International Journal of Academic Research in Business and Social Sciences
July 2013, Vol. 3, No. 7
ISSN: 2222-6990

Sampling Design, Validity and Reliability in General Social Survey

Wilfred Uronu Lameck
PhD candidate Faculty of Social science, VU University Amsterdam, Lecturer in the school of Public Administration and Management of Mzumbe University, Morogoro, Tanzania

DOI: 10.6007/IJARBSS/v3-i7/27 URL: http://dx.doi.org/10.6007/IJARBSS/v3-i7/27

Abstract

Following the proliferation of the research in the study of social network, the debate on the design and use of General Social Survey has dominated research articles addressing social network question. For example the three prominent authors have propelled a discussion on various anomalies of the findings from different articles on social change which as result raise a question on reliability and validity. Although in all the articles the design on sampling and measurement instrument are featured, the authors deviate from the corollary of sampling. Consequently, the findings have been unfolded and subjected to critical review and comments by others. In order to best understand the basis of this claim, these coalitions and competing argument of the authors have been brought to order. Using the lens of the findings presented by the authors, the review discloses the dangers of ineffective sampling and measurement instrument which consequently bear on lack of reliability and validity of study findings.

Keywords: General Social Survey, Sampling Design, Reliability and Validity

I. Introduction

Sampling design, reliability and validity are among the criteria used by social scientists to evaluate the quality of social science research. In this review, sampling design is defined as means of selecting primary unit for data collection and analysis which are appropriate for a specific research question.(Handwerker 2005) Reliability on the other hand refers to consistent of research instrument. This means that, the research instrument should produce the same score over repeated measures.(Croker & Algina, 1986) while validity means the extent to which the test measures what it claims to measure.(Gregory, 1992) On the contrary, many articles which use surveys as their design have been proved to deviate from key principles of sampling which consequently affect the quality of research findings. For example the article by McPherson, Smith-Lovin and Brashers (2006) replicated the GSS question used in (1985) to conduct another study in (2004) concerning change of core social network structures based on the social network question. In this study, the response rate indicated that, number of people who did not discuss important matter with anybody tripled and the mean network decreased for about a third from (2.94%) in 1985 to (2.08%) in (2004) which indicated larger social change in the past two decades. In response, (Bearman and Parigi, 2004) GSS study on social network

for *people who talk important matters* found out that, half of the people who do not talk, have nothing to talk about while others have no person to talk to, and suggested that, gender difference in network composition might be an artifact. Similarly, Fisher (2009) strongly criticized this study based on the fact that, the question used in the survey to measure the size of respondents produced results which were so inconsistent with other data and ignore what happened in America between (1980s) and (2004).These arguments validate the discussion on sampling design, validity, and reliability in these studies. It further discloses the need for social scientists to understand how to avoid such negative impact of ineffective sampling design, reliability and validity of the study findings. This review therefore, aims at among other things, disentangling the views of these authors and suggesting further improvement on the design of sampling and measurement instrument. In doing that, the competing argument of the authors are presented and discussed in the light of sampling design, reliability and validity of their study findings. Finally, the discussion is brought to order by trying to assess the direction of General social Survey based on sampling, reliability and validity as criterion for assessing the quality of research.

(II)Sampling design in General Social Survey

Like other social science research, sampling design in GSS is common preferred method of selecting study units out of the population. Although the theory states clearly the criteria of judging a good sample such as free sampling error, certainty, triangulation by using similar methods across different data sets, avoidance of exclusion and non response rate Firebaugh(2008), sampling in general Social Survey deviate from some of these key principles. For example, in GSS sampling from the population is picked at random. As a result, in most cases, it does not represent the characteristics of the population due to sampling error and lack of reality check. Second, in GSS sampling design does not pay attention to the use of similar methods of collecting data in different data sets. Consequently, lack of repetition of data collection methods in different data sets results to unreliable data. This implies that, lack of triangulation (collecting information from different categories of sample) affect reality check. Third, design of sampling in GSS is limited by exclusion error. In this regards, the move from population to sampling process excludes other members of populations, for example non institutionalized members of the populations exclude people who have no access to phone in the survey. Fourth, unit non response is another problem which is a consequence of missing data. For example in our case, non response rate of (2004) General Social Survey was very high which as a result affect variance and association of variables. This includes limitation to generalization of sample to populations which may also cause sampling error because as a sample becomes less representative, the error margin increases. In this case, the increase of percentage of non response rate to 25% of (2004) GSS affected the reliability of the sample and generalization of study findings on social network question to population. Lastly, GSS lacks larger sample and better longitudinal design including repeated cross sectional design which is more preferred for sampling design because it accommodate population characteristics and makes the sample more representative.

