

5

This chapter reviews the literature on survey fatigue and summarizes a research project that indicates that administering multiple surveys in one academic year can significantly suppress response rates in later surveys.

Multiple Surveys of Students and Survey Fatigue

*Stephen R. Porter, Michael E. Whitcomb,
William H. Weitzer*

As described in Chapter One, survey nonresponse has been increasing both in the United States and internationally, and much of this nonresponse is due to rising rates of refusal. In many discussions about the rise in survey nonresponse, survey fatigue is often cited as one possible cause. Steeh (1981, p. 53), for example, cites “overexposure to the survey process,” while de Heer (1999) notes some interesting variations in the number of surveys being conducted in different countries as influencing response rates across countries.

Despite the view that rising nonresponse rates are in part caused by an increase in the number of surveys, there has been almost no research on the impact of multiple survey requests on survey response. In part this is not surprising, because most research on survey nonresponse analyzes only one survey and thus focuses on a single point in time (Harris-Kojetin and Tucker, 1999). Yet the issue of survey fatigue will become increasingly important as the costs of designing and administering a survey decrease. A variety of software products now allow anyone with minimal technical skills to create and administer a simple Web survey.

This issue is also of vital importance to research in higher education, as the use of student surveys in assessment and institutional research continues to increase. A large array of national surveys that together can be used to describe and assess almost any facet of the undergraduate experience are currently available. Most colleges and universities have their own internally designed surveys as well. Add to the mix the growing pressures for assessment from outside groups such as legislatures and accrediting

agencies, and internal pressures from individual offices trying to show performance results and the pressure to administer multiple surveys can be intense.

Even if the number of surveys on a campus is limited, the timing of the surveys could be such that two surveys may overlap or be administered back-to-back. On many campuses this may even happen unknowingly, as different offices administer their particular survey unaware of the actions of other offices. Educational researchers must understand the impact of multiple surveys on response rates in order to appropriately design and implement their surveys. Understanding the impact of multiple surveys can allow an institution to juggle various demands and can also act as an impetus for the development of a survey research policy.

Quantifying the impact of survey fatigue is also useful for individual institutions grappling with demands for surveys from numerous internal constituencies. For larger schools, multiple surveys do not necessarily pose a problem, because many large samples of students can be drawn without surveying the same students twice. For smaller schools, however, conducting multiple surveys inevitably means that the same students will be surveyed multiple times.

This chapter describes two experiments conducted at a selective liberal arts college to quantify the impact of multiple survey requests on student survey response behavior. We seek to answer two questions. First, does implementing multiple student surveys have a negative effect on later surveys? Second, if so, does this effect vary by subpopulations of students?

Previous Research

Survey fatigue is one component of respondent burden, generally defined as the time and effort involved in participating in a survey (Sharp and Frankel, 1983). Much of the research on respondent burden has focused on interview length and has generally found that longer surveys result in lower response rates (see the discussion in Chapter One of this volume). Although little research has been done on the impact of multiple surveys on response rates, research on respondent burden sheds some light on the potential impact of survey fatigue. This research can be divided into three areas. The first looks at nonresponse in panel surveys, in which respondents are interviewed several times during a research project rather than just once. The second area of research uses surveys to query nonrespondents about reasons for their behavior. The third area actually analyzes the impact of multiple (different) survey requests on survey nonresponse, the topic of this chapter.

Panel Surveys. Some scholars have looked at respondent burden in panel surveys. Because panel surveys involve several survey iterations, refusals to participate are expected to rise over time because of the increasing burden on the respondent. Thus, increasing nonresponse is a common feature in panel surveys (Kalton, Kasprzyk, and McMillen, 1989). For

example, using several federal government household panel surveys, Atrostic, Bates, Burt, and Silberstein (2001) found refusal rates to increase with each subsequent interview, although this pattern began to taper off after the first few interviews.

Interestingly, one study compared response rates in an experiment that altered the protocol used when approaching members of the sample for a panel survey (Apodaca, Lea, and Edwards, 1998). Simply informing respondents that if they chose to participate in the current interview they would be contacted again several times over the next few years for additional interviews reduced response rates by 5 percentage points. These results, along with the general research on refusal rates in panel surveys, indicate that not only do respondents initially balk at participating in a survey with several interview components, but also some of those who initially agree refuse to participate in later interviews. Clearly multiple surveys are perceived as a burden by respondents, and we would expect response rates to decrease as the number of survey requests increases.

