



the **LEADERSHIP CENTER**

presents
the

LEADERSHIP & ORGANIZATION EXCELLENCE SERIES

**Nonprofit Management Inquiry:
Addressing Validity and Bias
and Survey Research Tool Design or Selection**

author:
M. Louise Walters

The Leadership Center
912.349.7989
Savannah, GA 31406 USA

© The Leadership Center

**Nonprofit management inquiry:
Addressing validity and bias
and survey research tool design or selection**

by:

M. Louise Walters

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

Capella University

School of Public Service Leadership

May 14, 2015

Introduction

In the last half century, there has been major growth in the interest in and focus on nonprofit organizations. Research is one of those interest areas. Because of growth in numbers and increased interest, in 1973, the Commission on Private Philanthropy and Public Needs was established by the United States government. The Commission was charged with studying philanthropy and the private sector (Shier and Handy, 2014). From that point on, the number of nonprofit organizations grew, and research about the nonprofit sector did, as well. Today, over 1.5 million nonprofit organizations are recognized by the United States Internal Revenue Service (IRS, 2015).

This paper begins with a definition of research from current day and partners it with a definition of research from 1918; the definitions are not dissimilar. The definition of a survey follows. The paper continues with a discussion of threats to internal and external validity that are inherent in survey research designs. The constraints and limitations in selecting a research tool for nonprofit management inquiry are analyzed. The paper looks at types of bias associated with sampling in survey research and the ways to control bias in experimental designs.

Definitions

Research. In the world of research, inquiry is the process that leads to discovery. Research is an invaluable part of humanity's search – and thirst -- for knowledge. Research is a process of study and investigation, of trial and error, and of efforts to establish facts and to reach conclusions (Research, 2015). Research also shows (and verifies) the paths taken and the steps used to collect and to analyze information. Ogden (1918) summarizes it well: Research is the discovery of truth.

Surveys. Surveys are a method of facilitating discovery. Survey results can prompt change or future action. Surveys are a way of collecting data from interviewees that represent a specific population. Asking specific questions provides information on the attitudes, opinions and beliefs of the interviewee. Survey questions can also provide demographic information. Nonprofit organizations undertake surveys for many reasons. Some of those reasons include the desire to evaluate existing programs, the need for input in the creation of a strategic plan, and the desire to measure or to analyze opinions with the anticipation of making future policy or operational decisions.

Threats to Internal and External Validity

Inherent In Survey Research Designs: A Discussion

In the social sciences, research begins with broad concepts that become distilled into research designs. Research designs are precise and address either quantitative, qualitative or mixed methodologies (SOPSL, 2015). Experimental, quasi-experimental and non-experimental are the types of research designs that comprise quantitative methodology. As non-experimental research, surveys are commonly used vehicles that test linear relationships and also analyze statistical differences or statistical independence. There are four types of validity: statistical conclusion validity, construct validity, internal validity and external validity. Non-experimental research includes internal and external validity (SOPSL, 2015).

There can be many things that impact research validity. The universal threats to internal and external validity are:

Conclusion of the Researcher. Is what the researcher discovers correct or incorrect?

Independent Variable. Did changes in the independent variable cause the variation in the dependent variable, or was there another reason or impact that caused the change?

Other Causes. Did the dependent variable change due to the variation in the independent variable, or was there another reason or impact that caused the change?

A Discussion of Internal Validity

Internal validity addresses cause and effect. In other words, did 'A' cause an impact that effected 'B'? Internal validity is the process by which a determination is made as to whether there is a causal relationship between or among two variables. Internal validity can be defined as controlling extraneous (peripheral, inessential or minor) variables, leaving one isolated variable (the independent variable) which can be manipulated by the researcher. Internal validity is the ability to say that the only variable being studied is the one that is the cause of the result. Considerations of internal validity include the design of the study, which variables were measured and which other variables were not, how the variables were measured, and what operational definitions were used. Internal validity assesses the credibility or trustworthiness of the data collected from participants (SOPSL, 2015). Internal validity also addresses the level of confidence the researcher has that the change in the dependent variable was not impacted by other factors.

