60	8.73 *
Bachelor's degree	.134	.067	1.14	2.00 *
Graduate degree	.144	.066	1.16	2.20 *
Unknown	.880	.182	2.41	4.84 *
Job role (reference=Primarily student working to meet expenses)				
Primarily employee going to school	.038	.057	1.04	0.65
No job	-.193	.067	0.82	-2.75 *
Hours worked (per 10 hours)	-.079	.019	0.92	-4.26 *
Income (per \$25,000)	-.016	.013	0.98	-1.24
Distance from home (per 200 miles)	.015	.006	1.02	8.23 *
Science/technology major	.185	.069	1.20	2.66 *

Table 5 (continued)

	<u>B</u>	<u>SE</u>	Odds Ratio	<u>t</u>
Taken distance education	.545	.066	1.72	8.23 *
Citizenship	-.414	.102	0.66	-4.08 *
<b>Social environment</b>				
Institution type (reference=Public 2-year)				
Public 4-year	.012	.115	1.01	0.11
Private not-for-profit 4-year	.215	.141	1.24	1.52
Private for-profit 2-years or higher	-.200	.120	0.82	-1.66
Private not-for-profit 2-year	-.149	.201	0.86	-0.74
Enrollment (per 1,000 students)	.015	.003	1.02	4.96 *
Percent living on campus (per 5 percent)	.021	.010	1.02	2.01 *
Student-faculty ratio (per 10 students)	-.000	.017	1.00	-0.02
<b>Technological environment</b>				
Institution offers distance education classes	-.116	.075	0.89	-1.55
Percent of science/tech degrees (per 5 percent)	.022	.010	1.02	2.28 *
Number of computers on campus (per 100 PCs)	.005	.002	1.01	2.57 *
Computer ownership required	-.080	.167	0.92	-0.48
Campus e-mail account provided	.202	.075	1.22	2.69 *
<b>Survey design</b>				
Contact calls	-.067	.023	0.94	-25.56 *
Application for and receipt of student aid (reference= Applied for and received aid)				
Applied for & didn't receive aid	-.159	.070	0.85	-2.27 *
Didn't apply & didn't receive aid	-.114	.063	0.89	-1.82
Received government assistance	-.027	.120	0.97	-0.22
Received incentive	3.01	.059	20.2	50.66 *

\*p < .05.

**Crossvalidation.** The cross-validation results showed that the logistic regression model was able to correctly classify 78% of the analysis group and the holdout group. Hence, the predictive ability of the model was similar for both of the samples. For both groups, the model was better at classifying those who were Web interview respondents. For the analysis group, 81% of Web respondents were classified accurately compared to 76% of telephone respondents. These rates were similar for the holdout group.

In general, the overall model built to predict mode of completion did an adequate job of predicting Web completion. Most of the variables from each factor were significant predictors of the outcome. Incentives were the most important predictor of Web completion.

## Discussion

This study examined how selected factors were related to whether a student participated in a survey and the mode of completion chosen using a modified version of Groves and Couper's conceptual framework on survey participation (Groves & Couper, 1998; Vehovar et al 2002). The four factors examined were respondent characteristics, the social and technological environments, and survey design features. These factors were brought together to determine how they were related to survey participation overall and by mode of completion for undergraduate college students. The results showed that all factors were important predictors of the outcomes, but that respondent characteristics and incentives were the most important predictors.

The results showed that incentives played a major role in getting students to respond to the student interview via Web, and capturing hard-to-reach and reluctant students. Incentives were also important to survey cooperation. The incentive variable was based on whether or not a person received an incentive for completing the survey. However, since all of the student interview nonparticipants did not receive an incentive because they never responded that variable was not included in the

logistic regression model to examine overall survey participation. About 66 percent of the participants received an incentive, which suggests that incentives played a role in gaining cooperation.

This finding is consistent with literature. Incentives have long been a proven method of increasing survey participation (Grove & Couper, 1998; James & Bolstein, 1992; Shettle & Mooney, 1999). Goyder (1987) found that no matter what other survey design features are in place, incentives always work. This was true for this study. Incentives outweighed the influence of other predictors on survey participation by mode of completion. They were more important than the respondent characteristics and the institution's social and technological environments.

Incentives were also important in directing those who decided to participate in the survey to use a particular mode. About 46 percent of respondents completed the survey via the Web. However, 88 percent of Web respondents received an incentive compared to 48 percent of telephone respondents. Hence, researchers would be able to emphasize the mode they want respondents to use by affixing an incentive to the mode.