Surveys About Surveys. Research on survey behavior is another area that sheds some light on survey fatigue. In a telephone follow-up of non-respondents to a mail survey, researchers queried for the most important reason why they had not responded to the survey (Sosdian and Sharp, 1980). Twenty percent said they “never got around to it,” implying a possible lack of time to participate, while 17 percent replied “too busy,” and 7 percent said that the survey came “at a bad time” in their personal schedule. Only 1 percent said that the survey was too long. Taken together, these comments indicate that time is an issue in survey response, with the implication that the more time is demanded, as in multiple surveys, the lower the response rate will be.

From a higher education perspective, a research project conducted by the U.S. Air Force Academy (Asiu, Antons, and Fultz, 1998) provides important information. Faced with anecdotal evidence that students were frustrated by the number of surveys being conducted, researchers used focus groups and a survey to determine students’ attitudes toward surveys at the Academy. The results are striking. Almost all of the respondents (97 percent) stated that they felt “somewhat” oversurveyed, with almost half (48 percent) stating that “yes, definitely” they felt oversurveyed. When asked, the students indicated that they should be surveyed only three or four times a year.

Interestingly, a content analysis of student definitions of the term *oversurveyed* revealed that students felt oversurveyed because of “the combination of frequent surveys that are perceived as irrelevant to daily student (cadet) life.” This result indicates that survey fatigue may depend on salience, and that the impact of multiple survey administrations may vary not just because of the number of surveys but because of their content as well.

Studies of Survey Fatigue. Our literature review revealed two studies of the impact of survey fatigue on nonpanel surveys, that is, multiple

surveys from different projects conducted over time. Each study reached a different conclusion about the effect of survey fatigue on response rates. The first study looked at a series of farm surveys conducted over time by the U.S. Department of Agriculture (McCarthy and Beckler, 1999). This study found no relationship between the number of times participants were previously contacted by the USDA and the response rates in a later survey. The second study asked respondents the number of times they had been previously contacted to participate in a survey and found a strong negative relationship between the number of previous survey contacts and participation in a later survey (Goyder, 1986).

From all of these studies we can conclude the following:

- The prospect of multiple surveys can reduce response rates.
- Nonrespondents often cite time concerns as reasons for nonresponse, implying that as the amount of time spent participating in surveys increases, survey nonresponse will increase.
- The effects of survey fatigue may be moderated by the salience of survey content.
- The number of previous surveys may have an impact on current survey response, although the evidence here is mixed.

Two Tests of Survey Fatigue

Using several undergraduate student surveys, we conducted two experiments at a selective liberal arts college to measure the impact of survey fatigue and whether it may have a greater impact on some subpopulations than on others. The first study looks at the impact of a paper survey administered immediately prior to a second paper survey. The second looks at the impact of three Web surveys administered during the fall semester on a Web survey administered during the following spring semester.

Experiment I. The first experiment took place in spring 2001 and used two surveys administered to the senior class ($n = 649$). The class was randomly divided into two groups, with the first group receiving two surveys and the second group receiving one survey. The first group was administered the College Student Experiences Questionnaire (CSEQ) during the last week of March and the first three weeks of April. This paper questionnaire is seven pages long and was administered with a prenotification e-mail, a first mailing of the survey via campus mail, an e-mail reminder to all members of the sample, a second mailing of the survey via campus mail to nonrespondents, and an e-mail reminder to nonrespondents. The response rate was 28 percent.

Beginning in the last week of April and extending into May, all seniors were administered the Senior Survey, an eight-page paper questionnaire asking questions about future plans and about satisfaction with various aspects of their undergraduate education. Survey administration consisted of a

Table 5.1. Experiment I: Senior Survey Response Rates

	Senior Survey response rate (%)		Differences in response rates between groups (percentage points)	Sample sizes	
	No prior surveys (A)	1 prior survey (B)	(B-A)	(A)	(B)
All students	67	57	10**	324	325
Gender					
Female	70	59	11*	173	174
Male	64	54	10†	151	151
Difference			1		
Race					
White	72	59	13**	197	220
Nonwhite	60	52	7	127	105
Difference			5		
1st semester GPA					
A	71	59	12*	159	172
B or less	63	54	9	165	153
Difference			4		

*Response rates differ significantly at the $p = .05$ level.