An example of internal validity is the process by which a manager seeks to determine to what extent a person's salary is the result of -- or is caused by -- productivity. In this example, the productivity is the independent variable and the salary is the dependent variable. All the other variables that could possibly impact the salary (such as motivation, or personality, or years with the organization, or number or types of promotions) are set aside.

Survey Research Design: Threats to Internal Validity

Campbell and Stanley (1963) list the effects that can jeopardize internal validity of research. Some of those factors include: maturation, confounding, history, testing effects, instrument effects, and mortality.

Maturation. Maturation of the participants occurs when they grow older, become more fatigued, grow wiser or become more experienced. For example, responses to a survey of the members of the sophomore class at Capella High School about where they plan to attend college will probably be different than those responses by the same class surveyed two years later when they are soon-to-be graduated high school seniors.

Confounding. This is a risk factor impacting an outcome. An example is the relation between eating smart, exercise, and obesity. Obesity is related to each variable. If we were trying to analyze the relationship between eating smart and obesity, exercise could be the confounder. Simundic (2013) points out that confounding effect occurs in nearly every study, so every effort should be made to identify it and minimize it. Peer reviews are a way to decrease the instances of confounding.

History. History includes something that happened that was not anticipated. Did that event affect the dependent variable? For example, responses by owners of ocean front property to a survey about quality of life will no doubt be different than responses to the same quality of life survey by the owners of the same ocean front property one year after a new oil drilling platform is built within their line of sight.

Testing Effects (or Repeated Testing). Testing effects can arise when repetitive testing causes changes in the results. For example, what if we look at responses by members of an engineering class to a test about time and motion, and we compare those results with the results of the same time and motion test conducted in the same class on the following day? The results

of the second test will be higher because the class members were already familiar with the questions because they had taken the same test on the previous day.

Instrument Effects (or Instrument Change). This includes influences or changes to the survey instrument (in other words, changing – however slight the change may be -- the language of a questionnaire) or changes to the scorers or administrators. For example, assume a videotape was made of an Olympic ice dancing competition. Then, a panel of judges views the videotape and is asked to score the performance. Then, a different panel of judges views the videotape and is asked to score the same performance. If the panels produce different scores, that is likely due to the difference in scorers.

Mortality (or Differential Attrition). For example, on Monday, we survey members of a water aerobics class that meets every Monday, Tuesday and Wednesday. If two people drop out of the class on Tuesday, and if we survey the class again on Wednesday, we would expect different responses to the survey because of attrition in the population.

A Discussion of External Validity

External validity addresses the extent to which the results of a study can be generalized to a population. Were the study results representative of only the population that was studied, or are the results representative of (or can the results be extended to) the larger population?

Campbell and Stanley (1963) identify factors that can influence external validity (or representativeness). Those factors include the reactive or interaction effect of testing, interaction effects of selection bias and the experimental variable, interaction effects of experimental arrangements, and multiple treatment interference. For example, assume a researcher wants to study a new blood pressure drug. The study is only conducted on men over the age of sixty-five who are overweight and who smoke. The study does not include men that are age sixty-five or

younger. The study also does not include any women. The results of the study are valid for the population of the study (men over the age of 65 who are overweight and who smoke). However, the results of the study cannot be generalized (extended) any farther because the study did not include any women, nor did the study include men that are age sixty-five or younger. Likewise, the survey did not include men over the age of sixty-five who are not overweight and smoke.

Survey Research Design: Threats to External Validity

Threats to external validity include small sample size, interaction effects of experimental arrangements, interaction effects of selection bias and the experimental variable, and multiple treatment interference.

Small Sample Size. A small convenience sample size (perhaps less than thirty respondents) may not provide an accurate representation of the larger population group. An inaccurate representation will taint the survey results.

Interaction Effects of Experimental Arrangements. For example, time poverty (not having enough time) is a common complaint. Studies that require a large block of time from respondents may discourage respondents from agreeing to take the survey.

Interaction Effects of Selection Bias and the Experimental Variable. The location where the study is being conducted can have an impact on the study. An example is conducting a nonprofit organization-focused survey on the floor of a convention trade show where there is enough commotion that survey participants become distracted and do not complete the questionnaire.

Multiple Treatment Interference. Interference occurs when several of the same treatments are applied in succession to the same respondents. This is important because the

effects of the initial treatment cannot be erased or undone, and that may impact subsequent treatments (Campbell and Stanley, 1963).

