Using Web Surveys to Reach Community College Students: An Analysis of Response Rates and Response Bias

Linda J. Sax, University of California Los Angeles
Shannon K. Gilmartin, University of California Los Angeles
Jenny J. Lee, University of California Los Angeles
Linda S. Hagedorn, University of Southern California

Research paper accepted for presentation at the annual meeting of the Association of Institutional Research (AIR)
May 2003

Introduction

As online surveys continue to capture the attention of institutional researchers, several questions about this new medium of data collection invariably surface, especially when online instruments are compared to traditional paper instruments. First is the issue of response rates. Do online surveys yield higher rates of response than do paper surveys? By which method can institutional researchers collect the most data? Second is the issue of nonresponse bias, or differences between survey respondents and nonrespondents (demographically, attitudinally, or otherwise). Is the nonresponse bias characteristic of online surveys similar to or different from that of paper surveys? Do online surveys steer data collection toward new (and possibly less skewed) respondent pools, or do they reproduce the respondent bias found in paper surveys? Still a third issue is response bias. That is, are there differences between online survey responses and paper survey responses, despite identical survey items? Close analysis of response bias is particularly critical when surveys are distributed as paper and electronic forms within a single administration, and clarifies further the methodological implications of data collection via the Internet.

With these issues in mind, the present study is designed to examine response rates, nonresponse bias, and response bias across two groups of community college students: those who received a district-wide follow-up survey of their college experiences via email, and those who received this survey by standard mail. The results of this study not only paint a clearer picture of differences and similarities between online surveys and paper surveys, but also inform efforts to equate online survey data with paper survey data in a single, mixed-mode administration. Further, by focusing this study on community

college students, we stand to learn more about a group of students who are notoriously difficult to locate and who historically have had lower-than-average survey participation rates.

Background of the Study

Though the body of literature on response rates, nonresponse bias, and response bias among online and paper surveys is not extensive, several studies in this burgeoning area of research merit discussion. These studies are reviewed below, following brief comments on the advantages and disadvantages of online data collection.

Online Surveys

Notwithstanding the increasing popularity of and reliance on the Internet, the use of online surveys for institutional research carries with it many challenges (Hamilton, 1999; Goree & Marszalek, 1995). One concern is that of access. Goree and Marszalek (1995) warn that access to computers is not equal—those with the most power in society enjoy the broadest access to new and different forms of technology, while those with the least power find themselves on the margins of the Information Age. Ebo (1998) agrees that disadvantaged or underrepresented populations have insufficient access to the resources of cyberspace, a finding also noted for college freshmen (Sax, Ceja, & Teranishi, 2001). Thus, the sample of individuals who respond to an online survey may not be entirely representative of the study's intended population. This reality must be addressed before generalizing online survey data to a larger group.

Other methodological challenges include concerns about data security, which could lead to nonresponse (Smith, 1997), and human subjects guidelines that are unclear about online research (Hamilton, 1999). However, the appeal of online surveys is

indisputable: completing a questionnaire on the Internet is more cost-efficient for many institutions and more convenient for many "computer savvy" subjects like college students (Carini, Hayek, Kuh, Kennedy, & Ouimet, 2003).

Response Rates, Nonresponse Bias, and Response Bias

Relatively few studies examine response rates, nonresponse bias, and response bias by electronic and paper modes of survey administration, although the findings of those that do cast doubt on methodological strengths of online data collection relative to more traditional formats. In a comparison of paper surveys to online surveys, Matz (1999) observed little difference in types of responses and respondent demographics by survey format. However, the paper survey yielded a higher rate of response than did the online survey. So too observed Underwood, Kim, and Matier (2000): among the college students in their study, rates of response were higher among those who received a paper survey than among those who received a survey by email. The authors also noted that response rates of women were higher than those of men regardless of survey format, as was true of the White, Asian American, and international students in their sample.