** Response rates differ significantly at the $p = .01$ level.

†Response rates differ significantly at the $p = .10$ level.

paper prenotification letter, a first mailing of the survey via campus mail, an e-mail reminder to nonrespondents, a second mailing of the survey via campus mail to nonrespondents, and a second e-mail reminder to nonrespondents. Additionally, an in-person request for survey completion was asked of all nonrespondents when they went to pick up their diploma at the end of the survey administration period. The overall response rate was 62 percent.

Both surveys used Dillman’s (2000) method of survey administration, which emphasizes several contacts with respondents, and the two surveys were conducted almost back-to-back. Table 5.1 illustrates the impact of the first survey administration: students who were mailed the CSEQ prior to the administration of the Senior Survey had a Senior Survey response rate 10 percentage points lower than seniors who were not asked to participate in the CSEQ (57 percent and 67 percent, respectively).

Looking at response rates by gender, race, and grade-point average (GPA), there are some interesting differences for some subgroups. Survey fatigue appeared to affect females and males equally. For whites alone, the prior survey had a statistically significant impact on response rates (13 percentage points), but the difference for nonwhites (7 percentage points) was not statistically significant. A better test of differential impact is whether

these two differences differ from one another. This difference, 5 percentage points, is not statistically significant. The same relationship holds for the two GPA groups. While the results hint at a larger impact for whites and A students, we cannot conclude that this is indeed the case. The lack of statistical significance is likely due to the small number of participants in our study.

Experiment II. The second experiment took place during the 2002–2003 academic year and used four Web surveys administered throughout the academic year to the class of first-year students. The first-year students were randomly selected into four groups. The first group (A) was asked to take only one survey during the entire academic year, a consortium survey about academic experiences called the Enrolled Student Survey. The second group (B) was asked to take two surveys, an internal survey about campus dining services and the Enrolled Student Survey. The third group (C) was administered the dining and Enrolled Student Surveys as well as a national drug survey called the Core Alcohol and Drug Survey. Finally, the fourth group (D) was administered four different surveys: an internal survey evaluating our new student orientation program, the Core survey, the dining survey, and the Enrolled Student Survey. Table 5.2 shows the experimental design and sample sizes for each group. The orientation survey was conducted in the last two weeks of October, the Core survey in the first two weeks of November, the Dining Services Survey in the last two weeks of November, and the Enrolled Student Survey in March.

The surveys were conducted in a similar manner, with students notified via an e-mail that contained a hyperlink to the survey Web site. The first, third, and fourth surveys consisted of an initial e-mail and two reminder e-mails to nonrespondents. The Core survey was conducted anonymously, with an initial e-mail and only one reminder to all members of the sample.

Table 5.2 presents the response rates for the orientation, Core, dining, and Enrolled Student Surveys. It is fairly clear that the response rate drops for an experimental group if there was a previous survey. Group A took only one survey and had a response rate of 60 percent. Group B took two surveys and the response rate dropped from 68 percent to 63 percent. Group C took three surveys and the response rate fluctuated from 54 percent to 58 percent to 47 percent. Group D dropped from 70 percent to 44 percent to 46 percent to 47 percent. One inconsistent data point for Group C is the dining survey response rates. Here the response rate is higher for the second survey, most likely due to the high salience of the dining survey among students. Also note that the decline tends to level off at the mid-40 percent range for Group D. It is reasonable to speculate that there are “hard-core” survey responders who will not be fatigued by multiple surveys, hence the impact of survey fatigue may not be strictly linear.

By looking at the first diagonal in Table 5.2, which contains the response rates for the first survey administered to each experimental group,

Table 5.2. Experiment II: First Year Student’s Response Rates

Group	Response rate (%)				
	Sample size	Orientation (late Oct.)	Core (early Nov.)	Dining Services (late Nov.)	Enrolled Student (March)
A	144	—	—	—	60
B	144	—	—	68	63
C	144	—	54	58	47
D	144	70	44	46	47
Total	576	70	49	52	54

it is possible to see the effect of each survey’s salience and timing on survey response. If survey attributes did not influence survey participation, the response rates across this diagonal would not vary, but they do. Looking at the first survey response rate for Groups A through D, the results are varied: 70 percent (orientation), 54 percent (Core Survey), 68 percent (Dining Services), and 60 percent (Enrolled Student). However, by looking at the timing and salience of the specific surveys, a pattern does emerge.

Although the expectation might have been that the administration of the first survey to each group would produce equivalent responses rates, it is understandable that new students might be more inclined to fill out a specific survey about their orientation experience than the other students were to fill out the first survey they received. Similarly, it is not surprising that first-semester students would answer a short, specific, and salient survey about dining services at a greater rate than second-semester students responded to a long, varied survey in March (the Enrolled Student Survey). Thus we see strong evidence that factors such as survey content and timing of administration can affect rates of participation.