Review of Various Types of Bias Associated with Survey Research – With an Emphasis on Bias in Sampling

There are, in general, two sources of research study error (or bias) that can threaten validity of a study: systematic bias (or error) and random bias (or error). Selection bias, measurement bias and intervention bias are the most common of many biases that can impact research studies.

In his seminal work titled *On Errors in Surveys*, Deming (1994) identified thirteen factors that can impact surveys. Those factors are worth mentioning here and include: variability in response, differences between kinds and degrees of canvass, interviewer and interviewer variation, auspices, questionnaire design and tabulation plan imperfections, changes in the universe prior to tabulating, nonresponse and omissions, late reports, unrepresentative selection of data or period covered, sampling errors, processing errors, and errors of interpretation.

Gerhard (2008) points out that the words ‘prejudice’ and ‘unfairness’ are synonyms for the word ‘bias’ which refers to a systematic error in research. There are many occasions when unintended bias occurs that has the potential to adversely impact the efforts of the researcher.

It is incumbent on the researcher to ensure that every effort has been addressed to eliminate opportunities for bias.

Bias in Sampling

As an overarching definition, sampling bias occurs when the information collected does not accurately depict the group from which the data was collected.

Sampling is a method researchers employ to efficiently and accurately obtain the opinions or demographic profiles of people in an identified group. That identified population could be reflective of age, sex, education, primary residence, religion, political persuasion, birth location, job title or any of dozens of combinations.

Sampling in survey research is a method of data collection that, when accurately applied, represents a total population and, more specifically, every point in the sample also needs to be representative of the population. If the survey sample is not accurate, selection bias occurs.

As Simundic (2013) further clarifies, bias happens when the population that is being studied is not reflective of its general population.

If sampling bias occurs, and if it is left unaddressed, or if the error goes unnoticed, the bias will prejudice findings. That might occur because the sample statistics will underestimate or overestimate a parameter (also known as a measurable characteristic) like a standard deviation or a mean.

Therefore, researchers using surveys to gather information are wise to be mindful of sampling bias which can occur when the sample does not accurately reflect the entire population. This can happen because of a flawed study design or inconsistent data collection procedures.

Some types of survey sampling bias are: selection error bias, measurement error bias, coverage error bias, implicit bias, sample mortality bias, sympathetic bias, interviewer bias, and questionnaire and tabulation design imperfection bias.

Selection Error Bias. For example, if we surveyed one group of people that includes all women with Bachelor's degrees, and we surveyed another group of people that includes all women with both Bachelor's degrees and Master's degrees, we would expect to see differences in responses.

Measurement Error Bias. As an example, let's assume a scientist has failed to calibrate an analytical balance scale and there is a five percent error in each measurement, so things seem to weigh five percent more than they actually weigh. Every weight measurement by the scientist will have a measurement error bias.

Coverage Error Bias. An example of coverage bias is a survey conducted by telephone. The bias is present because a telephone is needed in order for respondents to participate in the survey (Gordoni and Shavit, 2014). People who are not listed in the telephone book will be left out of the study, and this could constitute bias.

Cut-off Bias. Cut-off bias refers to closing the selection of participants prematurely (Smith and Noble, 2014). For example, assume that two hundred sixth graders are sent to the lunch room at 11:00 a.m., two hundred seventh graders are sent at 11:30 a.m., and two hundred eighth graders are sent at noon. If the surveyor randomly selects thirty students for a survey at 11:40 a.m., the eighth graders will be left out. Consequently, a representative sample will not have been identified at the middle school.

Implicit Bias. Bendick and Nunes (2012) caution researchers to be mindful of stereotypical behavioral traits that can have an influence on either the interviewee or the interviewer. An example is the interviewer who is physically intimidating enough that potential interviewees will decline the interview.

Sample Mortality Bias. As discussed previously, if the water aerobics class loses two participants between Monday and Wednesday, surveys completed on Monday and Wednesday are likely to have different results than surveys completed on Wednesday. Suchman (1962) advises that because sampling mortality bias is serious, the researcher must make every effort to secure the highest percentage of survey returns as possible.

Sympathetic Bias. Sympathetic bias occurs when the motivating force becomes a growing affection between the investigating statistician and the client (Levy and Peart, 2008).