More recently, Sax, Gilmartin, and Bryant (forthcoming) randomly assigned a sample of nearly 5,000 college students at 14 four-year institutions to one of three survey administration groups: (1) paper survey only, (2) paper survey with the option to complete the questionnaire online, and (3) online survey only. The authors found that the rate of response was highest among students who received the paper survey with online option, and was lowest among students who received the online version of the instrument only. Like the students in Underwood, Kim, and Matier's (2000) study, women responded in greater numbers than did men; response rates also were highest among

Asian American students, as compared to other racial/ethnic groups. In terms of nonresponse bias, being female increased the odds of response across all administration groups. Other predictors varied by group, but these were few in number, and did not yield enough evidence to conclude that nonrespondents to online surveys were substantially different than were those to paper surveys. Relatedly, Carini, Hayek, Kuh, Kennedy, and Ouimet (2003) observed that survey format (online versus paper) did not appreciably impact responses among a national sample of college students, although subjects tended to respond more favorably to some questions when completing the questionnaire online.

Objectives

Building on the work of Sax, Gilmartin, and Bryant (forthcoming), Carini, Hayek, Kuh, Kennedy, and Ouimet (2003) and others, the present study is designed to compare community college students who received a follow-up survey of their college experiences via email to community college students who received this survey via standard mail. The study addresses three questions:

1. Do response rates differ by mode of survey administration?
2. Do the predictors of response differ by mode of survey administration? (nonresponse bias)
3. Are item-by-item responses to online surveys different than item-by-item responses to paper surveys? (response bias)

The goal of this study is to determine if different modes of survey administration yield substantively similar survey data. Similar data imply that online surveys are methodologically equivalent to paper surveys, but do little to reduce traditional biases in the respondent pool and types of survey responses. Disparate data imply that online

surveys are not equivalent to paper surveys, but might increase the representation of certain groups who otherwise might not respond to the survey itself.

Methodology

Sample

Data for this study draw from the 2001 "Transfer and Retention of Urban Community College Students " (TRUCCS) baseline survey and the 2002 TRUCCS follow-up survey. Funded by the U.S. Department of Education, TRUCCS is designed to examine the myriad factors that influence persistence, transfer, and achievement among students enrolled in the Los Angeles Community College District (LACCD). In keeping with this goal, the TRUCCS surveys include a range of questions about students' family life, employment history, classroom experiences, educational goals, and personal values. TRUCCS represents a collaboration between the University of Southern California (USC), the University of California Los Angeles (UCLA), and LACCD.

In Spring 2001, the TRUCCS baseline survey was administered to a stratified sample of 5,001 students at nine LACCD campuses. Members of the TRUCCS project team at USC and UCLA distributed paper surveys in 241 classrooms; students were instructed to complete the survey as part of a larger study of community college student experiences and educational pursuits. To maximize variation in the sample, a proportionate mix of remedial, standard, vocational, and gateway courses were selected as sites for survey administration. Subsequent analyses confirmed that students who were enrolled in these courses resembled the larger LACCD population in terms of race, ethnicity, age, and primary language.

So to examine these students' experiences longitudinally, subjects who completed the TRUCCS baseline survey were mailed or emailed the TRUCCS follow-up survey in Winter and Spring 2002, or approximately one year after the baseline survey was distributed. Follow-up surveys were administered by mail or email depending on the type of contact information that students provided on the baseline survey. In other words, surveys were sent via email to students who listed a valid email address, and via standard mail to students who did not list a valid email address, or did not list an email address at all (the drawbacks associated with this nonrandom assignment of administration mode are described in the results and discussion sections). Second and third waves of the survey were distributed to first-wave nonrespondents, sometimes via email and standard mail if students provided both types of contact information. However, the sample for the present study is comprised of 4,387 students who received the 2002 TRUCCS follow-up survey as a paper or electronic questionnaire (and for those who returned the follow-up survey, via the mode in which they were initially contacted). The remaining 614 students either 1) received the follow-up survey as a paper and electronic instrument, 2) did not provide any valid address at which to contact them for the follow-up study, or 3) were contacted by telephone in the final months of data collection to maximize overall response. These students were excluded from this dataset in order to calculate more accurate rates of response and "cleaner" estimates of bias.