Finally, regardless of the number of previous surveys, no survey achieved a response rate below 44 percent. This finding likely indicates that while survey fatigue has an impact on response rates, there may be a hard-core group of responders who will reliably complete our surveys.

Table 5.3 provides more detail about the impact of one and two previous surveys on the Dining Services Survey response rate. The first row of this table illustrates the most linear survey fatigue finding, with response rates decreasing as the number of prior surveys increases from zero to one to two (68 percent, 58 percent, and 46 percent, respectively). Here receiving invitations to participate in two previous surveys lowered response rates in the dining survey by 22 percentage points. The immediately prior survey was the Core Survey, which was long and personally intrusive. Another factor leading to the consistent results is that all three surveys were administered during one semester.

Most of the subgroups (gender, race, and GPA) follow the overall trend and show statistically significant declines in response rates. As in the first

Table 5.3. Experiment III: Dining Services Survey Response Rates

	Dining Services Survey response rates (%)			Impact on Dining Services Survey response rate (%)			Sample sizes		
	No prior surveys (Group B)	1 prior survey (Group C)	2 prior surveys (Group D)	(C-B)	(D-A)	Total impact (D-B)	(B)	(C)	(D)
All students	68	58	46	10†	12*	22**	144	144	144
Gender									
Female	69	58	49	11	9	20*	70	77	72
Male	68	57	43	11	14	25**	74	67	72
Difference				0	5	5			
Race									
White	69	57	52	12†	5	17*	91	88	95
Nonwhite	66	59	35	7	24*	31**	53	56	49
Difference				5	19	14			
1st semester GPA									
A	74	60	57	14†	3	17	61	67	49
B or less	64	56	40	8	16*	24**	83	77	95
Difference				6	13	7			

*Response rates differ significantly at the p .05 level.

**Response rates differ significantly at the p .01 level.

†Response rates differ significantly at the p .10 level.

experiment, there are no significant differences in declines between subgroups. Again, it is likely that the small number of participants had an impact when the findings were in the predicted direction but not significant.

Table 5.4 provides additional details about the impact of one, two, and three previous surveys on the Enrolled Student Survey response rate. This, the most complicated table, provides the most nuanced view of survey fatigue. The column measuring total impact (D – A) is consistently in the predicted direction and statistically significant. Here we see a significant decline, but less than in the previous table. Two things are worth noting in the first row of the table. First, the decline is not as large as in the previous table. This may be due to the administration of this survey in the spring, while all previous surveys were administered in the fall. It could be that as time passes, the impact of previous survey administrations tends to wear off. Second, the decline appears to level off after the second survey. It may be that there is a hard-core group of survey “cooperators” who are relatively unaffected by multiple survey administrations.

The column measuring the impact of the Dining Services Survey (B–A), however, is not significant overall and significant only for one subgroup (males). This may be attributed to the short, specific, and salient nature of

Table 5.4. Experiment II: Enrolled Student Survey Response Rates

	Enrolled Student Survey response rates by group (%)				Differences in response rates between groups (percentage points)				Sample sizes			
	No prior surveys (A)	1 prior survey (B)	2 prior surveys (C)	3 prior surveys (D)	(B-A)	(C-B)	(D-C)	Total impact (D-A)	(A)	(B)	(C)	(D)
All students	60	63	47	47	3	16**	0	13*	144	144	144	144
Gender												
Female	71	63	55	53	8	8	2	18*	77	70	77	72
Male	46	62	39	42	16†	23**	3	4	67	74	67	72
Difference					24*	15	5	14				
Race												
White	60	66	50	55	6	16*	5	5	89	91	88	95
Nonwhite	60	57	43	33	3	14	10	27**	55	53	56	49
Difference					10	2	15	22†				
1st semester												
GPA												
A	66	74	51	59	8	23**	8	7	59	61	67	49
B or less	55	54	44	41	1	10	3	14†	85	83	77	95
Difference					9	13	12	7				

*Response rates differ significantly at the $p = .05$ level.

**Response rates differ significantly at the $p = .01$ level.

†Response rates differ significantly at the $p = .10$ level.

the Dining Services Survey. There is also no effect when looking at the impact of the orientation survey (D–C), which may possibly be attributed to the specific and salient nature of that survey. Finally, the greatest contribution to the overall difference comes from the Core Survey (C–B), which as stated above is long, although specific, but intrusive in nature.

