SAMPLING BIAS AND ITS IMPLICATIONS FOR RESEARCH VALIDITY

Enoch Daniel[1], Abubakar Musa Tafida[2], Suleiman Aliyu Abubakar[3]*, Ahmed Osu Abdulrahman[4] & Salisu Oloko Abubakar[5]

[1]Nigeria Security and Civil Defense Corps, Keffi Area Command, Nasarawa State, Nigeria

[2]Department of Psychology, Nasarawa State University Keffi, Nasarawa State, Nigeria

[3]Department of Psychology, Federal University of Lafia, Nasarawa State, Nigeria

[4]Department of Crime Management, Federal Polytechnic Nasarawa, Nigeria

[5]Department of Crime Management, Federal Polytechnic Nasarawa, Nigeria

*suleiman220@gmail.com

ABSTRACT: This paper explores sampling bias and its implications for research validity. It starts by explaining the concept of sampling bias. The paper further defines sampling bias in research as the collection of samples that do not accurately represent the entire group. The paper further explains the types of sampling bias and the causes of sampling bias. It also discusses the sampling bias and implications for research validity with some examples. Furthermore, it looks at the approaches to mitigate sampling bias, which includes Probability sampling, Weighting and adjustments, response rate optimization, and pilot testing. Lastly, the paper gave some recommendations for the best practices in sampling from a population in order to avoid sampling bias.

Keywords: Sampling Bias, Research Validity, Probability Sampling, Response Rate Optimization, Mitigation Strategies

INTRODUCTION

Sampling bias, a pervasive issue in research, occurs when a sample is selected in a way that systematically favours certain characteristics, leading to an unrepresentative sample (Cochran, 1977). This bias can result in inaccurate or misleading conclusions, undermining the validity and reliability of research findings (Groves, 2006). Sampling bias in research is the collection of samples that do not accurately represent the entire group. A biased sample is the result of collecting a sample from a population that is not random and tends to produce a particular outcome (Tourangeau, 2014).

According to Kish (1965), sampling bias in research is the collection of samples that do not accurately represent the entire group. Sampling bias occurs during data collection. The reason the sample is biased is that the data collected has a higher chance of occurring than other possible data.

Sampling bias is a statistic computed of the sample may be systematically erroneous. Sampling bias can lead to a systematic over- or underestimation of the corresponding parameter in the population. Sampling bias occurs in practice as it is practically impossible to ensure perfect randomness in sampling. If the degree of misrepresentation is small, then the sample can be treated as a reasonable approximation to a random sample. Also, if the sample does not differ markedly in the quantity being measured, then a biased sample can still be a reasonable estimate.

The word bias has a strong negative connotation. Indeed, biases sometimes come from deliberate intent to mislead or other scientific fraud. In statistical usage, bias merely represents a mathematical property, whether it is deliberate or unconscious or due to imperfections in the instruments used for observation (Tourangeau, 2014). While some individuals might deliberately use a biased sample to produce misleading results, more often, a biased sample is just a reflection of the difficulty in obtaining a truly representative sample, or ignorance of the bias in their process of measurement or analysis.

This paper aims to systematically explore the types, causes, and implications of sampling bias on research validity and provide practical strategies to mitigate these biases**.**

Types of Sampling Bias

1. **Selection bias:** Selection bias refers to situations where research bias is introduced due to factors related to the study's participants. Selection bias can be introduced via the methods used to select the population of interest, the sampling methods, or the recruitment of participants. It is also known as the selection effect (Berk, 1983).

Examples of studies that encountered selection bias:

COVID-19 Testing and Health Awareness in the U.S.

Liamputtong (2020) examined selection bias in COVID-19 testing behaviours, focusing on how health-conscious individuals are more likely to seek testing and participate in health-related studies. The study, a cross-sectional survey of 1,200 participants, showed that individuals already practising preventive health measures were overrepresented. This skewed the data, suggesting that higher health awareness correlated with better health outcomes, although this may not reflect the general population. Such selection bias affected both the internal validity and generalizability of the findings (Liamputtong, 2020).

Cancer Screening Programs and Referral Bias

A study on cancer screening programs in high-risk populations noted a referral bias due to healthcare professionals preferentially recommending screening for individuals with family histories of cancer or symptoms. Conducted across multiple hospitals with 800 patients, the study found that participants with referrals had a higher incidence of cancer. This selection bias potentially inflated cancer detection rates compared to an unbiased sample of the population (Bryman, 2020).

Fitness Level Study with Location Bias

Atkinson et al. (2021) analysed fitness levels using participants from a gym setting, which introduced location bias. The study, with 300 participants, measured physical fitness indicators. Because gym-goers typically have higher fitness levels, the results misrepresented average fitness, as data collection in gyms naturally excluded fewer active individuals. This setting-specific bias limits the findings' applicability to broader populations (Atkinson et al., 2021).

