

The Limitations of Online Surveys

Chittaranjan Andrade¹

ABSTRACT

Online surveys are growing in popularity, perhaps because they are an easy, convenient, and inexpensive means of data collection. Online surveys commonly suffer from two serious methodological limitations: the population to which they are distributed cannot be described, and respondents with biases may select themselves into the sample. Research is of value only when the findings from a sample can be generalized to a meaningful population. When the population addressed by the survey cannot be described, and when the sample is contaminated by respondents with biases, findings from online surveys cannot be generalized and may therefore mislead.

Keywords: online survey, online questionnaire, internet survey, limitations, sample, population

An online survey collects information from people who respond to a form or instrument that is distributed through internet channels. Online surveys have several advantages.¹ They are easy to conduct using free platforms such as SurveyMonkey, Google Forms, and Typeform. They can be quickly created and distributed with global reach; this is important when information is urgently sought, such as in the context of the current Covid-19 pandemic. There are few to no administrative costs. Analysis is usually easy to perform on the same platform at which the survey was conducted.

Online surveys are becoming increasingly popular. There were 1682 PubMed hits for “online survey” (search phrase entered with quotes) in 2016; this num-

ber increased to 1994 in 2016, 2425 in 2017, 2872 in 2018, and 3182 in 2019. On August 15, 2020, the number of hits for 2020 was already 2742; when annualized, this number projects to 4387. In contrast, the growth for the search term “meta-analysis” has been lower, from 17,806 in 2015 to 24,033 in 2019; and “randomized controlled trial” has shown negative growth, from 31,689 in 2015 to 26,285 in 2019. This article explains why, despite their popularity, online surveys may be methodologically unsound.

Generalizing from Sample to Population

In a hypothetical study that we conducted at our hospital in 2019, patients with schizophrenia expressed positive

attitudes toward long-acting injectable antipsychotic medications. This finding truly applies only to the patients whom we studied. There is no assurance that the same finding will be obtained in a different set of patients studied in the same hospital, let alone in patients studied elsewhere in the world and in later years.

Research findings are of scientific value only if they can be generalized. At the very least, it should be possible to generalize from the sample to the population from which the sample was drawn. This can happen only if the sample is representative of the population, which requires two conditions to be fulfilled. The first condition is that the population must be known; it is not possible to generalize the findings of a study to a population that cannot be defined. The second condition is that a valid method of sampling should have been adopted; a method that recruits a sample that is overrepresented for a particular characteristic cannot represent the population.

An Absurd Example

Imagine that we create a questionnaire to study how Covid-19 has affected the mental health of the population. We leave copies of our questionnaire on a park bench. These sheets of paper are

¹Dept. of Clinical Psychopharmacology and Neurotoxicology, National Institute of Mental Health and Neurosciences, Bangalore, Karnataka, India.

HOW TO CITE THIS ARTICLE: Andrade C. The Limitations of Online Surveys. *Indian J Psychol Med.* 2020;42(6): 575–576

Address for correspondence: Dept. of Clinical Psychopharmacology and Neurotoxicology, National Institute of Mental Health and Neurosciences, Bangalore, Karnataka 560029, India. E-mail: andradec@gmail.com

Submitted: 18 Aug. 2020
Accepted: 18 Aug. 2020
Published Online: 13 Oct. 2020



Copyright © 2020 Indian Psychiatric Society - South Zonal Branch

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-Commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

ACCESS THIS ARTICLE ONLINE
Website: journals.sagepub.com/home/szj
DOI: 10.1177/0253717620957496

blown by the wind to different destinations. Most people regard the questionnaire as litter. Some pick it up, look at it, and drop it. A few take the trouble to answer the questions and return the questionnaire. We analyze the data in the returned forms and conclude, perhaps, that Covid-19 has compromised the mental health of persons in our city or country.

These conclusions are not valid because we have no idea at all about the population that had access to the questionnaire, let alone whom the questionnaire actually reached; so we cannot define or describe the population to which the findings can be generalized. The questionnaire may have collected demographic details, but we cannot even generalize the findings to persons with these demographics because we did not select a random sample of persons with these demographics (we cannot assume that persons who responded to the survey were representative of their population demographics).

Sampling bias could have arisen in several ways. For example, the questionnaire would only have reached persons who were downwind and would only have been completed by those who were literate and those who were sufficiently interested in the topic to take the time and trouble to respond. What is the value of findings that can only be generalized to “persons who are sufficiently interested in the topic to respond to it”?

Online Surveys

The absurd example in the previous section mirrors what happens in online surveys. Online surveys are commonly

distributed almost by the wind, through individual emails, mailing lists, and social media platforms. Requests are made by the investigators to forward the survey to other channels. There is no way of identifying, understanding, and describing the population that could have accessed and responded to the survey, and to whom the results of the survey can be generalized.

Furthermore, online surveys are completed only by persons who are literate and who have access to the internet, and by those who are sufficiently biased to be interested in the subject; why else would they take the time and trouble to respond? As an example, patients who were traumatized by a medical procedure are more likely to respond to an online survey about that procedure than patients who underwent the procedure uneventfully. The traumatized patients want to be heard; other patients feel no pressure to respond. Thus, persons with bias are overrepresented in online survey samples. The survey findings are thereby skewed. Because there is no way of knowing the motives of those who responded, there is no way of understanding the extent of bias in online surveys.²⁻⁴

The Best Case Scenario

If the online survey is individually distributed to, say, all the members of a society whose names are listed in a membership directory, we know the population that the survey has reached. If most of the members respond to the survey, we can reasonably conclude that the opinions of the sample are similar to those of most if not all of the population. This is the best case scenario; the sample and

population can both be described, and the sample is probably representative of the population. However, if only a small proportion of members respond, it becomes possible that the results are contaminated by respondent bias and are therefore untrustworthy.

Concluding Notes

The results of online surveys must be regarded as tentative when the surveys are distributed to an unknown audience and when voluntary participation in the survey can result in respondents with biases selecting themselves into the sample.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Evans JR and Mathur A. The value of online surveys. *Internet Res* 2005; 15: 196–219.
2. Wright KB. Researching internet-based populations: advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *J Comput-Med Commun* 2005; 10: JCMC1034.
3. Ameen S and Praharaj SK. Problems in using WhatsApp groups for survey research. *Indian J Psychiatry* 2020; 62: 327–328.
4. Menon V and Muraleedharan A. Internet-based surveys: relevance, methodological considerations and troubleshooting strategies. *Gen Psychiatry* 2020; 33: e100264.