International Journal of Academic Research in Business and Social Sciences
July 2013, Vol. 3, No. 7
ISSN: 2222-6990

(III)Reliability and Validity of General Social Survey questions on social network

The design of research instrument in social science research determines the quality of research findings. Although the research instrument in GSS of (2004) was clear, the question used in the survey to measure social network did not pay attention to reliability because the wording and arrangement of questions was not properly done, for example the question on organization membership was in the wrong sequence and it required much details which influence low response rate followed by missing data. Fischer (2009)Second, the question on social network lacks external validity, for example the results were so inconsistent with other data and they ignore what happened in America in (1980s) and (2004) regarding social change. In the same vein, it was discovered that, forty one cases in that survey were wrongly coded and have been corrected but does not suffice the conclusion which was given in (2004).Fischer (2009) Third, the initial results, in question (128) about *num given* and number of respondents were coded zero which affect association of variables. In addition, the GSS of (2004) varied greatly in isolation with other studies, for example the isolated in (1985) GSS was (8.1 %) and in (2004) GSS was (22.6 %). This represent a big social change which is not socially realistic because the scale of change suggested threefold increase which is impossible in social network . Fischer (2009) Fourth factor is triangulation. In this case, the GSS question was not used to collect data from different sources of data sets or different categories of sample which could explain the validity and reliability of the question on social network. For example, the questions might be used in sample categories with different demographic characteristics. Lack of triangulation therefore raises questions on whether the questionnaire used would produce the same results in different data sets. Fifth, the check on incomplete data by checking the variance and correlations between variables was not properly done during coding. This results to compilation of data with some data missing and therefore the findings were invalidated. (Fisher 2009)
Sixth, external checks by comparing the findings with previous research and internal check to see if the question capture what it intends to measure was not properly done. Therefore GSS of (2004) lacks internal validity because it contradict the (1985) GSS percentages on non response in social network question.(Fisher 2009) Seventh, the four measures of social involvement concerning isolated and non isolated and alternative measure which is different from name isolation is not precise measure of the network and it also contradict other data. For example the question on *do you have any friends you feel close to?* in GSS of (1998) was 9% who said "no" which is similar to (1985).On the contrary , about one third of (2004) isolates by McPherson et al (2006) was 25 % and with *social evening question* does not provide evidence for network shrinkage. Fischer (2009) Lastly, the GSS conducted international survey program with the question that comes closer to *num given* which reads *suppose you feel depressed who would you turn first for help*? in (1986) the response to "no" was 2%.In (2002) the response to the same question was (4 %) and in (2000) to (2004) another survey indicated that it was only (2 %) which had no annual contact. All this information makes it difficult to reconcile with (25 %) of GSS in (2004) by McPherson et al (2006) and Fischer (2009)

(IV)The direction of research on social network question

To the larger extent, the existing arguments of the authors provide a general picture on the reliability and validity of GSS. Although some of authors have come up with critical comments on this arguments to their counter parts, it would be fair to say that Fischer (2009) argument is valid since he has critical argument on reliability and validity of GSS of McPherson ,Smith –Lovin & Brashears(2006). Their report on the shrink of social network in America by using (2004) GSS is challenged based on external validity and reliability. This section is dedicated to discussion on McPherson ,Smith –Lovin and Brashears(2006) concerning GSS question, followed by critiques by fisher (2009). To start with McPherson, Smith-Lovin & Brashers (2006),the findings indicated that, they replicated the questions of GSS to assess the change of core social network structures based on the social network question. In their survey, the results of the question on ***The number of people who did not discuss important matter with anybody*** tripled and the mean network decreased for about a third from (2.94) in GSS of (1985) to (2.08) in GSS of (2004) and both kin and non kin were lost. On contrary, GSS of (1985) model of respondents had three confidents of which the lost of non kin leads to more confidents network on spouse and parents, few contacts and voluntary associations for neighborhoods. Some changes reflects the demographic of the U.S. populations, decrease of educational heterogeneous and ratio heterogeneous increase. The survey revealed that, there has been larger social change in the past two decades and the number of people who have important matter to discuss with others and alternative discussion partners has decreased in the community and neighborhood ties. This implies that, one is more connected through larger ties than to family members. Although they tried to establish some reason for this pitfalls based on demographic shift of American discussion on network and social environment surrounding most of American, their study has being challenged based on contraction and lack of validity and reliability.

