Identifying and Avoiding Survey Nonresponse and Noncompliance

Allix Bishop

Faculty Advisor: Dr. Christy Craig

Introduction

In the last five years, researchers have seen a notable increase in difficulty of collecting valid and useful survey data. Respondents, likely attracted to the prospect of receiving prizes or other compensation after survey completion, submit survey responses that are irrelevant to the questions asked, relevant but unusable due to the respondent not identifying as a member of the target population, or comprised of nonsense (such as paragraphs of lorem ipsum text or copy-pasted from other sources). As a result of experiencing a high volume of invalid survey responses to a qualitative survey aiming to research identity narrative formation among Irish women, this project highlights the forms in which invalid survey responses appear and analyzes the ways in which this might be avoided. Previous research has suggested the use of internet protocol (IP) address analysis and attention checks to prevent survey nonresponse as well as responses from non-target populations and invalid responses. For this project, tactics such as attention checks, skip logic, CAPTCHA, and highly specific population-relevant questions were experimented with to lessen invalid responses from both human respondents and bots, as well as to lessen the amount of data cleaning needed to find relevance in responses. Analysis of these methods and a conclusion of suggested best practices to avoid survey nonresponse and noncompliance are offered for future surveying.

Literature Review

As a result of the increasing popularity of web-based survey methods, researchers have found that there are new issues to contend with when attempting to confirm the validity of survey responses and whether survey respondents are truly of the target population for that survey. Survey methods research has previously been primarily focused on comparing traditional and web-based surveys for legitimacy and quality, rather than on how best to determine whether respondents are of the target population for a given survey. Online survey responses do have the benefit of being tied to internet protocol (IP) addresses, which can help ensure that survey responses are in fact valid and different from other responses, but this cannot be solely relied upon to confirm validity because respondents can use virtual private networks (VPNs) to mask their online identities. The issue of insufficient effort responding (IER) has also arisen as a problem within online surveying, wherein respondents complete surveys with minimal effort, resulting in invalid responses . Methods of IER like straightlining, where respondents respond to multiple survey questions with the same answer, and random responding, where participants respond to survey items with answers that are entirely random but may be varied so that the survey response appears genuine. Because survey responses that come from respondents who are not the target population do not actually represent the target population, data derived from them cannot be used for the purpose of answering the original research questions. There are validation tools that can be used to catch invalid responses, but they often do not catch noncompliance, and so they cannot be relied upon. Nonresponse, wherein participants leave survey items or entire survey units blank, is another rising issue within survey research that appears to occur more often in online surveys than in other methods.

Methods

This project began with a focus on a narrow list of specific questions that were designed to prompt respondents to answer expansively, because previous research has found that the way that questions are worded in online qualitative surveys can be more important than it is in other methods, due to an inability to change or edit questions as the survey continues based on respondent answers. The survey was designed in Qualtrics and utilized skip logic to prevent responses from participants who were not part of the target population based on qualifications of age (over 18), citizenhood or identity (Irish), and gender (identifies as woman). It was initially posted on Reddit, in two subreddits: r/mnawesome and r/survey. R/mnawesome is a subreddit dedicated to Irish women, and r/survey is a subreddit dedicated to survey postings.

The initial survey was closed after receiving over 200 responses in a matter of hours. Upon analyzing these responses, it became clear that most of them were from non-targeted populations or contained other types of nonresponse or noncompliance. For the second attempt at the survey, additional skip and display logic were added to narrow down the number of responses entirely; CAPTCHA was added to prevent bot responses, and additional questions were added to further verify respondent fit for the target population. Upon examining previous research for additional suggestions on preventing survey noncompliance, attention check questions were also added to both assist in verifying that respondent fit and to cut down on nonresponse.

For both attempts at the survey, responses were analyzed for nonresponse and noncompliance. Some responses were excluded for duplicating other responses either entirely or variably by item, or for not falling within the target population, or for including item responses that were entirely nonsense (lorem ipsum text) or not truly written by the respondents themselves.

Figure 1 - Attempt #1 Response Totals

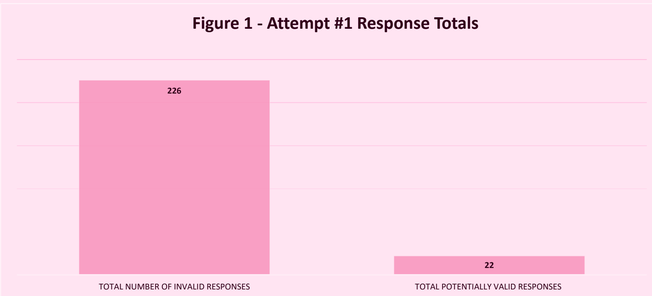
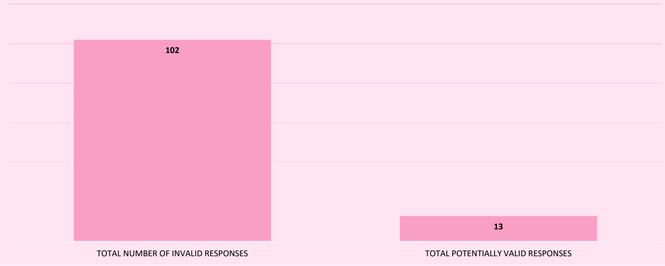


Figure 2 - Attempt #2 Response Totals



Analysis

Figure 1 and Figure 2 display the difference between the total number of responses and the total number of potentially valid responses for each attempt at the survey. While the methods used to limit noncompliance and nonresponse were ultimately useful, very few of the overall responses were determined to be potentially valid. Responses that were excluded varied from containing identical answers to other responses to those that contained answers that indicated the respondent was not part of the target population based on age, gender, or citizenship. These questions, originally included on the survey to function as a way of excluding nontarget populations, also worked well as attention-check questions.

