

College Students' Beliefs and Values (CSBV) Pilot Survey Methodology

With a goal of developing an institutional sample that reflected diversity in type, control, selectivity, and geographical region, between late February and early March 2003 we contacted representatives at approximately 150 colleges and universities via electronic mail to invite their institutions to participate in the study. Our intention was to select roughly equal numbers of institutions within each category of type, control, selectivity, and geographical region. Ultimately, 47 schools were able to work within a very tight timeframe to secure any necessary institutional approval to participate in the study and to provide us with the updated student contact information that we needed.

At each institution, our goal was to randomly sample an average of 250 third-year students. Participating institutions facilitated our direct mail administration of the survey by providing us with updated (i.e., current academic year) local mailing addresses for students who: (a) had completed the CIRP survey at that institution as entering freshmen in Fall 2000; (b) were still enrolled as of Spring 2003; and (c) had given the Higher Education Research Institute (HERI) permission when they completed the 2000 CIRP to contact them again for research purposes. To determine the number of student names that we would need to send each institution for address updating in order to yield 250 juniors (or "enrolled students") per institution, we calculated each institution's expected six-year retention rate using a formula devised by Astin and Oseguera (2002). As it turned out, this method yielded reasonably close to 250 useable addresses at most institutions. The number of names sent to each institution was then calculated as 250 divided by the estimated proportion retained. In addition to local mailing addresses, 32 of the institutions also provided us with students' email addresses.

In late March 2003, we sent postcards to a total of 12,030 students introducing the study and notifying them that they would receive a questionnaire and more information about the project within the next two weeks. Then, in early April 2003, we mailed each student the four-page questionnaire along with a cover letter explaining the purpose of the study. On the reverse side of this letter was information for students pertaining to their rights as participants in the study.

To explore the comparative effects of differential monetary incentives on response rates, we categorized institutions by type (public university, private university, public college, private nonsectarian college, Catholic college, other religious college) and selectivity (low, medium, high, very high). Within each type/selectivity category, we then randomly assigned individual institutions to one of three monetary incentive groups (\$0, \$2, \$5). To the greatest extent possible, and within the limits of our available resources, we assigned at least one institution of each type and selectivity level to each of the three cash incentive groups. For example, students at one moderately selective public college received a \$5 cash incentive while their counterparts at each of two other moderately selective public colleges received a \$2 cash incentive and no monetary incentive, respectively. Overall, students at 13 institutions received a \$5 cash incentive. Those at 17 institutions received \$2, while those at the remaining 18 institutions received no incentive. All incentives were included inside the envelope containing the first survey packets that students received.

Two weeks after the initial questionnaire was mailed, we selected a sample of students to receive an email reminder from the total population of students for whom we had email addresses. Two weeks later, a second questionnaire (without monetary incentives) along with a modified cover

letter and information sheet was mailed to nonrespondents; one week later we sent a second email reminder to a selected group of non-respondents for whom we had valid email addresses.

Of the 11,547 students in the sample pool whose survey envelopes were not returned as undeliverable, we ultimately received useable questionnaires from 3,680 students, representing a 32 percent overall response rate.¹ Women were about 50% more likely than men were to respond, and there was considerable variation in the overall response rates for men and women who received different monetary incentives. Money did have a substantial effect: \$2 increased the rate of response by about half, and \$5 increased it by more than two-thirds. There were positive (but smaller) effects of email reminders. In absolute terms, students who received reminder emails were about 4% more inclined to respond than their peers who did not to return their questionnaires (33.5 percent versus 29.4 percent, respectively).

To approximate the results that would have been obtained if all 3rd year full-time students at each participating institution had responded, we employed a multi-stage weighting procedure. The first set of weights was designed to adjust for non-response bias. Equations derived from regression analyses were used to compute response probabilities for men and women. Independent variables consisted of most items from the 2000 entering freshman survey, whereas the dependent variable was a dichotomy

¹ One college was dropped from the final sample due to an inexplicably low response rate.

(responded versus did not respond).²

After the regressions were run, all non-respondents were dropped from the file. The weights derived were the reciprocal of the probability of response. In other words, respondents who were typical of those freshmen with a low probability of response received high weights, and respondents most resembling those with a high probability of response received low weights. The second set of weights was designed to bring the respondent file counts up to the “population,” which in this case was the total number of first-time, full-time freshmen from Fall 2000 who were still enrolled in Spring 2003 at this sample of institutions. The final weight used for each respondent consisted of the product of the two weights.

To keep the degrees of freedom at an appropriate level for purposes of statistical inference, these final weights were “normalized,” such that their sum equaled the actual number of respondents (i.e., $N = 3,680$). The weight variable used in deflating the sample to 3,680 was derived by dividing the original weight by the ratio of the weighted sample to the unweighted sample.

²In recent years it has become fashionable to use logistic regression instead of ordinary least squares (OLS) regression when the dependent variable is a dichotomy. However, since an extensive empirical comparison of the two methods using CIRP data (Dey & Astin, 1993) shows that they yield essentially identical results, we chose to use OLS regression because the SPSS program includes important options (e.g., “Beta in” for variables not in the equation) not available in the logistic regression program. Further, a recent methodological study using the same data employed in the present study shows that OLS regression and logistic regression produce cross-validated results that are essentially identical (Oseguera & Vogelgesang, 2003).

References

- Astin, A. W., & Oseguera, L. (2002). *Degree attainment rates at American Colleges and Universities*. Los Angeles: Higher Education Research Institute, UCLA.
- Dey, E. L., & Astin, A. W. (1993). Statistical alternatives for studying college student retention: A comparative analyses of logit, probit, and linear regression. *Research in Higher Education*, 34, 569-581.
- Oseguera, L., & Vogelgesang, L. (2003). *Statistical alternatives for studying college student retention: Logistic versus linear regression - An update*. Unpublished manuscript. Los Angeles: Higher Education Research Institute, UCLA.