Why Fisher (2009) might be right about critiques of the GSS of (2004)

First of all, according to fisher (2004) GSS in our case is criticized based on non response rate. In this survey , the respondents provided few names almost one third compared to respondents of (1985) GSS on the same question .This includes, the percentage of *no names at all* which increased from 10% in GSS of (1985) to (25%) in 2004 GSS. This implies huge increase which is not valid in social sciences and hence raise a doubt for example, respondents with *no names* at all double from (1) in (10) to (1) in (4) which contradict other findings. Second, the question used in the survey to measure the size of respondents produced results which are different from other data and ignored what transpired in America in (1980s) and in (2004).In addition, the forty one cases which were discovered wrongly coded and corrected is an indication that, the findings from such survey can not provide a reliable conclusion. Third, GSS (2004) *question (128)* about *num given* and number of respondents coded zero contradict the (1985)& (1987) survey. All of these changes including the scale of change suggested is against normal social change. Fourth, four measures of social involvement create a doubt for example, *isolated and non isolated* and alternative measure which is different from name *isolation* may not be precise measure for network isolation. This includes the contradiction with other data for example, the question on *do you have any friends you feel close to*? in GSS of (1998) .The (9%) response was *no* option, which is similar to GSS of (1985) .Unlike others, about one third of

GGS of (2004) isolates or *no* option was (25 %).Fifth, the GSS on international survey program with the question that comes closer to *num given* which reads, *suppose you feel depressed who would you turn first for help?.* In 1986, the response in this questions to *no* option was (2%) and (2002) was (4 %)and in (2000) to (2004) in another survey indicated that it was only (2 %) which had no annual contact. This raises a reconciliation problem with (25 %) increase in (2004).

Lastly the (2004) GSS contradict other survey for example, it reveals the (23%) on the question *on friendly interested* in GGS of 2004.This was coded zero compare to 6 and (3.7) in (1980s) survey while in education attainment of 16 percent of respondents with postgraduates were coded as giving no names .In (1980s) only (2) respondents were coded as zero while(22 %)married women gave no names compare to (5 %) in (1985).(Fischer, 2006)

(V) Improvement on sampling

First of all, the GSS sampling can be improved by increasing representation through inclusion of sample selection which represent population characteristics . Since sampling is done at random, the use of larger sample and where possible larger longitudinal design and repeated cross sectional design will be appropriate. Second, non response rate of the sample can be improved by redesigning a questionnaire which reflect sample characteristics. For example, lack of response in (2004) GSS was observed to be related to type and design of question in the questionnaire. Therefore, the questionnaire should correspond with kind of sample used in the study. Third, training of interviewer can also be used as a strategy to increase response rate. It is not mention directly but the knowledge of the interviewer on the questionnaire design is important because it increases the quality of questionnaire and hence increasing the response rate. Fourth, sampling error can be reduced by taking larger sample. The larger sample in this regards help to accommodates all characteristics of population and therefore increase representativeness of sample. Fifth strategy is to reduce total uncertainty by replication and triangulation. This implies collection of data from different sources by using different methods or using one method to collect data from different data sources. In this regards, the collection of data in GSS can be improved by comparing the data collected from different samples within the same populations .Sixth strategy is to avoid exclusion error. For example the use of telephones may exclude those who have no access to telephones during data collection and hence increasing non response. Alternatively, the emphasis would be on the use of face to face administered questionnaire which accommodate all members of the population .Lastly ,Gender differences must be considered in sampling for example the GSS on the question of *people who talk important matters,* revealed that male discuss important matters with their fellow men than they do with their spouses. Therefore, it is suggested that gender difference in network composition might be an artifact of data collection. (Bearman and Parigi 2004)