Research Methods

As part of this study, three sets of analyses were conducted:

- Descriptive analyses to calculate response rates by mode of follow-up survey administration, sex, and race/ethnicity. These included frequencies and crosstabulations.
- Logistic regression analyses to explore nonresponse bias by mode of follow-up survey administration. These analyses compared the predictors of response to the follow-up survey across two groups: students who received the survey as a paper form (Group A), and students who received the survey as an electronic form (Group B). A total of four logistic regression analyses were performed. The first two analyses regressed each dependent variable ("Paper Response to the Follow-Up Survey," for students in Group A, and "Email Response to the Follow-Up Survey," for students in Group B) on 29 independent variables using stepwise procedures ($p<.05$). Next, predictors of each dependent variable were pooled and force-entered into a second set of logistic regressions in order to compare the same predictors across each group. Missing values on independent variables were replaced with the mean of each variable by administration group (missing values for any given variable did not exceed 15 percent of the sample).
- Independent sample t-tests to determine response bias by mode of follow-up survey administration. Here, mean responses to 113 items on the follow-up survey were compared across two groups: students who submitted the paper form (Group A) and students who submitted the electronic form (Group B). Those items with statistically significant mean differences ($p<.01$) between Groups A and B were flagged for discussion.

Variables for the Logistic Regression Analyses

As noted above, the dependent variables for these analyses were "Paper Response to the Follow-Up Survey" (1= "no," 2= "yes") for students in Group A, and "Email Response to the Follow-Up Survey" (1= "no," 2= "yes") for students in Group B. Based on findings from previous studies of online and paper surveys (Matz, 1999; Sax, Gilmartin, & Bryant, forthcoming; Underwood, Kim, & Matier, 2000), a total of 29 independent variables were selected for the stepwise logistic regression analyses, all of which drew from the TRUCCS baseline dataset. These included race/ethnicity, sex, age, average income, plans to attend the same college next semester, number of other colleges/universities attended, and degree aspirations. Hours per week spent on campus, doing housework or childcare, and working at a job also were included in these analyses, as were students' average grades in high school, level of math preparation, reasons for attending their current college, and length of commute to campus. Two measures of disability, one measure of computer ownership, one measure of English speaking ability, and one measure of place of residence were tested in these analyses as well.

The remaining four variables were factors derived from two principal components factor analyses of 71 items on the TRUCCS baseline survey (each factor analysis used varimax rotation techniques, and items with factor loadings of .40 or below were dropped from these analyses to maximize reliability). These included: 1) "Academic involvement: Interaction with instructors/academic counselors," a five-item factor that measures how often respondents interacted with instructors and counselors in the past week; 2) "Academic involvement: Studying with others," a five-item factor that measures how often respondents interacted with other students on academic matters in the past week; 3)

"Views: Determined and confident," a nine-item factor that measures the degree to which respondents are committed to doing well in school and achieving their goals; and 4) "Positive attitude towards school," a two-item factor that measures the degree to which respondents enjoy and feel comfortable with their coursework. Appendix A provides a complete list of all independent variables and coding schemes; Appendix B describes the items that comprise each factor, and lists factor loadings and Cronbach's alpha values.

Results

Response Rates

As shown in Table 1, the average response rate across both modes of follow-up survey administration was 21.3 percent. This rate is surprisingly similar to the response rate of 21.5 percent reported by Sax, Gilmartin, and Bryant (forthcoming), who conducted a one-year follow-up study of college students at four-year campuses. However, Sax, Gilmartin, and Bryant noted that response rates were lowest among students in their sample who comprised the online-only administration group, whereas response rates were highest among the online-only administration group in the TRUCCS follow-up sample. In fact, response rates for the online-only group were double that of the paper-only group in this study (31.5 percent versus 15.7 percent). This difference likely owes to the point that students who returned a follow-up survey via email were those who had provided the TRUCCS project team with a valid email address on the baseline questionnaire in Spring 2001. Therefore, the TRUCCS study appears to have avoided one of the pitfalls of many online surveys: low response rates due to incorrect or infrequently used email addresses (as discussed in Sax, Gilmartin & Bryant).

Response rates broken out by gender and race/ethnicity are provided in Table 2. Regardless of mode of contact, women displayed higher rates of response than did men, a finding consistent with recent research on gender differences in response to paper and web surveys (Sax, Gilmartin & Bryant, forthcoming; Underwood, Kim & Matier, 2000). Interestingly, the gender gap in response rates is narrower in the email administration group than in the paper administration group, suggesting that online survey administration yields a better gender balance among respondents than does paper survey administration. Underwood, Kim, and Matier also reported a smaller gender gap in web responses as opposed to paper responses. However, online response rates yielded larger gender differences than did paper response rates in Sax, Gilmartin, and Bryant. Clearly, the jury is still out on precisely how Internet surveys affect the gender balance in respondent pools.

