

## SURVEY RESEARCH AND RESPONSE BIAS

Anne G. Scott, Lee Sechrest, University of Arizona  
Anne G. Scott, CFR, Educ. Bldg. Box 513, University of Arizona, Tucson, AZ 85721

**KEY WORDS:** Missing data, higher education

Missing data on surveys, whether due to missing cases or due to skipped items, can pose problems for both choosing a statistical analysis and interpreting results. Technically, there are many ways of dealing with missing data (c.f., Kalton & Kasprzyk, 1982); however, bigger issues are, Why are there missing data? and, Do these missing data represent a bias in the results? In higher education, surveys are frequently used to augment basic information on students available through Student Information Systems (SIS) (e.g., demographic information, incoming academic characteristics such as ACT/SAT scores and high school GPAs, current grades, and major). They often provide information to administrators on particular problems that students may be experiencing (e.g., course non availability, poor advising, and difficulty transferring credits from two-year institutions). The ramifications of biased survey results can be bad policy decisions. Therefore, it is imperative that researchers understand the extent to which survey design and administration can impact the findings.

Non response can mean either refusing to fill out and return surveys or skipping items on a survey. Return rates on surveys can range from below 10% to over 90% depending on various factors including attractiveness of the survey, perceived importance of the subject matter, and the method of distribution used (Dillman, 1978). The potential impact of missing data, or non response, for specific items on surveys is analogous to that of omissions in the sampling frame, i.e., part of the target population has been omitted from the sample (Henry, 1990). Results on items for which a substantial number of responses (usually over 10%) are left blank no longer necessarily

accurately represent the opinions or attitudes of the target population.

The best way to deal with non response is to minimize it from the beginning. Obviously, the smaller the number of non responses (of either type), the lesser will be any bias resulting from non response. Often missing data are due to factors that are under the researcher's control. Modifying some of these factors, such as survey length and the wording or items, can decrease non response for specific items (Sudman & Bradburn, 1988). Changing others, such as the method of survey administration, can improve the overall return rate. On the other hand, when it is not possible to optimize survey administration or format, then it is important to know how results may be biased so that the presentation of results is not misleading.

To minimize survey bias, one must understand both the causes and the effects of non response. This paper explores several survey variations -- distribution strategies, survey length, and type of items -- tested over several years and examines both the reasons for non response and the effects of non response on the overall results.

### Experiment 1: Survey Distribution Methods

At our institution, surveys are given at two key points in a student's career: when they first arrive (New Student Survey) and when they have upper division standing (Student Experiences Questionnaire, or SEQ). Although an appropriate time (during the required upper division writing exam, or UDWPE) was found for distributing the SEQs, the method of distribution to be used was initially undetermined. Because of the additional time and effort that adding another task to the testing process would entail, the English Composition Board wanted to rely on returns by mail. In order to convince the Board to incorporate the survey into their testing process it was necessary

---

We thank Vanessa Hayden at the University of Arizona for reviewing and assisting in the preparation of this manuscript.

to demonstrate that use of different distribution methods would result in different return rates. Because the UDWPE is administered monthly (with no discernable difference in the composition of students taking the test at each time point, with the exception of the summer administration), different methods were used at different time points during the year.

### Method

Three different methods were used to distribute the SEQ during part of the 1990-91 school year. For each method, the survey was identical and consisted of 96 multiple choice items. Instructions on the surveys informed students to write, using either pen or pencil, directly on the survey.

Method 1 involved giving surveys to students as they signed up to take the UDWPE. As students came to the Composition Board Office to register for the UDWPE, they were then handed the SEQs, and instructed to fill them out and return them to the secretary. There were chairs in the corridor that students could use while filling out the surveys. One hundred surveys were distributed using this method.

Method 2 involved mailing surveys to students along with their UDWPE test results about a month after they had taken their exam. Stamped, self-addressed envelopes were provided for returning the survey. Two hundred surveys were distributed using this method.

Method 3 consisted of two subgroups, each of which involved handing out surveys as students entered the testing room to take their UDWPE. There were two different test sites, with different instructors overseeing the

distribution and providing instructions. The Group 1 instructor told the students that participation was voluntary but gave no other instructions. The Group 2 instructor was given a prepared paragraph describing both the purpose of the survey and its importance to planning at the university (and that participation was voluntary). Both groups of students were told that they had 15 minutes prior to the UDWPE to fill out the survey and that if they did not finish the survey at that time they could complete it after they had finished their UDWPE. The total number of surveys distributed by this method was 436: 206 for Group 1 and 230 for Group 2.

### Results

As displayed in Table 1, the various methods resulted in very different response rates. These response rates ranged from 11% for the first method to 78% for Group 2 of the third method (specific instructions given).