(VI)Improvement on questionnaire design

First of all, it is important to note that, the wording of the question is important in the questionnaire and if not properly done it is likely to affect the response rate.(Firebaugh 2008)For example in GSS of (2004) wording of question on the organization membership affect

the response rate as it requires much details which is different from (1985) GSS. The design of the organization membership questions in the beginning is alleged to have caused low response rate as compared to the following question which had high response rate in (2004) GSS. This is because, the question was more specific and more demanding to respondents including the need for address which made respondents reluctant to give names (Fisher 2009).Therefore the questions which requires too much details should be avoided especially in the beginning of the questionnaire.

Furthermore, general contents and ordering of the question is important because the contents of the questionnaire may influence different response. For example, the puzzle on *important matter topic which is* argued to be so broad to potentially *unimportant* . (Bearman and Parigi 2004) For example some people do not report to talk anything to anybody because they have nothing to talk or someone to talk to. In this question, the findings revealed that half of the people reported not to talk anything, are not isolated but have nothing important to talk about while another half have unimportant matters like the state of the economy, the failure of the space program and moral decay but have no body to talk to. (Bearman and Parigi 2004)

Moreover, the preceding questions can influence response rate for example, in GSS of (2004) 15 % of respondents who reported to belong to *one* , *two* or *more organization* gave name to *no one* in the *num give* and five times as many of respondents claimed no confidants in GSS of (2004) which was not plausible. Finally, the design of question in GSS should also reflect *what people talk about* and not only *important matter question*. GGS instrument gives little attention to what people talk about. Knowing what people talk about, can help to understand the importance of the matter that people talk about. While some people assume that, the important matters discussed are related to important matters which generate the substantive outcome such as borrowing money, learning about jobs, others assume that *people talk about important matter* to people who are important to them. (Bearman and Parigi 2004)

(VII)Conclusion

Therefore sampling, validity and reliability seem to be reliable criteria for assessing the quality of research. Although many researchers are aware of the importance of their importance, very few of them abide to the key principles guiding the application. Consequently, most of their research lack high quality design especially on sampling and research instrument which as a result affect the quality of research results. Following this flaw, which includes among other things, unrealistic contradiction of research results with existing findings, the critical reviewers have raised a question as to whether the findings from such studies can be generalized to population or test the existing theory. The proper design of research which ensures representativeness of sample, triangulation of sample and research instrument, wording and proper design of contents of measurement instrument can to larger extent reduce such flaws and increase the quality of research findings in social research.

Acknowledgement

I wish to extend my sincere gratitude to Dr.Renne Bekker, Assistant Professor in the faculty of social science of VU University in Amsterdam for his guidance in survey design and my fellow PhDs student who encourage me to rewrite and public this review article.

Corresponding Author

Wilfred Uronu Lameck
PhD candidate Faculty of Social science, VU University Amsterdam
Lecturer in the school of Public Administration and Management of Mzumbe University, Morogoro, Tanzania
E-mail:wilfredlameck@yahoo.com

References

Bearman,P, Parigi P (2004) Social Forces ,University of Carolina Press, 83:535-557
Croker,L & Algina,J(1986) Introduction to Classical and Modern test theory, Philadelphia: Harcourt Brace Jovanovich college Publishers.
Fisher C (2009) The 2004 Findings of shrunken Social Networks: An Artifact? *Am Sociol.Rev*.74:657-669
Firebaugh G (2008). Seven rules of Social Research ,Princeton University press, UK
Gregory,R.J.(1992) Psychological testing: History ,Principles and applications, Boston, Alln and Bacon
Handwerker,P.W(2005) Sample design, *Enclopedia of social measurement* 3:429-436
McPherson,Smith-Lovin, Brashers (2006). Social Isolation in America: Changes in core Discussion Networks over Two Decades, *Am. Sociol.Rev*. 71:353-375