Most of the subgroups (gender, race, and GPA) follow the overall trend and show statistically significant declines in response rates. Looking at differences between groups, there is some evidence that survey fatigue may have a differential impact among students. We can see that survey fatigue affected nonwhites more than whites, but only at an alpha level of .10. Again, it is likely that the small number of participants had an impact where the findings were in the predicted direction but not significant.

Conclusion

Although the demand for student surveys is growing, little research examines the impact of survey fatigue on response rates. Will administering multiple surveys to students eventually result in less cooperation? On the basis of the research presented here we would answer with a qualified yes. Multiple surveys do appear to suppress response rates. Yet the impact of multiple surveys is not linear. Our results indicate that survey fatigue may have the biggest impact on surveys conducted back-to-back. Surveys conducted in a previous semester may not affect response rates, or the impact may be minimal. Similarly, the impact may not be strictly linear and instead may level off over time.

Some of the results given here are obscured by salience effects as well as by timing effects. Clearly some surveys interest students more than others, and it may be that these surveys do not cause as much survey fatigue as less relevant surveys do. More research in this area is needed.

Experienced panel researchers write, “After cooperating for what can be some years of a panel, respondents may become bored or uninterested in taking part any further or *simply feel that they have ‘done enough’*” (Laurie, Smith, and Scott, 1999, p. 270, emphasis added). In e-mails we have received from students who were targeted for multiple survey administrations, this feeling appears to be quiet common. Institutional researchers must be careful not to evoke such a feeling among students; otherwise, survey fatigue may become more of a problem and negatively affect future research efforts.

References

- Apodaca, R., Lea, S., and Edwards, B. “The Effect of Longitudinal Burden on Survey Participation.” Paper presented at the annual conference of the American Association for Public Opinion Research, St. Louis, Mo., May 1998.
- Asiu, B. W., Antons, C. M., and Fultz, M. L. “Undergraduate Perceptions of Survey Participation: Improving Response Rates and Validity.” Paper presented at the annual meeting of the Association of Institutional Research, Minneapolis, Minn., May 1998.

- Atrostic, B. K., Bates, N., Burt, G., and Silberstein, A. "Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights." *Journal of Official Statistics*, 2001, 17(2), 209–226.
- de Heer, W. "International Response Trends: Results of an International Survey." *Journal of Official Statistics*, 1999, 15(2), 129–142.
- Dillman, D. A. *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley, 2000.
- Goyder, J. "Surveys on Surveys: Limitations and Potentialities." *Public Opinion Quarterly*, 1986, 50(1), 27–41.
- Harris-Kojetin, B., and Tucker, C. "Exploring the Relation of Economic and Political Conditions with Refusal Rates to a Government Survey." *Journal of Official Statistics*, 1999, 15(2), 167–184.
- Kalton, G., Kasprzyk, D., and McMillen, D. B. "Nonsampling Errors in Panel Surveys." In D. Kasprzyk, G. Duncan, G. Kalton, and M. P. Singh (eds.), *Panel Surveys*. New York: Wiley, 1989.
- Laurie, H., Smith, R., and Scott, L. "Strategies for Reducing Nonresponse in a Longitudinal Panel Survey." *Journal of Official Statistics*, 1999, 15(2), 269–282.
- McCarthy, J. S., and Beckler, D. G. "An Analysis of the Relationship Between Survey Burden and Non-response: If We Bother Them More, Are They Less Cooperative?" Paper presented at the International Conference on Survey Non-response, Portland, Oregon, October 1999.
- Sharp, L. M., and Frankel, J. "Respondent Burden: A Test of Some Common Assumptions." *Public Opinion Quarterly*, 1983, 47(1), 36–53.
- Sosdian, C. P., and Sharp, L. M. "Nonresponse in Mail Surveys: Access Failure or Respondent Resistance." *Public Opinion Quarterly*, 1980, 44(3), 396–402.
- Steeh, C. G. "Trends in Nonresponse Rates, 1952–1979." *Public Opinion Quarterly*, 1981, 59, 66–77.

STEPHEN R. PORTER is director of institutional research at Wesleyan University.

MICHAEL E. WHITCOMB is assistant director of institutional research at Wesleyan University.

WILLIAM H. WEITZER is senior associate provost and dean of continuing studies at Wesleyan University.