

# Participant Dropout as a Function of Survey Length in Internet-Mediated University Studies: Implications for Study Design and Voluntary Participation in Psychological Research

Michael Hoerger, M.A.

## Abstract

Internet-mediated research has offered substantial advantages over traditional laboratory-based research in terms of efficiently and affordably allowing for the recruitment of large samples of participants for psychology studies. Core technical, ethical, and methodological issues have been addressed in recent years, but the important issue of participant dropout has received surprisingly little attention. Specifically, web-based psychology studies often involve undergraduates completing lengthy and time-consuming batteries of online personality questionnaires, but no known published studies to date have closely examined the natural course of participant dropout during attempted completion of these studies. The present investigation examined participant dropout among 1,963 undergraduates completing one of six web-based survey studies relatively representative of those conducted in university settings. Results indicated that 10% of participants could be expected to drop out of these studies nearly instantaneously, with an additional 2% dropping out per 100 survey items included in the study. For individual project investigators, these findings hold ramifications for study design considerations, such as conducting a priori power analyses. The present results also have broader ethical implications for understanding and improving voluntary participation in research involving human subjects. Nonetheless, the generalizability of these conclusions may be limited to studies involving similar design or survey content.

## Introduction

INTERNET-MEDIATED STUDIES HAVE ALLOWED investigators to accomplish a number of research feats, ranging from collecting data on broad public samples of hundreds of thousands of individuals to conducting multicultural research involving specialized or difficult-to-reach groups.<sup>1-5</sup> As compared to traditional laboratory-based research, web-based research often confers a number of advantages, including increased ease of participant recruitment, reductions in research time and expenses, environmental benefits, improved sample heterogeneity, reduced barriers to research access for participants, greater ease of gathering time-sensitive data, and automatization of data entry and basic analyses.<sup>6-12</sup> However, these advantages are accompanied by a different set of technical, ethical, and methodological risks.<sup>13</sup>

Although basic guidelines on core technical and ethical issues are readily available,<sup>1,8,13</sup> important methodological concerns continue to be addressed on a rolling basis. Initial studies evaluating Internet-mediated research methodology

have focused on core psychometric issues, such as reliability, validity, and measurement equivalence.<sup>14-17</sup> For example, in a multinational study of 361,703 participants, Gosling et al.<sup>15</sup> found that the results of web-based studies were fundamentally similar to those obtained from traditional paper-and-pencil measures, with participant response sets having minimal impact. Over time, this line of research examined more specific psychometric issues, such as how participant response speed affects measurement reliability<sup>18</sup> or how the latency between online study phases impacts sample characteristics.<sup>19</sup> Although initial studies have allayed basic psychometric concerns, an essential complication of web-based research is the problem of participant dropout. In particular, as participant dropout increases, the sample completing the study potentially becomes less representative of the recruited population, decreasing the generalizability of the results of the study.

Although many Internet-mediated studies involve administering lengthy surveys to college students, the majority of web-based studies focusing on the specific issue of dropout have been confined to psychoeducational public-health

studies designed to serve community samples (e.g., smoking cessation interventions, depression education, psychoeducation on risky sexual behavior, etc.). For example, researchers have examined how response rates vary among different samples recruited for participation and attempted to improve recruitment strategies to decrease early dropout during initial study phases.<sup>20,21</sup> Additionally, investigators have studied patient factors associated with dropout during the course of multiple-stage longitudinal treatment studies conducted using web-based interventions.<sup>22,23</sup> In addressing dropout concerns, these studies hold important implications for online intervention research. Unfortunately, web-based treatment research represents only a fraction of the vast array of studies being conducted on the web.

In contrast to these applied studies drawing upon community samples, Internet-mediated psychology research more often involves basic research using convenience samples recruited from university participant pools.<sup>12</sup> For example, it is relatively common for participants to be asked to spend in excess of an hour completing multiple self-report surveys of personality and related constructs, often consisting of hundreds of items in total length. To the author's knowledge, no studies to date have closely examined participant dropout during the course of completing basic online survey batteries commonly used in university research. Yet, this type of investigation would have several pragmatic implications. First, being able to estimate dropout rates is important for conducting a priori power analyses, necessary for recruiting samples that will have adequate power for the rejection of one's null hypothesis. Second, substantial non-random dropout threatens the generalizability of results produced using this modality. Third, the nature and course of dropout behavior carries implications for how researchers and participant pool overseers ethically credit and compensate participants for incomplete study participation.