Interviewer Bias. Interviewer bias can happen when the individual conducting the interview deviates from the protocol and begins asking questions which 'lead' the respondent to answer in a certain way. Ngongo, Frick, Hightower, Mathingau, Burke and Breiman (2015) observe that an interviewer can influence respondents to 'edit' their comments to how they think the interviewer wants them to respond.

Questionnaire and Tabulation Design Imperfections Bias. Words, phrases or questions that are unclear or ambiguous force the respondent to self-interpret. Emotionally toned words or suggestions are also types of bias.

Bias in Analysis

As with bias in the sampling process, bias can also occur during data analysis. Some opportunities for bias in data analysis that can be important include data fabrication, data abuse, data manipulation, and processing errors.

Data Fabrication. This occurs when bogus data is created and reported.

Data Abuse. This occurs when there is wholesale elimination of data when the data does not support the hypothesis.

Data Manipulation. This occurs when the researcher divides the population into subgroups that were not part of the original hypothesis, and, the researcher performs multiple tests until the data becomes statistically significant.

Processing errors. There are times when responses to surveys are coded by more than one person. Processing errors usually occur because of a difference in opinion as to how the

information should be coded. Another type of processing error occurs when the wrong keyboard letter or number is entered.

Bias in Data Interpretation. There are ample opportunities for bias to creep into data interpretation. Researchers must make every effort to present correct data, to develop correct and unbiased interpretations that accurately reflect results, and to provide accurate and truthful conclusions.

Bias in Definitions. For example, the term ‘Americans’ can have more than one definition. Some people would say that ‘Americans’ include everyone who is a resident in the continental United States, Alaska and Hawaii. Someone else might say that ‘Americans’ include all United States citizens who live in the continental United States, Alaska, Hawaii, other United States territories (such as Puerto Rico), and United States citizens living in other countries. If we wanted to take a survey of ‘Americans’ we would need to develop clear definitions of the group we want to survey, and we would need to apply that definition consistently during our research.

An Analysis of Methods of Controlling for Bias

Bias can occur in any research study. It is important that the researcher acknowledges the presence of bias and makes the effort to either eliminate or minimize (control) bias (Smith and Noble, 2014).

Several methods are available to researchers in their effort to control for bias in experimental survey designs.

Recognizing Bias Up Front. Outline the types of bias that might occur in advance of and during development of the study.

Well-Designed Protocol. A protocol of all the steps in the process is essential.

Sample Size. A sample size that is too small can result data that are inadequate to draw valid conclusions. In turn, if conclusions are drawn, they may be invalid.

Peer Review. Peer reviews by fellow researchers and by institutions and their review boards are helpful.

There are a number of options available for reducing bias. One is to include the recognition of the presence of bias. Another is the process of addressing the elimination or minimization (control) of bias.

An Analysis of Constraints and Limitations in Selecting a Research Tool for Nonprofit Management Inquiry

Some of the biggest challenges facing nonprofit organizations are trying to do too much, trying to respond to too many demands, and trying to touch too many audiences with too few resources. Nonprofit executives are expected to prudently manage the financial and human resources (both staff members and volunteers) of their organization in order to have as much of an impact as possible. It is with those challenges in mind that the following discussion occurs of constraints and limitations (money, time poverty, participant reluctance, incorrect information) in the selection of a research tool for nonprofit inquiry management.

Money. Nonprofit organizations, particularly those that have limited financial resources and an already over-burdened staff, may be inclined to search out an existing research survey questionnaire and attempt to adapt it to their own needs. Ritchie and Sherlock (2009) warn about the dangers of arbitrarily modifying existing language to fit circumstances other than those for which the original instrument was intended. Bielefeld (2006), too, cautions nonprofit organizations about what he terms the “uncritical use” (p. 397) of adapting existing research from other sources. The substitution of a single word or phrase – even if it is a minor change --

can easily alter the meaning of the question which, in turn, changes the response. The result can be incorrect information. Incorrect information that is applied to strategic (or business) plans or programs may take the nonprofit organization in a wrong direction.