These results demonstrated to the Composition Board that the method used in distributing surveys clearly made a difference in the response rate. In order to further improve survey return rate, the Board agreed to include the SEQs in their packet containing the writing exam so that students could begin filling them out as soon as they entered the test room and to provide written instructions to staff who would be distributing the packets. Because of this change the response rate has been consistently high (90-95%).

*Effects of Non-response.* Next, the responses made by students responding to each method were examined. Means and standard deviations were calculated for each item on the

Table 1: Return rate by different distribution methods

Method 1: Return at sign up	Method 2: Return by mail	Method 3: Passed out prior to UDWPE	
		Group 1 No Instruction	Group 2 Detailed Instruction
n=100 Return 11%	n=200 Return 28%	n=206 Return 30%	n=230 Return 78%

Table 2: Responses on the SEQ

Item	Group 1		Comparison (full sample)	
	Mean	S.D.	Mean	S.D.
How often did you attend campus social events during the past year?*	2.7	1.1	2.1	0.9
If you could start over, would you still attend the UofA?*	2.6	1.2	1.8	0.8
Would you recommend the UofA to a friend who is thinking about attending college?*	2.3	1.1	1.7	0.8
How committed do you feel the UofA is to your success?*	2.7	0.8	2.1	0.8

\* significant,  $p < .05$

survey, for each method. The means and standard deviations were then compared to those obtained for the full sample ( $n=320$ ) in order to determine whether there were any significant differences among responders. The only group that differed was in Group 1. Table 2 displays those items that differed for that group. Note that Group 1 made higher responses on all items. A higher score for the first two items indicates more activity and for the last two items indicates dissatisfaction.

Another difference for Group 1 was in the number of respondents who marked "bureaucracy" as being their biggest frustration on campus. All respondents in Group 1 (100%) checked off this item as opposed to 34% of the other group ( $X^2_{1, n=371} = 18.24, p < .01$ ). Students in Group 1 also differed in the number of comments written on the survey. All respondents in this group (100%) wrote in comments. Since the SEQ is a multiple choice survey and there are no open ended items, comments written in are fairly rare (only 138 out of 3976 (3.5%) wrote in comments for the entire year).

Although the sample size for Group 1 is very small, the differences in responding are pronounced. One interpretation of these results is that the more inconvenient it is for students to complete and return a survey, the more likely it is that only those students who have strong opinions (usually negative) will complete them.

This interpretation is also congruent with results from our other surveys: On mail-in surveys quite a number of responders take the time to write comments and even attach additional pages; yet very few of the responders do so with other methods of survey administration.

#### Experiment 2: Survey Length

Another factor that may affect non response and is also under the control of the researcher is survey length. Generally, the shorter a survey, the more likely respondents are to complete it. Non response on a long survey may be due to its formidable nature, which can result in total non response, or to other factors such as motivation, boredom, difficulty of the task, as well as time constraints. In order to determine whether there is any non response bias due to survey length, response rates of two different versions of the SEQ (one short, one long) were compared.

#### Method

During 1991-92 the SEQ contained 94 multiple choice items. The following year (1992-93) the SEQ was cut to 36 items. Survey format and administration was identical for both survey administrations (see methods section above). The time allotted was three times as long for the first year (longer SEQ). During 1991-92, 4863

students were surveyed and during 1992-93, 3239 were surveyed.

### Results

Both survey administrations resulted in overall high (94-96%) response rates. However, response rates for the individual items on the longer version ranged from 99% for those items that were placed at the beginning of the survey to 77% for items placed near the end (last 18 items). For the shorter version, there was virtually no difference in response rate (94-98%) for any of the individual items, regardless of where they were placed on the survey.

Next, differences in responses made by non responders were compared to those of completers on the first two-thirds of the longer version (items 1-76). No difference in responding was found for any of the items. Students who finished the survey were similar in most ways (i.e., ethnicity, HSGPA, SAT/ACT scores, current GPA) to those who did not. The only difference was gender: Females were more likely to finish than males (53.6% of females completed the survey as opposed to 46.6% of males ( $X^2_{1, n=873} = 5.62, p < .01$ )).

These findings indicate that although not all responders had completed the survey, there is no apparent bias, except for gender, on the items at the end of the survey. This information can be useful in certain situations when there is a great deal of information to be obtained and many items are needed on a survey. In some cases, complete data for each subject are necessary in order to link survey data to other data (e.g., grades and persistence) and run predictive models. At other times, when the purpose is merely to collect information on a sample of students, missing data near the end of the survey on part of the sample is not a limiting factor. It may be useful to keep in mind, however, that if a survey must be long, vital items should not be placed near the end. Most researchers agree that placing sensitive items last is a good idea since by that time responders have "bought in" to the survey and not likely to be turned off by those items immediately. However, if these sensitive items (for example,

ethnicity and income) are important for linking to previous items then they need to either be placed earlier in a survey or on a shortened version of that survey.