Investigating dropout during the course of completing lengthy online surveys is also important for a broader ethical understanding of voluntary participation in psychological research. Demand characteristics, politeness expectations, obedience to authority, and conformity norms are known to significantly impact behavior, largely preventing any dropout from occurring during the course of standard research studies conducted in person in psychology labs.<sup>3,12,24</sup> Participants are more often freed from these constraints in the web-based setting. The Internet affords participants greater perceived privacy and proximity from the project investigator, perhaps allowing for participation in research inherently more voluntary.<sup>1</sup> If, as studies would suggest, web-based research is largely analogous to lab-based research but with decreased situational demand, then dropout rates in Internet-mediated research studies could arguably be thought of as rough indicators of involuntary participation rates in standard lab-based research studies. The present investigation summarizes existing data on dropout rates across a variety of Internet-mediated survey studies conducted in the university setting in an effort to shed light on these important ethical and methodological issues.

## Method

### *Participants and procedures*

A total of 1,963 college undergraduates from a large Midwestern university participated in one of six survey-based

studies administered online. Studies were administered online via SurveyMonkey.com and ranged in length from 243 to 535 survey items, measuring constructs primarily related to personality and emotion. All studies were approved by a university IRB, and participants were compensated for their involvement with psychology subject pool credit.

**Participant responding.** Each participant response was coded as an "item" completed, regardless of the type of question asked or response format. For example, checking a consent box after reading an online consent form was considered the first item completed. Participants typically then provided various forms of contact or demographic information, followed by completing a battery of surveys, usually with Likert-type response scales. No data are available on participants who failed to respond to a single item, namely those who read but did not complete a consent form.

**Study 1: Valentine's Day I.** Participants completed a number of measures related to personality, dysphoria, and anticipated reactions to Valentine's Day, totaling 243 items. Of the 710 participants, 670 completed the entire study (94.4%), and based on the responses of those providing demographic information, participants were predominantly young ( $M = 19.9$ ,  $SD = 3.6$ ) and female (74%).

**Study 2: Valentine's Day II.** In a study similar to Study 1, participants completed a 311-item battery of measures. Of the 336 participants, 287 finished the study (85.4%). Participants were young ( $M = 18.7$ ,  $SD = 1.0$ ) and the majority were female (67%).

**Study 3: Impulsivity study.** Participants completed a 365-item battery of surveys examining the construct validity of various measures of impulsivity, gratification delay, ego control, and related measures. Of the 90 participants, 71 completed the study (78.9%). Again, participants were primarily young ( $M = 19.4$ ,  $SD = 2.4$ ) and the majority female (63%).

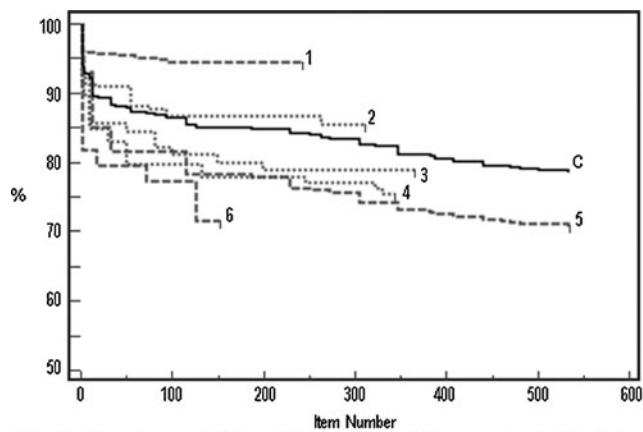
**Study 4: Picture rating study I.** Participants completed a 343-item series of measures related to personality and emotional intelligence, which were compared to predicted and actual emotional reactions to various emotionally evocative images. Of the 118 participants, 89 completed the entire study (75.4%). Most were young ( $M = 19.2$ ,  $SD = 2.2$ ) and female (61%).

**Study 5: Picture rating study II.** In a study similar to Study 4, participants completed a 535-item battery of surveys. Of the 621 participants, 440 completed the entire study (70.9%). Yet again, participants were generally young ( $M = 19.6$ ,  $SD = 2.9$ ) and female (71%).

**Study 6: Birthday study.** Participants with November/December birthdays completed a 152-item study examining personality differences in emotional reactions to their own birthdays. Of the initial 88 participants, 63 completed the entire battery of surveys (71.6%). They consisted mainly of young ( $M = 20.5$ ,  $SD = 3.3$ ) females (76%).

## Results

Survival functions were estimated using the Kaplan-Meier estimator. Figure 1 shows the rate at which voluntary par-



Note. % = Percentage maintaining participation. Curves 1-6 represent each of six studies, respectively, and curve C represents a composite Kaplan-Meier curve.