Time Poverty. Lack of time coupled with the demand for people's attention are factors that cause people to prioritize, resulting in reluctance to participate in research. When the concept of surveying as a barometer or predictor of public opinion was introduced, people participated to a larger extent than they do now (Groves, 2006). Perhaps their participation was due to the novelty of public opinion and consumer surveys, or it could be that people had more discretionary time than they have today, or it could be that people today are tired of being surveyed.

Participant Reluctance. Today, some people are hesitant to contribute their thoughts, ideas and opinions. The reasons for that hesitancy include concerns about maintaining personal privacy or personal safety, or lack of interest. Response rates may increase if there are a variety of ways in which the target population can respond anonymously. Those options include mail surveys and computer-based surveys (Rudig, 2010).

Undue Influence. Undue influence can occur before or during the time a decision is made to conduct a study. Perhaps the staff is too over-burdened with administrative functions to create or implement a study, so the executive director attempts to subvert the activity with the intent of killing the proposal to do a survey. Another opportunity for undue influence is during the selection of a research firm. The membership director, or a member of the board of directors, or the president of a company in the top five percent of dues-paying members says "I have a friend who moonlights as a researcher" or "I know the perfect company."

Selection of a Research Tool for Nonprofit Management Inquiry

It is the goal of the researcher to provide accurate and timely results of research. The type of information the researcher seeks has an impact on the type of research tool that is selected. Contextual considerations that will impact the selection of a research tool include issues like population characteristics, cultural context, research goals and administrative issues (Switzer, Wisniewski, Belle, Dew and Schultz, 1999). Research tools having applicability for nonprofit management inquiry include interviews, forms of inquiry, and observations.

Interviews. These include one-on-one interviews and focus groups that can be a good way to acquire more in-depth information.

Forms of Inquiry. Forms of inquiry include questionnaires, score cards, rating scales, check lists, and internet key word searches.

Observation. Types of observation include covert, overt and researcher participation.

Choosing a research tool is an important responsibility. Decision makers should keep in mind the findings of Li and Van Ryzin (2012) who report that, of all the research options available to nonprofit organizations, survey forms of research were the most popular. The researchers also found that of the various forms of surveys, web surveys and mail surveys were the most popular.

Conclusion

This discussion addresses non-experimental research, where the researcher must be aware of issues of internal and external validity. Threats to internal validity can include: maturation, confounding, history, testing effects, instrument effects, and mortality. Threats to external validity can include small sample size, interaction effects, and multiple treatment interference.

There are several types of bias that can impact research. Bias can fall into two types: systematic bias and random bias, and researchers must take steps to eliminate it. Survey

sampling bias can include: selection error bias, measurement error bias, cut-off bias, implicit bias, sample mortality bias, sympathetic bias, interviewer bias, and questionnaire and tabulation bias. Bias can occur in analysis, including: data fabrication, data abuse, data manipulation, processing error bias, bias in data interpretation and bias in definitions

There can be constraints and limitations on selecting a research tool for nonprofit management inquiry, including: lack of money, time poverty, participant reluctance, and undue influence. Research tools that are applicable to nonprofit management include: interviews, forms of inquiry and observations.

References

- Bendick, M. and Nunes, A. P. (2012). Developing the research basis for controlling bias in hiring, *Journal of Social Issues*, 68(2), 238-262.
- Bielefeld, W. (2006). Quantitative research for nonprofit management. *Nonprofit Management & Leadership*, 16(4), 395-409. doi:10.1002/nml.117
- Campbell, D. T. and Stanley, J. C. (1963). *Handbook of Research and Teaching* (1st ed.). Boston, MA: Houghton Mifflin Company
- Damali, B. Z. A., and McGuire, F. A. (2013). The enabling potential of constraints. *Journal of Leisure Research*, 45(2), 136-149. Retrieved from <http://search.proquest.com.library.capella.edu/docview/1431922363?accountid=27965>
- Davern, M. (2013). Nonresponse rates are a problematic indicator of nonresponse bias in survey Research, *Health Services Research*, 48(3), 905-912. doi:10.1111/1475-6773.12070
- Deming, W. E. (1944). On errors in surveys, *American Sociological Review*, 9(4), 359-369.
- Gerhard, T. (2008). Bias: Considerations for research practice, *American Journal of Health-System Pharmacy*, 65, 2159-2168. DOI 10.2146/ajhp070369
- Godoni, G., Oren, A. and Shavit, Y. (2011). Coverage bias in telephone surveys in Israel, *Field Methods*, 23(2), 188-203. doi:10.1177/1525822X10387573
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys, *Public Opinion Quarterly*, 70(5), 646-675. doi:10.1093/poq/nna33
- Hodgkinson, G. P. (1992). Development and validation of the strategic locus of control. *Strategic Management Journal*, 13(4), 311-317. Retrieved from <http://search.proquest.com.library.capella.edu/docview/224964542?accountid=27965> .