As shown in Figure 3 and Figure 4, both surveys contained a large amount of noncompliance in the form of respondents who were not part of the target population. Unit responses were also deemed invalid if they involved straightlining or random responding, and unit responses that contained items that duplicated other items were also excluded for noncompliance. Because previous research indicated that verifying validity of responses based on IP address was possible, survey responses were also analyzed for duplicate IP addresses and removed.

Some of the unit responses that were excluded due to being identical to other unit responses might have been legitimate, if perhaps the respondent chose to submit their response multiple times in an effort to gain compensation for the survey, but because they could not be verified as individual responses they could not be considered valid for the purposes of the original project. Overall, noncompliance proved a bigger issue than nonresponse in this particular survey.

Figure 3 - Attempt #1 Invalid Response by Type

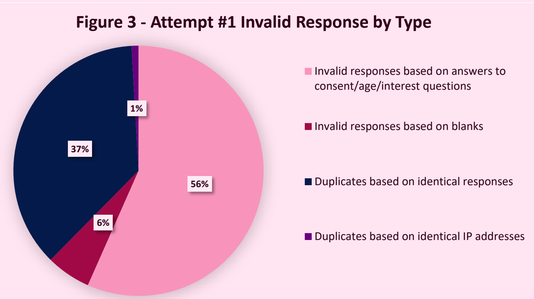
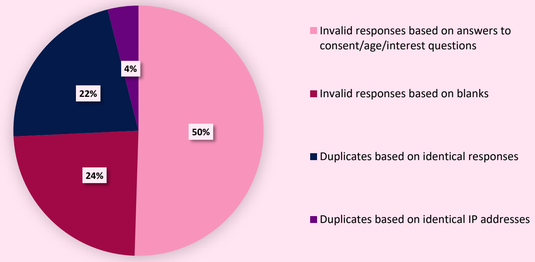


Figure 4 - Attempt #2 Invalid Response by Type



Conclusions

Initially, this project sought to understand the ways in which Irish women form their identities, with an emphasis on the ways that religion and media might change the ways that those identities form. As a result of the overwhelming number of responses indicating noncompliance, the questions of why this was occurring and how to mitigate it arose. Previous research about noncompliance and nonresponse in online qualitative research is limited, and solutions to the issues that noncompliance and nonresponse contribute to (a lack of valid data being the most relevant issue for this project) were available, but not especially effective.

The survey data acquired during the original research process was valuable because it presented a variety of different types of nonresponse and noncompliance to examine. Ideally, the data gathered and analyzed for noncompliance and nonresponse will help in future survey endeavors by providing an example of methods that do not work for qualitative research. Possible future adjustments to the original survey might include the addition of more attention-check or test questions that might help eliminate identical or duplicate unit or item responses. Examining IP addresses for duplicates was useful in the data analysis process, but ultimately was not as useful as previous research has suggested. Without additional ways to verify individual responses, web-based surveys will continue to prove difficult to pull useable data from.

References

[1]Bernerth, Jeremy B., Herman Aguinis, and Erik C. Taylor. 2021. “Detecting False Identities: A Solution to Improve Web-Based Surveys and Research on Leadership and Health/Well-Being.” Journal of Occupational Health Psychology 26(6):564–81.
[2]Silber, Henning, Joss Robmann, and Tobias Gummer. 2022. “The Issue of Noncompliance in Attention Check Questions: False Positives in Instructed Response Items.” Field Methods 34(4):346–60. : [3]DeSimone, Justin A., Alice J. DeSimone, P. D. Harms, and Dustin Wood. 2018. “The Differential Impacts of Two Forms of Insufficient Effort Responding: IMPACT OF DIFFERENT TYPES OF IER.” Applied Psychology 67(2):309–38.
[4]Schonlau, Matthias, and Vera Toepoel. 2015. “Straightlining in Web Survey Panels over Time.” Survey Research Methods Vol 9:125-137.
[5]Čehovin, Gregor, Michael Bosnjak, and Katja Lozar Manfreda. 2022. “Item Nonresponse in Web versus Other Survey Modes: A Systematic Review and Meta-Analysis.” Social Science Computer Review 089443932110562.
[6]Silber, Roßmann, and Gummer, 348
[7]Bernerth, Aguinis, and Taylor, 566
[8]Braun, Virginia, Victoria Clarke, Elicia Boulton, Louise Davey, and Charlotte McEvoy. 2020. “The Online Survey as a Qualitative Research Tool.” International Journal of Social Research Methodology 1–14.
[9]Silber, Roßmann, and Gummer, 347
[10]Bernerth, Aguinis, and Taylor, 566

FORT HAYS STATE UNIVERSITY
Forward thinking. World ready.

Department of Sociology