FIG. 1. Voluntary participation rates as a function of survey items completed.

participation levels declined as a function of the number of survey items completed, with curves 1-6 representing each of the six individual studies. Curve C represents a composite function, based on the Kaplan-Meier estimator, which factors in the concept of "censoring" (that participants completing the entirety of brief studies could have been estimated to complete an additional number of items given the chance), as well as sample size differences across studies. Essentially, the Kaplan-Meier estimator allows for the best estimate of dropout behavior, while extrapolating from studies with different survey lengths. Based on the Kaplan-Meier estimator, 6% of participants dropped out immediately after providing consent (item 1), and a cumulative total of 10.1% of participants discontinued within the first dozen responses. After this initial item set, dropout decelerated significantly, with a cumulative 13.2% of participants having dropped out after 100 responses, and 20.7% after 500 responses.

Given the regularity of dropout behavior, a regression equation was developed, using the number of survey items completed to predict dropout rates. A simple linear equation adequately characterized the data,  $R^2 = 90.5$ ,  $p < 0.001$ , and various non-linear equations (logarithmic, exponential, geometric, and quadratic functions) did not account for significant additional variance. The linear equation states:

$$\hat{Y} = 0.10 + 0.02 \times X$$

$$\text{Predicted Percentage Dropout} = 10\% + 2\% \\ \times (100 \text{ items completed})$$

That is, 10% of participants can be expected to drop out automatically, with an additional 2% of participants discontinuing per 100 survey items administered. Post hoc analyses were used to determine if dropout behavior varied as a function of age or gender, though no statistically significant or practically relevant differences were noted.

## Conclusion

Internet-mediated research has allowed investigators to conduct studies of enormous power, often with diverse groups and while consuming resources more efficiently than

lab-based studies.<sup>1-5</sup> Past researchers have already addressed initial technical, ethical, and methodological concerns.<sup>1,8,13-17</sup> The issue of participant dropout has important ramifications for power, generalizability, compensation considerations, and our understanding of voluntary participation. Although many web-based studies involve college students completing a lengthy battery of online surveys, the few existing studies of participant dropout in the context of web-based research are confined to multistage public-health intervention studies.<sup>20-23</sup> The present investigation was the first known to report on participant dropout behavior within the context of prototypical web-based survey studies involving student samples. The current results have encouraging implications for Internet-mediated studies and also have ramifications for voluntary participation in standard lab-based research.

In particular, most participants who discontinued did so early on in the studies. Approximately 10% of participants can be expected to drop out almost immediately, with an unremarkable 2% dropping out for every 100 items of survey content. These results are directly relevant to conducting a priori power analyses. For example, if a researcher needs 400 participants to achieve a desired level of power and a survey is 300 items in length, the present findings indicate that 477 participants should be recruited for the study.

Beyond facilitating power analyses, the present results have encouraging implications for Internet-mediated research from both an ethical and methodological perspective. Methodologically, the results are advantageous for handling incomplete data for individual cases; most participants who drop out do so early and can be discarded entirely from analyses, whereas very few drop out late and necessitate the imputation of data. Ethically, although past investigations have highlighted the difficulty of participants in understanding consent forms,<sup>25,26</sup> the present results are more encouraging. Participants who voluntarily discontinued typically choose to do so immediately after reading the consent form or after answering a rather small subset of questions. That is, the decision to drop out of lengthy survey studies is based primarily upon the initial information available, rather than fatigue, unanticipated survey content, or potential harm experienced during latter portions of survey studies.

Additionally, the present investigation suggests strategies for facilitating voluntary participation in standard university research conducted in the laboratory setting. These findings make clear that a sizeable subset of university research participants typically prefer to discontinue participation after reading a consent form or completing a few initial items. Unfortunately, during single-session studies conducted in standard university psychology laboratories, dropout is unlikely to occur due to high situational demand.<sup>3,12,24</sup> To improve voluntary participation, researchers with high ethical standards may wish to break up study participation into two phases—one involving the gathering of consent, contact, and demographic information, and the other involving completing survey materials—as this would afford disinterested participants greater opportunity to discontinue voluntarily. Because the bulk of voluntary dropout occurs during the first dozen items completed, researchers who provide early opportunities for discontinuation could foreseeably request that participants complete a rather extensive battery of measures, as the number of survey items completed has a relatively

marginal impact on the desire to maintain participation in research.

The central limitation of this investigation is the restricted scope of studies included in the analyses. Dropout data were gathered from six studies, all conducted at the same Midwestern university by the same principal investigator, primarily involving surveys on personality and mood. Studies involving more controversial topics or complicated instructions could yield higher dropout rates than those reported here. Thus additional research examining dropout rates across more heterogeneous populations and research settings, such as in applied medical settings rather than psychology departments, would be necessary for drawing broader conclusions about the generalizability of these results.