Aside from incentives, respondent characteristics were the most important set of variables. This study found that civic participation was important to overall survey participation. For this study, the number of community service hours performed and whether the respondent had ever served in the military were used to determine civic activity. The results showed that for overall student interview participation, military service<sup>1</sup> was an important predictor of participation. For mode of completion, community service hours performed was a negative predictor of completing the survey via the Web and military service was not a significant predictor.

The finding of the importance of community service in survey participation is consistent with the research literature. Groves and Couper (1998) found that households with military members were more likely to cooperate in a survey. This finding may be due their belief in the legitimacy of the federal government as a sponsor of the survey. Porter and Whitcomb (2003) also found that socially engaged students were more likely to participate in a survey than their peers. In addition, Groves, Singer, and Corning (2000) found that community service was important to survey participation and that the impact of incentives was smaller for those involved in community service. Hence, many individuals did not need an incentive in order to decide to participate. Kropf and Blair (2005) had similar findings. They found that for nonresponse follow-up, while incentives were important in gaining the cooperation of those who refused, that letters stressing the good of the survey to society also help gain participation.

In general, the results suggest that while civic duty may motivate a person to decide to participate in a survey, it may require additional contact to get the person to do the survey. However, community service hours may get in the way of completing a self-administered interview.

Two respondent characteristics variables were found to be significant predictors in both outcomes: gender and degree program. The results showed that males were less likely than females to participate in the student interview; and if they did participate, they were less likely to do the self-administered Web-based instrument.

Research has consistently shown that a higher proportion of females complete surveys regardless of the mode (Collins et al., 2003; Sax, 2003). However there is mixed research on whether females are more likely than males to complete a Web survey. Researchers have found that in the general population, Web completers are younger, male, and more affluent (Palmquist & Stueve, 1996). Studies involving college students have shown mixed results. While Smith and Leigh (1997) found males were more likely than females to complete the Web survey, Underwood et al (2000) found that college women were more likely to respond to any mode at a higher rate than men, including Web. Sax (2003) also found the same result in a study of first-time, full-time freshmen. When students were given the choice of a paper-and-pencil survey or a Web-based survey, men were more likely than women to choose the Web survey.

Degree program was also an important predictor of survey participation. Students pursuing a bachelor's degree were more likely than those pursuing a certificate or associate's degree to participate in a student interview. This is contrary to findings by Groves and Couper (1998). They found that households in which the key respondent had a lower education were more likely to respond. Also, Laurie, Smith, and Scott (1999) found that individuals with less education were more likely to

---

<sup>1</sup> Military service indicator was available for both student interview participants and nonparticipants. Community service hours performed was available only for student interview participants.

continue to participate in panel studies. Perhaps examining data on the highest degree of the parents would have led to findings consistent with the literature.

While students with a higher degree pursuit were more likely to participate in a survey, they were also more likely to do it by Web. This study found that the odds of completing a Web survey were 1.3 times higher for those pursuing a bachelor's degree than for those pursuing a certificate or associate's degree. Hence, students in bachelor's degree programs may have the necessary technical knowledge and skills or attend institutions that provide sufficient technological support that makes it easier for students to respond to the survey early via the Web.

This study examined several variables pertaining to the social environment of the students' institutions. The results showed that the institution type and urbanicity were important in whether a student participated in the interview. It was expected that the urbanicity of the school would be a predictor of survey participation. Research literature has shown that variables such as perceived legitimacy of social institutions and social cohesion may impact a person's sense of social responsibility and thus, his or her decision to participate in a survey (Goyder, 1987; Groves, Cialdini, & Couper, 1992). Also, urbanicity consistently has been found to influence survey participation. Research shows that people in large cities tend to be less willing to complete a survey than those in rural areas (Brehm, 1993; Goyder, 1987). The results of this study were consistent with the literature. The results showed that students attending schools in non-urban areas were more likely than students in large urban areas to participate in a survey. However, the variables pertaining to student-faculty ratio and percentage of undergraduate living on campus were not significant predictors of student interview participation.