Racial/ethnic differences in rates of response produce a different pattern of results for the email and paper samples. White/Caucasian students had the highest response rates regardless of mode of administration (18.8 percent paper and 33.7 percent email). Mexican/Mexican-American students, on the other hand, had the lowest rate of response to the paper survey (14.7 percent), but the highest rate of response to the online survey (tied with White students at 35.4 percent). Asian students—who have demonstrated some of the highest rates of response to paper and email questionnaires (Underwood, Kim & Matier, 2000)—yielded the lowest rate of response to the email survey (24.6 percent) and the second-lowest to the paper survey (15.0 percent). Racial/ethnic variations in response rates partly owe to the "paper-email split" listed in Table 2. This "split" is simply the distribution of the mail-out sample by mode of administration, and is

determined by one criterion: whether or not students provided a valid email address on the baseline questionnaire. These distributions show that African American, Mexican American, and Latino/a students were least likely to have self-selected into the email sample to begin with (i.e., they were least likely to have provided a valid email address on the baseline questionnaire), suggesting that these students may not have had regular access to the Internet or did not rely heavily on email communication. So, while response rates to the online survey would have been lower had all 4,387 students received the follow-up questionnaire via email (i.e., not just those who provided valid email addresses at initial point of contact), these rates probably would have been disproportionately low among African American, Mexican American, and Latino/a students given their lower likelihood of providing valid email addresses on the pretest questionnaire.

Nonresponse Bias

Logistic regression analyses conducted for each group identify predictors of response/nonresponse. As discussed in the methods section, these analyses force-entered a common set of independent variables, each of which had predicted response for either the paper-only or email-only samples in an initial set of logistic regressions. Table 3 provides the logistic regression coefficients, standard errors, and odds ratios for each of the seven independent variables that predicted paper or email response. The logistic regression coefficients signify whether the relationship between a predictor variable and survey response is positive or negative, and give some indication as to the strength of that association. Odds ratios are somewhat different, in that they are centered around 1, with odds ratio greater than 1 indicating that higher scores on a predictor variable increase the

odds of response, and odds ratios less than 1 suggesting that higher scores on a predictor variable decrease the odds of response.

Only two variables significantly predicted survey response for both the paper and email samples: age and average high school grades. Each variable predicted higher rates of response to the paper and online questionnaires. The role of high school grades is not surprising, given prior work that documents better response rates among higher-achieving students (Dey, 1997; Sax, Gilmartin & Bryant, forthcoming). Prior research also has reported age as a predictor of survey response, but typically in general household surveys (e.g., Lepkowski & Couper, 2002). That age is a predictor of survey response among college students is a more novel finding, since most surveys of college students are conducted on samples of students at four-year campuses (wherein the range of student age is fairly narrow). The greater variation in age among community college samples allows us to see this variable in a new light: as compared to their younger peers, older students may be more likely to respond to follow-up questionnaires regardless of whether the survey is sent via standard mail or email.

Two variables positively predict response to the paper survey but are not significant in predicting response to the email survey: being female and being White/Caucasian. These findings are consistent with the results of other standard mail surveys with respect to the role of gender (Dey, 1997; Sax, Gilmartin & Bryant, forthcoming) and race (Dey, 1997; Johnson, O'Rourke, Burris & Owens, 2002).

Three variables are significant predictors of response among the online sample only. The first is the positive effect of being Mexican American. As noted earlier, response rates among Mexican Americans were significantly higher in the email sample

than in the paper sample; these regression results confirm that there is a unique positive effect of being Mexican American on the likelihood of responding to the online questionnaire. In a review of the role of racial/cultural differences in nonresponse, Johnson, O'Rourke, Burris, and Owens (2002) reported no studies in which Mexican American students had higher rates of response. It is unclear at this point why the TRUCCS survey would have produced such unique results, except that Mexican Americans were less likely than most groups to have provided email addresses to begin with. In other words, the unavoidable sampling bias may be one explanation for this group's higher rates of response to the email questionnaire. However, that this scenario does not play out for African American students—the group least likely to have placed themselves in the email respondent pool—is a finding that needs to be explored further.

The remaining two variables are degree aspirations, which positively predict response, and attending college because it was "something to do," which negatively predicts response. In other words, response to the email survey was less likely from students with lower degree aspirations and/or who may view college a way to keep themselves occupied. Although these seem to be logical predictors of response to a survey about college experiences, it is not clear why they would relate solely to survey response via email.