- Hodgkinson, G. P. (1993). Doubts about the conceptual and empirical status of context-free and firm-specific control expectancies: A reply to Boone and de Brabander. *Strategic Management Journal*, 14(8), 627. Retrieved from <http://search.proquest.com.library.capella.edu/docview/224669834?accountid=27965> .
- IRS (2015). Exempt organizations business master file extract (EO BMF). Retrieved from: <http://www.irs.gov/Charities-&-Non-Profits/Exempt-Organizations-Business-Master-File-Extract-EO-BMF>
- Koryak, O., Mole, K.F., Lockett, A., Hayton, J.C., Ucbasaran, D. and Hodgkinson, G. P. (2015). Entrepreneurial leadership, capabilities and firm growth, *International Small Business Journal*, 33(1), 89-105. doi: 10.1177/0266242614558315
- Levy, D. M., & Peart, S. J. (2008). Sympathetic bias. *Statistical Methods in Medical Research*, 17(3), 265-77. doi: <http://dx.doi.org/10.1177/0962280207080642>
- Ngongo, C. J., Frick, K. D., Hightower, A. W., Mathingau, F. A., Burke, H. and Breiman, R. F. (2015). The perils of straying from protocol: sampling bias and interviewer effects, *Public Library of Science, ONE 10(2)*. doi:10.1371/journal.pone.0118025
- Ogden, H. N. (1918). The Purpose of Research, *Science - New Series*, 48(1248), 525-532. Retrieved from: <http://www.jstor.org/stable/1642926>
- Research (2015). Definition. Merriam-Webster Dictionary. Retrieved from: <http://www.merriam-webster.com/dictionary/research>
- Ritchie, W. J. and Sherlock, J. J. (2009). Adapting surveys for nonprofit research, *Nonprofit Management & Leadership*, 19(3), 387-403. DOI: 10.1002/nml.225
- Rudig, W. (2010). Assessing nonresponse bias in activist surveys, *Quality and Quantity*, 44(1), 173–180. DOI 10.1007/s11135-008-9184-9

- Shier, M L. and Handy, F. (2014). Research trends in nonprofit graduate studies: A growing interdisciplinary field, *Nonprofit and Voluntary Sector Quarterly*, 43(5), 812–831. doi: 10.1177/0899764014548279
- Šimundić, A.M. (2013). Bias in research, *Biochemia Medica*, 23(1), 12-17.
<http://dx.doi.org/10.11613/BM.2013.003>
- Smith, J. and Noble, H. (2014). Bias in research, *Evidence Based Nursing*, 17(4), 100-101. doi: 10.1136/eb-2014-101946
- SOPSL (2015). Acceptable SOPSL research methods and research design, Capella University.
Retrieved from:
<http://www.capella.edu/iGuidePA/PDF/academics/AcceptableMethodsPSL.pdf>
- Speer, P. W., & Christens, B. D. (2013). An approach to scholarly impact through strategic engagement in community-based research. *Journal of Social Issues*, 69(4), 734-753.
- Suchman, E. A. (1962). An analysis of bias in survey research, *The Public Opinion Quarterly*, 26(1), 102-111.
- Switzer, G. E., Wisniewski, S. R., Belle, S. H., Dew, M. A., & Schultz, R. (1999). Selecting, developing, and evaluating research instruments. *Social Psychiatry and Psychiatric Epidemiology*, 34(8), 399-409. Retrieved from
<http://search.proquest.com.library.capella.edu/docview/236072541?accountid=27965>
- Woodman, R. W. (2014). The role of internal validity in evaluation research on organizational change interventions, *Journal of Applied Behavioral Science*, 50(1), 40-49.
doi: 10.1177/0021886313515613J