### Experiment 3: Types of Questions

There are always certain trade offs in asking particular questions on surveys. One issue relevant to surveys in higher education is whether to ask for the students' identification numbers (SIDs), which enable the researcher to link survey information to other existing data on campus (e.g., demographic information, incoming academic characteristics such as ACT/SAT scores and high school GPAs, current grades, major), or to leave the survey completely anonymous. In higher education, obtaining the student identification number allows the researcher to ask only those items currently of interest, skipping demographic items that are easily obtained from the Student Information System (SIS). On the other hand, leaving the survey anonymous may improve the honesty and accuracy of students' responses, especially to sensitive questions.

Ethnicity is one variable that researchers in higher education find useful, for various reasons. For example, researchers often need data on ethnicity to evaluate programs involved in the recruitment and retention of under represented students. In addition, certain policy decisions may have differential impact for students of different ethnicity, making data on these issues by ethnicity imperative. Results from past surveys have demonstrated that asking ethnicity, even when the item is placed at the end of the survey and phrased as being an important piece of information in order to be "certain that we are representing all students on our campus," is potentially irritating or worse to some students. On the other hand, if that information is necessary, the only other alternative is to ask for identification of students. Therefore, it is a trade off between irritating and potentially losing some respondents and obtaining information that may be biased because of lack of anonymity. In order to determine what type of bias, if any, in

responding is introduced into the results through the use of one approach over the other, the two approaches were compared.

### *Method*

For the 1992-93 SEQ, two different versions of the survey were produced. One survey requested SIDs and no further demographic information (n=78). The other survey requested students to mark their ethnicity, but not their SID. The request for ethnicity was placed at the end of the survey and was prefaced with the explanation that the university was concerned that it respond to the needs of all students on campus (n=52). The two different versions of the SEQs were randomly distributed to students taking their required UDWPE. Both surveys assured confidentiality of responses and were described as being completely voluntary.

### *Results*

Both surveys had a 95% return rate, indicating that there was no difference in overall response rate. Next, responses for each group (SID versus anonymous) were compared. Responses to all of the items were almost identical with the exception of one item: "How satisfied are you with the attitude of faculty toward undergraduate students?" Students who were not asked to provide their SIDs (anonymous) were more likely to respond that they were dissatisfied on that item (45% versus 27%;  $X^2_{1, n=130} = 4.28, p < .05$ ). This finding is interesting because there was no difference between the groups on items asking for overall ratings of the university, implying that there probably is very little or no bias on those items regardless of whether the survey is anonymous or not. On the other hand, students are reluctant to report negatively about those people on campus that they currently have contact with, i.e., faculty, as reflected by the differences of response for the two groups. Since this survey is filled out during mid-semester students may believe that there could be repercussions to their responses.

In order to determine if students find either SID or ethnicity more offensive, a third version of the SEQ that asked both items was administered at another testing session (n=1128). This survey was administered identically to other SEQ administrations. For this survey administration, the proportions of students providing valid SIDs and of those providing ethnicity were almost the same (93% of the students provided their ethnicity and 91% provided valid SID numbers). Fifteen students refused to fill in their ethnicity but had provided their SID so it was possible to determine their ethnicity: Of those, 80% were white; 13%, Hispanic; and 7%, Foreign.

In order to determine why students react to the ethnicity question, the surveys that had comments written in on that item were reviewed (n=21). Eight of these students did not provide their SID numbers and four did not provide their ethnicity. The ethnicities of those who did reply were distributed across all categories (White was marked by six; African American, one; Hispanic, three; Native American, one; Asian American, one; other, five). Comments ranged from a more specific description of their ethnicity (e.g., Navajo, Puerto Rican) to negative statements about that item (such as "I don't think that is any of your concern,") accompanying a refusal to respond.

Although some of the students were not willing to provide either ethnicity and/or their SID numbers, the actual number of these students is quite small and not likely to affect any findings overall. Currently we do not request SIDs on surveys that are likely to be sensitive or that we have reason to believe may be biased if we requested identifying information.

### *Conclusions*

This paper has addressed some common non response situations in higher education survey research. Some of these different methods resulted in differences that are quite dramatic, such as distribution method, while others do not appear to have significant impact on findings, for example, the length of the

survey. Although it is not possible to predict and test all biases that might be present, it is important to continue to study this phenomenon and, perhaps more importantly, to be constantly aware that findings may be due at least to some part to the methods used rather than to some real phenomenon.

#### References

- Dillman, D. A. (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley & Sons.
- Henry, G. T. (1990). *Practical Sampling*. Beverly Hills, CA: Sage.
- Kalton, G., & Kasprzyk, D. (1982). Computing for Missing Survey Responses. *Proceedings of the Section on Survey Research Methods*. American Statistical Association, 22-31.
- Sudman, S., & Bradburn, N. M. (1988). *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.