Response Bias

The issue of response bias is addressed in Table 4. Evidence of response bias exists if the item-by-item responses to the mail survey differ significantly from the item-by-item responses to the paper survey. The top portion of the table describes those items for which mean responses were higher ($p < .01$) in the paper group than in the email

group, whereas the bottom portion of the table lists items for which responses were higher among email respondents. In nearly all cases, these mean differences are fairly small (despite their statistical significance), but are important since they shed light on potential limitations to combining responses to paper and email surveys.

Overall, responses to the online questionnaire differed from responses to the paper questionnaire in three primary ways. First, as compared to paper respondents, online respondents were more likely to be employed, to report problems in finding time for school, and to report that job responsibilities were a problem. Second, online respondents were more likely to be unmarried and to be transfer-seeking students, the latter defined by having higher degree aspirations, having applied for admission to a four-year and having indicated problems getting information about transfer. Other differences that may relate to transfer aspirations include concerns about the quality of teaching and the college staff, since students who intend to transfer may have higher expectations in this regard.

However, we questioned whether these statistically significant mean differences reflected true differences in the ways that people respond to web or paper surveys, or whether they reflected differences in who received web or paper surveys. To explore this issue, we examined differences between the paper mail-out sample and the online mail-out sample based on their responses to the initial Spring 2001 questionnaire. As suspected, individuals in the email sample (i.e., those who provided valid email addresses on the baseline survey) were more likely to be employed (at the part-time level), to be unmarried, to not have children, and to be transfer-seeking than students who, by default,

were placed into the paper sample. This finding confirms that differences between paper and email responses are largely the result of sampling bias and not response bias.

Discussion

This paper explored three primary questions in a longitudinal study of community college students: (1) Do online surveys yield higher rates of response than do paper surveys? (2) Is the nonresponse bias characteristic of online surveys similar to or different from that of paper surveys? (3) Are there differences between online survey responses and paper survey responses, despite identical survey items?

Results indicate that response rates to the online survey were higher than those found for the paper survey regardless of race or gender of respondent. As discussed earlier, the fact that this pattern differs from that reported in recent research on college students is likely attributable to the fact that the online mail-out sample was comprised entirely of individuals who had provided a valid email address on the initial (Spring 2001) questionnaire. While on the one hand this fact points to sampling bias, it also serves as an important lesson in survey administration: To achieve higher response rates and reduce costs in follow-up surveys, it is wise to collect multiple forms of contact information in the baseline questionnaire or interview. Students who are contacted via valid email addresses are more likely to respond and do not incur the expense of being sent a paper questionnaire via standard mail.

When considering predictors of nonresponse, we find that some of the bias traditionally produced in paper surveys is reproduced in email surveys, such as age and prior academic achievement. This is important since it indicates that new modes of survey administration do not help us to reach certain groups of students who tend to be

underrepresented in more traditional survey formats—most critically, lower-achieving students.

Bias connected to race/ethnicity also exists in both modes of administration, but the patterns are not uniform (i.e., the positive effect of being White/Caucasian on the likelihood of response to the paper survey, and the positive effect of being Mexican American on the likelihood of response to the online survey). Notably, the study suggests that online methodologies may yield more balanced samples with respect to gender.

The issue of response bias was more difficult to assess in the current study. Although statistically significant mean differences did emerge between item responses to the paper and email surveys, especially with respect to marital status, employment, and transfer-aspiration, further investigation revealed that these disparities reflect differences in the students who had self-selected into the mail and email samples. Based on this finding, the next step in our research is to match paper and email samples on the basis of key variables such as age, marital status, employment status and GPA. This selection process will enable us to conduct cleaner analyses of both response bias and nonresponse bias.

Interestingly, other than the response differences attributable to selection bias, there was little difference in student responses to the items in the paper and email surveys. This is certainly good news for those engaged in the administration of both online and standard mail questionnaires, since it suggests that we can safely aggregate data from both modes of administration.

In sum, this study suggests that online survey methodologies may be a more effective mode of reaching community college students than paper surveys sent via standard mail if one has valid email contact information. In that sense, the study provides evidence of the value of collecting both mailing address and email address at the point of initial contact with the student. However, an important lesson to be learned from the present study is that if students do self-select into paper and email follow-up samples, it compromises one's ability to conduct research on response bias and nonresponse, since such studies would ideally be conducted on controlled or matched samples of students. Perhaps this points to an inherent tension between the need to advance the study of survey methodologies and the basic need to collect data (i.e., getting the highest response rate possible).