Vehovar et al. (2002) hypothesized that one's computer literacy and orientation towards computer usage were important influences on the mode used to participate in surveys and that they were more important to cooperation than the respondent's demographics. This hypothesis was supported by this study. The results showed that students with a science/technology major or who had ever taken distance education classes, or attended institutions that provided campus e-mail accounts were more likely than their respective peers to completed a Web survey. This suggests that the students' technical experience and the technological environment of their institutions were important in predicting Web completion. Sills and Song (2002) in a Web-based survey of international undergraduate and graduate students found that computer science and engineering majors were more likely than other majors to respond to their survey because of the importance of technology to their major. Yun and Trumbo (2000) found that those who responded by e-mail and Web were more connected to the Web, used the Web more frequently in their daily lives, and had a higher level of education. Also, Kwak and Radler (2002) found that students who quickly adapted to new technology, spent more time on the Internet and had advanced Internet services were more likely to do the Web survey.

Vehovar et al. (2002) also indicate that a person's technological environment may impact a person's ability to respond via the Web. If a person has inadequate computer equipment and a slow Internet connection, then length of time to complete the interview increases thus frustrating potential respondents. These issues make it difficult or impossible to complete the survey. Detailed information on the technical experience and equipment of subjects is not generally known about sample members, but it is information that can be requested from those who complete the Web-based instrument. Unfortunately, no information was available on the nonrespondents who had tried to complete the Web survey, but gave up because of technical problem or how many telephone respondents wanted to do the Web instrument early, but had experienced problems.

### **Limitations**

This study used data from a national survey of postsecondary students; therefore, the results may not be directly generalizable to different populations. Students may behave differently than other population groups in their ability and motivation to participate. Also, they may adapt quickly to new technology, such as the Internet, which makes it a viable and valuable option for collecting data from these individuals.

Another limitation of this study was the lack of information on the interviewers who conducted the telephone interviews. Survey research is clear on the importance of interviewers in gaining survey participation. This study was not able to measure the effect of the interviewer on how and when respondents participate. However, it was able to indirectly observe the interviewer-respondent interaction by looking at whether responses to interviewer-administered surveys differed from responses to self-administered surveys, especially in regards to sensitive questions.

This study provided a profile of those who did not respond to the student interview. The main data set contained information on sample members who did not complete a student interview. While data were available for the majority of these nonparticipating sample members, the results of this study may not be generalizable to all nonresponding sample members.

## **Implications**

The results of this study can help inform research on survey participation and lead to a better understanding of how selected factors influence cooperation. Specifically, the results give survey administrators additional guidance on the explicit factors and variables that are associated with survey participation. Results show whether or not the factors behave differently depending on how and when an individual responds. From these results, researchers may be able to reduce nonresponse and overall costs in their studies by tailoring features of the survey to different groups of sample members given their social and technological environment.

The information learned in this study provides guidance to survey researchers on where they should initiate their contacting efforts. They may gain a better idea of how to allocate their resources. If certain groups are more difficult than others to contact and gain their participation, researchers know that they will need to start engaging these people earlier in the survey process. For example, they could start contacting these hard-to-reach people early or consider extending the data collection cycle with the goal of capturing more of these types of nonrespondents.

## **Further Research**

This study examined the impact of selected factors on overall survey participation and by mode of completion. However, researchers will likely have different variables comprising the factors based on their sample. For example, the social environment in this study uses variables pertaining to the institution the student is attending. Most of these variables were not determined to be significant predictors of survey participation. Future research should look at whether the social environment of where the student attended high school would have been more important to survey participation.

Researchers should collect more comprehensive information on the technical experience or computer literacy of all of its respondents regardless of the mode they used to complete the survey. Only a few variables were available for this study, but looking at additional variables might uncover more specific reasons for why students do not respond by Web. This study never had information on the technical equipment (including Internet connection) available to students, nor whether they ever experienced technical problems; thus, ignoring any impact that these issues might have on nonresponse.

Future research should experiment with different aspects of the survey design. For example, the effectiveness and costs of telephone prompting should be examined to determine if prompting leads to earlier response with and without an incentive and whether gaining these responses would actually be cost effective. This will provide better information on whether using a combination of personal contact and incentives are both needed for high response rates. These procedures will be examined in the next cycle of the National Postsecondary Student Aid Study.