### Disclosure Statement

No competing financial interests exist.

### References

- Hoerger M, Currell C. (in press) Ethical issues in Internet research. In Knapp S, VandeCreek L, Handelsman M, et al., eds. *APA Handbook of ethics in psychology*. Washington, DC: American Psychological Association.
- Hoerger M. Delaying gratification inventory. Unpublished doctoral dissertation, Central Michigan University, 2009.
- Nosek B, Banaji M, Greenwald A. Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics* 2002; 6:101–15.
- Srivastava S, John OP, Gosling SD, et al. Development of personality in adulthood: Set like plaster or persistent change? *Journal of Personality & Social Psychology* 2003; 84: 1041–53.
- Fernández MI, Varga LM, Perrino T, et al. The Internet as recruitment tool for HIV studies: Viable strategy for reaching at-risk Hispanic MSM in Miami? *AIDS Care* 2004; 16: 953–63.
- Hoerger M, Quirk SW, Lucas RE, et al. Immune neglect in affective forecasting. *Journal of Research in Personality* 2009; 43:91–4.
- Birnbaum MH. (2001) *Introduction to behavioral research on the Internet*. Upper Saddle River, NJ: Prentice Hall.
- Fraley RC. (2003) *How to conduct psychological research over the Internet: A beginner's guide to HTML and CGI/Perl*. New York: Guilford.
- Hanggi Y. Stress and emotion recognition: An Internet experiment using stress induction. *Swiss Journal of Psychology* 2004; 63:113–25.
- Murray DM, Fisher JD. The Internet: A virtually untapped tool for research. *Journal of Technology in Human Services* 2002; 19:15–18.
- Nosek B, Banaji M. E-research: Ethics, security, design, and control in psychological research on the Internet. *Journal of Social Issues* 2002; 58:161–76.
- Skitka L, Sargis E. The Internet as psychological laboratory. *Annual Review of Psychology* 2006; 57:529–55.
- Kraut R, Olson J, Banaji M, et al. Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the conduct of research on the Internet. *American Psychologist* 2004; 59:105–17.
- De Beuckelaer A, Lievens F. Measurement equivalence of paper-and-pencil and internet organizational surveys: A large scale examination in 16 countries. *Applied Psychology: An International Review* 2009; 58:336–61.
- Gosling S, Vazire S, Srivastava S, et al. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist* 2004; 59:93–140.
- Lieberman D. Evaluation of the stability and validity of participant samples recruited over the internet. *CyberPsychology & Behavior* 2007; 11:743–5.
- Strickland O, Moloney M, Dietrich A, et al. Measurement issues related to data collection on the World Wide Web. *Advances in Nursing Science* 2003; 26:246–56.
- Montag C, Reuter M. Does speed in completing an online questionnaire have an influence on its reliability? *CyberPsychology & Behavior* 2008; 11:719–21.
- Wesemann D, Grunwald A, Grunwald M. A comparison of different survey periods in online surveys of persons with eating disorders and their relatives. *Telemedicine & e-Health* 2009; 15:751–7.
- Hoonakker P, Carayon P. Questionnaire survey nonresponse: A comparison of postal mail and Internet surveys. *International Journal of Human-Computer Interaction* 2009; 5:348–73.
- Van Horn P, Green K, Martinuseen M. Survey response rates and survey administration in counseling and clinical psychology: A meta-analysis. *Educational & Psychological Measurement* 2009; 69:389–403.
- Jain A, Ross M. Predictors of drop-out in an Internet study of men who have sex with men. *CyberPsychology & Behavior* 2008; 11:583–6.
- Lieberman M. Psychological characteristics of people with Parkinson's Disease who prematurely drop out of professionally led Internet chat support groups. *CyberPsychology & Behavior* 2007; 10:741–8.
- Quirk SW, Subramanian L, Hoerger M. Effects of situational demand upon social enjoyment and preference in schizotypy. *Journal of Abnormal Psychology* 2007; 116:624–31.
- Bhansalia S, Malhotra S, Pandhia P, et al. Evaluation of the ability of clinical research participants to comprehend informed consent form. *Contemporary Clinical Trials* 2009; 30:427–30.
- Mann T. Informed consent for psychological research: Do subjects comprehend consent forms and understand their legal rights? *Psychological Science* 1994; 5:140–3.

Address correspondence to:

Michael Hoerger  
 Doctoral Candidate in Clinical Psychology  
 Department of Psychiatry  
 University of Rochester Medical Center  
 300 Crittenden Blvd, Psych Box  
 Rochester, NY 14642

E-mail: michael\_hoerger@urmc.rochester.edu