References

Carini, R.M., Hayek, J.C., Kuh, G.D., Kennedy, J.M., & Ouimet, J.A. (2003). College student responses to web and paper surveys: Does mode matter? Research in Higher Education, 44, 1-19.

Dey, E.L. (1997). Working with low survey response rates: The efficacy of weighting adjustments. Research in Higher Eduation, 38 (2), 215-227.

Goree, C. T., & Marszalek, J. F. (1995). Electronic surveys: Ethical issues for researchers. College Student Affairs Journal, 15 (1), 75-79.

Ebo, B. (Ed.). (1998). Cyberghetto or cybertopia: Race, class, and gender on the internet. Westport, Connecticut: Praeger.

Johnson, T.P., O'Rourke, D., Burris, J. & Owens, L. (2002). Culture and survey nonresponse. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J.A. Little (Eds.), Survey Nonresponse (pp. 55-69). New York: Jon Wiley & Sons.

Hamilton, J.C. (1999, December 30). The ethics of conducting social-science research on the Internet. The Chronicle of Higher Education, p. B6.

Lepkowski, J.M. & Couper, M.P. (2002). Nonresponse in the second wave of longitudinal household surveys. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J.A. Little (Eds.), Survey Nonresponse (pp. 259-272). New York: Jon Wiley & Sons.

Matz, C.M. (1999). Administration of web versus paper surveys: Mode effects and response rates. Masters Research Paper. University of North Carolina, Chapel Hill. ED439694.

Sax, L.J., Ceja, M., & Teranishi, R.T. (2001). Technological preparedness among entering freshmen: The role of race, class, and gender. Journal of Educational Computing Research, 24 (4), 1-21.

Sax, L. J., Gilmartin, S. K., & Bryant, A. N. (forthcoming). Assessing response rates and nonresponse bias in web and paper surveys. Research in Higher Education, 44.

Smith, C.B. (1997). Casting the net: Surveying an Internet population. Journal of Computer Mediated Communications, 3 (1), 77-84.

Underwood, D., Kim, H., & Matier, M. (2000). To mail or to web: Comparisons of survey response rates and respondent characteristics. Paper presented at the Annual Forum of the Association for Institutional Research. Cincinnati, OH. ED446513.

Table 1. Response Rates to Follow-Up Survey, by Mode of Survey Administration

	Total Number of Students Contacted	Total Number of Surveys Returned	Response Rate
Group A: Paper-Only	2832	445	15.7
Group B: Email-Only	1555	490	31.5
Total	4387	935	21.3

Note: Mail-out and respondent samples exclude students who 1) received the survey as a paper and electronic form, 2) did not provide a valid address at which to contact them for follow-up, and/or 3) were contacted by telephone as a late-administration effort to maximize overall response.

Table 2. Response Rates to Follow-Up Survey, by Survey Mode, Sex, and Race/Ethnicity[1]

	Group A: Paper-Only			Group B: Email-Only			
	Total Number of Students Contacted	Total Number of Surveys Returned	Response Rate	Total Number of Students Contacted	Total Number of Surveys Returned	Response Rate	Paper/Email Split
Sex							
Women	1686	317	18.8	924	311	33.7	65 - 35
Men	1060	117	11.0	604	172	28.5	69 - 31
Race/Ethnicity							
White/Caucasian	350	68	19.4	288	102	35.4	55 - 45
Black/African American	487	90	18.5	187	58	31.0	72 - 28
Mexican/Mexican American	982	144	14.7	461	163	35.4	68 - 32
Latino/a	630	98	15.6	309	97	31.4	67 - 33
Asian	214	32	15.0	207	51	24.6	51 - 49

[1]Due to small counts, Race: Pacific Islander and Race: Asian Indian were not included in these response rate calculations. Mail-out sample counts by sex may not sum to full sample because some respondents did not mark their sex on the survey.

Note: Mail-out and respondent samples exclude students who 1) received the survey as a paper and electronic form, 2) did not provide a valid address at which to contact them for follow-up, and/or 3) were contacted by telephone as a late-administration effort to maximize overall response.