The impact of using large incentives in longitudinal studies requires further examination. Offering an incentive in the base year essentially sets the minimum incentive that can be used for subsequent follow-up surveys. In addition, if different levels of incentives are offered through out the base year data collection cycle, respondents might alter their behavior to get the highest incentive. For example, if a small incentive is offered at the beginning of data collection and a higher incentive is offered later, in subsequent follow-up surveys subjects may delay responding until a higher incentive is offered. Thus requiring more contact calls and higher incentive amounts to get more response. In addition, subjects may feel that the incentive amount should increase every time they participate. Hence, the cost effectiveness and dynamics of using large incentives may be different for longitudinal studies.

Most research only examines the background characteristics of respondents. Hence, researchers have a one-dimensional portrait of who responds to their surveys. Future research should examine the personality characteristics of respondent to really gain an understanding of the internal characteristics that lead to participation.

## **Disclaimer**

This paper is intended to promote the exchange of ideas among researchers and policy makers. The views expressed in it are part of ongoing research and analysis and do not necessarily reflect the position of the U.S. Department of Education.

## **References**

- Brehm, J. (1990). Opinion surveys and political representation (Doctoral dissertation, The University of Michigan, 1990). *Dissertation Abstracts International*, 5.2, 282.

- Collins, R. L., Ellickson, P. L., Hays, R. D., & McCaffrey, D. F. (2000). Effects of incentive size and timing on response rates to a follow-up wave of a longitudinal mailed survey. *Evaluation Review*, 24, 347-363.
- Cominole, M., Siegel, P., Riccobono, J., Dudley, K., Franklin, J., Roe, D., Gilligan, T., & Kuhr, B. (2006). *2004 National Postsecondary Student Aid Study (NPSAS:04) full scale methodology report*. Washington, DC: National Center for Education Statistics.
- Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, 64, 464-494.
- De Heer, W. F. (1999). International response trends: results of an international survey. *Journal of Official Statistics*, 15, 12-142.
- Fowler, F. J. (2002). *Survey research methods*. Thousand Oaks, CA: Sage Publications.
- Goyder, J. C. (1987). *The silent minority: Nonrespondents on sample surveys*. Boulder: Westview Press.
- Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understand the decision to participate in a survey. *Public Opinion Quarterly*, 59, 475-495.
- Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: Wiley & Sons.
- Groves, R. M., Singer E., & Corning, A. D. (2000). Leverage-saliency theory or survey participation: Description and an illustration. *Public Opinion Quarterly*, 64, 299-308.
- James, J. M., & Bolstein, R. (1992). Large monetary incentives and their effect on mail survey response rates. *Public Opinion Quarterly*, 56, 442-453.
- Kropf, M. E., & Blair, J. (2005). Eliciting survey cooperation: Incentives, self-interest, and norms of cooperation. *Evaluation Review*, 29, 559-575.
- Kwak, N., & Radler, B. (2002). A comparison between mail and web surveys: Response pattern, respondent profile, and data quality. *Journal of Official Statistics*, 18(2), 257-274.
- Laurie, H., Smith, R., & Scott, L. (1999). Strategies for reducing nonresponse in a longitudinal panel survey. *Journal of Official Statistics*, 15, 269-282.
- Palmquist, J., & Stueve, A. (1996). Stay plugged into new opportunities. *Market Research*, 8, 13-15.
- Porter, S. R., & Whitcomb, M. E. (2003). The impact of lottery incentives on student survey response rates. *Research in Higher Education*, 44, 389-407.
- Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education*, 44, 409-432.
- Shettle, C., & Mooney, G. (1999). Monetary incentives in U.S. government surveys. *Journal of Official Statistics*, 15(2), 231-250.
- Sills, S. J., & Song, C. (2002). Innovations in survey research: An application of web-based surveys. *Social Science Computer Review*, 20, 22-30.
- Smith, M. A., & Leigh, B. (1997). Virtual subjects: Using the Internet as an alternative source of subjects and research environment. *Behavior Research Methods, Instruments, & Computers*, 29, 496-505.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn and Bacon.

- Underwood, D., Kim, H., & Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the 40<sup>th</sup> Annual Forum of the Association for Institutional Research, Cincinnati, OH, May 21-24, 2000.
- Vehovar, V., Batagelj, Z., Lozar Manfreda, K., & Zaletel, M. (2002). Nonresponse in web surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 229-242). New York: Wiley & Sons.
- Yun, G. W., and Trumbo, C. W. (2000). Comparative response to a survey executed by post, e-mail, & web form. *Journal of Computer-Mediated Communication*, 6(1). Retrieved October 27, 2002, from <http://www.ascusc.org/jcmc/vol6/issue1/yun.html>.