Table 3. Predictors of Response to Follow-Up Survey, by Mode of Survey Administration

	Logistic Regression Coefficients and Standard Errors (in parentheses)		Odds Ratios	
Independent Variable	Group A (paper)	Group B (email)	Group A (paper)	Group B (email)
Sex: Female	**.590 *** (.119)**	.193 (.118)	**1.805**	1.213
Age	**.246 ** (.085)**	**.384 *** (.097)**	**1.279**	**1.468**
Race/Ethnicity: White/Caucasian	**.316 *** (.153)**	.280 (.147)	**1.372**	1.324
Race/Ethnicity: Mexican/Mexican American	.019 (.116)	**.368 ** (.126)**	1.019	**1.445**
Degree aspirations	.041 (.039)	**.219 *** (.048)**	1.042	**1.244**
Average grade in high school	**.061 *** (.029)**	**.071 *** (.030)**	**1.063**	**1.074**
Reason for attending this college: Something to do	-.019 (.017)	**-.054 ** (.020)**	.981	**.947**
Constant	**-3.915 *** (.474)**	**-3.769 *** (.511)**	**.020**	**.023**

Note: * p<.05 ** p<.01 *** p<.001

Table 4. Response Bias in Follow-Up Data, by Mode of Survey Administration

Variables with statistically significant mean differences by mode (p<.01)	Mean of paper respondents (SD in parentheses)	Mean of online respondents (SD in parentheses)
Paper > Online		
Identity: Primarily a student who is a parent[1]	1.06 (.24)	1.02 (.15)
Employment status: Not employed and not looking for work[1]	1.14 (.34)	1.08 (.27)
Experience since Jan. 2001: Filled out a form for financial aid[1]	1.56 (.50)	1.41 (.49)
Experience since Jan. 2001: Marriage[1]	1.08 (.27)	1.03 (.17)
Current religious affiliation: Christian Science[1]	1.04 (.20)	1.00 (.05)
Paper < Online		
Identity: Primarily a student who is employed[1]	1.30 (.46)	1.40 (.49)
Identity: Primarily a parent who is an employee[1]	1.01 (.11)	1.05 (.22)
Degree aspirations[2]	5.47 (1.45)	5.94 (1.13)
Experience since Jan. 2001: Application for admission to a four-year college[1]	1.17 (.38)	1.25 (.43)
Problem in obtaining education: Finding time for school[3]	2.24 (1.27)	2.49 (1.29)
Problem in obtaining education: Quality of teaching[3]	1.68 (.90)	1.89 (1.05)
Problem in obtaining education: College staff[3]	1.51 (.83)	1.73 (1.05)
Problem in obtaining education: Lack of information about transfer[3]	1.90 (1.15)	2.11 (1.24)
Problem in obtaining education: Job responsibilities[3]	2.22 (1.28)	2.60 (1.34)

Note: Means were compared using an independent sample t-test. Levene's test was used to determine equality of variances.

[1]Dichotomous variable: 1= "not marked" 2= "marked"

[2]Seven point scale: 1= "will take classes but do not intend degree" to 7= "doctoral or medical degree"

[3]Five-point scale: 1= "not a problem" to 5= "very large problem"

Appendix A

Variable List and Coding Schemes

Dependent Variables	**Coding Scheme**
Paper Response (among students in Group A)	Dichotomous variable: 0= "no," 1= "yes"
Email Response (among students in Group B)	Dichotomous variable: 0= "no," 1= "yes"
Independent Variables	**Coding Scheme**
Sex: Female	Dichotomous variable: 1= "male," 2 = "female"
Race/Ethnicity: White/Caucasian	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Black/African American	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Mexican/Mexican American	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Latino/a[1]	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Asian[2]	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Pacific Islander[3]	Dichotomous variable: 1= "not marked", 2= "marked"
Race/Ethnicity: Asian Indian[4]	Dichotomous variable: 1= "not marked", 2= "marked"
Age	Three-point scale: 1= "20 or younger" to 4= "40 or older"
Average income	Fourteen-point scale: 1= "Less than $6,000" to 14= "$200,000 or more"
Plan to attend the same college next semester	Dichotomous variable: 1= "no", 2= "yes"
Number of other colleges/universities attended	Three-point scale: 1= "None" to 3= "2 or more"
Degree aspirations	Six-point scale: 1= "Take classes only/Vocational certificate" to 6= "Doctoral or medical degree"
Hours per week: Work at a job	Nine-point scale: 1= "0, none, or didn't have time" to 9= "46 hours or more"
Hours per week: Do housework or childcare	Nine-point scale: 1= "0, none, or didn't have time" to 9= "46 hours or more"

Hours per week: Spend time on this campus (including time in class)	Nine-point scale: 1= "0, none, or didn't have time" to 9= "46 hours or more"
Obstacle to education: Understanding the English language	Five-point scale: 1= "Not a problem" to 5= "Very large problem"
Length of commute to campus	Six-point scale: 1= "Less than 15 minutes" to 6= "More than 2 hours"
Disability: Mobility impaired	Dichotomous variable: 1= "not marked", 2= "marked"
Disability: Attention deficit disorder	Dichotomous variable: 1= "not marked", 2= "marked"
Average grade in high school	Nine-point scale: 1= "D or lower (Poor)" to 9= "A or A+ (Extraordinary)"
Level of math preparation	Seven-point scale: 1= "Basic math/business math/pre-algebra" to 7= "Calculus"
Live alone while attending this college	Dichotomous variable: 1= "not marked", 2= "marked"
Own a computer with Internet access	Dichotomous variable: 1= "no", 2= "yes"
Reason for attending this college: Something to do	Two-item composite measure
Reason: I couldn't find a job	Seven-point scale: 1= "Very unimportant" to 7= "Very important"
Reason: I couldn't find anything better to do	Seven-point scale: 1= "Very unimportant" to 7= "Very important"
Academic involvement: Interaction with instructors/ academic counselors	*Five-item factor – see Appendix B*
Academic involvement: Studying with others	*Five-item factor – see Appendix B*
Views: Determined and confident	*Nine-item factor – see Appendix B*
Positive attitude towards school	*Three-item factor – see Appendix B*

[1]Includes South American, Central American, and Other Latino/Hispanic
[2]Includes Chinese, Japanese, Korean, Thai, Laotian, Cambodian, and Vietnamese
[3]Includes Filipino, Samoan, Hawaiian, Guamanian, and Other Pacific Islander
[4]Includes South Asian (Indian subcontinent), Arab, and American Indian

Appendix B

Factors: Loadings, Coding Schemes, and Cronbach's Alphas

Factor	Loading
Academic involvement: Interaction with instructors/academic counselors (α=.74)	
Class-related activity in past week (for course in which student completed survey): Ask the instructor questions[1]	.78
Class-related activity in past week (for all courses): Talk with an instructor before or after class[1]	.72
Class-related activity in past week (for course in which student completed survey): Speak up during class discussions[1]	.68
Class-related activity in past week (for all courses): Talk with an instructor during office hours[1]	.63
Class-related activity in past week (for all courses): Speak with an academic counselor[1]	.53
Academic involvement: Studying with others (α=.73)	
Hours per week: Study with students from this course[2]	.74
Class-related activity in past week (for all courses): Study in small groups outside of class[1]	.71
Hours per week: Study with students from other courses (not this course)[2]	.62
Class-related activity in past week (for course in which student completed survey): Telephone or email another student to ask a question about your studies[1]	.59
Class-related activity in past week (for all courses): Help another student understand homework[1]	.44
Views: Determined and confident (α=.84)	
View: I expect to do well and earn good grades in college[3]	.72
View: Understanding what is taught is important to me[3]	.71
View: It is important to me to finish the courses in my program of studies[3]	.71
View: I feel most satisfied when I work hard to achieve something[3]	.70
View: I am very determined to reach my goals[3]	.68
View: Success in college is largely due to effort (has to do with how hard you try)[3]	.64
View: I keep trying even when I am frustrated by a task[3]	.60
View: I always complete homework assignments[3]	.55
View: I know I can learn all the skills taught in college[3]	.46
Positive attitude towards school (α=.60)	
View: I enjoy doing challenging class assignments[3]	.66
View: My teachers here give me a lot of encouragement in my studies[3]	.55

[1]Six-point scale: 1= "0, or didn't have time" to 6= "5 or more times"
[2]Nine-point scale: 1= "0, none, or didn't have time" to 9= "46 hours or more"
[3]Seven-point scale: 1= "Strongly disagree" to 7= "Strongly agree"

Note: Students' raw scores on items to comprise each factor were summed to compute factor scores. In the event that items within a factor were scaled differently (e.g., the items in "Academic involvement: Studying with others"), students' scores were standardized and then summed to compute factor scores.