These examples underscore the importance of considering selection methods to enhance the reliability and applicability of study results by minimizing bias in participant selection.

Examples of studies that encountered selection bias in Nigeria:

Gender Inequality in Medical and Dental Institutions

A 2022 study examined the effects of gender inequality within Nigerian medical and dental research institutions using a qualitative approach. Researchers conducted interviews with faculty members from 17 universities, targeting both male and female professionals to understand the gendered dynamics in academic health fields. Participants were selected through a mix of purposive and convenience sampling, with snowball sampling also employed to broaden the pool. The findings showed that selection bias was evident in recruitment and promotions, with limited opportunities for female faculty members to advance to higher academic positions. Institutional culture and traditional expectations about gender roles were noted as significant barriers (PLOS ONE, 2022).

Tenant Selection Bias in Enugu Metropolis

A study focused on tenant selection criteria among estate surveyors in Enugu highlighted how biases affect property allocation. The survey included estate managers who rated potential tenants based on factors such as ethnicity, occupation, and family size. The study revealed that ethnicity was a significant factor influencing tenant selection, with participants often favouring tenants from certain ethnic backgrounds over others. This selection bias raised concerns about equitable access to housing, demonstrating how socio-cultural biases impact rental markets in Nigeria (Oyedeji, 2022).

COVID-19 Self-Testing Uptake

In a cross-sectional study, researchers explored willingness to use COVID-19 self-tests among the Nigerian population. Using online and physical outreach, researchers surveyed individuals in urban and semi-urban areas, though the study acknowledged a limitation due to the underrepresentation of rural participants. This self-selection led to a form of selection bias, as the sample may not fully represent the views of rural communities where healthcare access is more limited. The study's results indicated high interest in self-testing but emphasized the need for more inclusive sampling strategies to guide future public health policies (BMJ Open, 2022).

These studies illustrate the presence of selection biases arising from institutional preferences, socio-cultural biases, and sampling limitations, underscoring the importance of representative samples.

Selection bias may threaten the validity of your research, as the study population is not representative of the target population. Selection bias occurs when the selection of subjects into a study (or their likelihood of remaining in the study) leads to a result that is systematically different to the target population. Selection bias often occurs in observational studies where the selection of participants is not random, such as cohort studies, case-control studies, and cross-sectional studies. It also occurs in interventional studies or clinical trials due to poor randomisation.

Selection bias is a form of systematic error. Systematic differences between participants and non-participants or between treatment and control groups can limit your ability to compare the groups and arrive at unbiased conclusions. This paper aims to systematically explore the types, causes, and implications of sampling bias on research validity and provide practical strategies to mitigate these biases. Several potential sources of selection bias can affect the study, either during the recruitment of participants or during the process of ensuring their retention.

Types of selection bias

Selection bias is a general term describing errors arising from factors related to the population being studied, but there are several types of selection bias:

a. **Ascertainment bias** occurs when some members of the intended population are less likely to be included than others. As a result, your sample is not representative of your population.
b. **Attrition bias** occurs when participants who drop out of a study are systematically different from those who remain.
c. **Self-selection bias (or volunteer bias)** arises when individuals decide entirely for themselves whether or not they want to participate in the study. Due to this, participants may differ from those who don't—for example, in terms of motivation.

d. Survivorship bias is a form of logical error that leads researchers who study a group to draw conclusions by only focusing on examples of successful individuals (the "survivors") rather than the group as a whole.

2. Non-response bias: Nonresponse bias is observed when people who don't respond to a survey are different in significant ways from those who do. Non-respondents may be unwilling or unable to participate, leading to their under-representation in the study. Nonresponse bias occurs when survey participants are unwilling or unable to respond to a survey question or an entire survey. Reasons for nonresponse vary from person to person.

To be considered a form of bias a source of error must be systematic in nature. Nonresponse bias is not an exception to this rule. If a survey method or design is created in a way that makes it more likely for certain groups of potential respondents to refuse to participate or be absent during a surveying period, it has created a systematic bias (Lindsay & Ehrenberg, 1993)

Take these two examples, for instance:

i. Asking for sensitive information: Consider a survey measuring tax payment compliance. Citizens who do not properly follow tax laws will be the most uncomfortable filling out this survey and be more likely to refuse. This will obviously bias the data towards a more law-abiding net sample than the original sample. Nonresponse bias in surveys asking for legally sensitive information has been proven to be even more profound if the survey explicitly states that the government or another organization of authority is collecting the data.
ii. Invitation issues: Many researchers create nonresponse bias because they do not pretest their invites properly. For example, a large portion of young adults and business sector

workers answer the majority of their emails through their smartphones. If the survey invite is provided through an email that doesn't render well on mobile devices, smartphone users' response rates will drop dramatically. This will create a net sample that underrepresents the opinions of the smartphone user demographic.

3. Measurement bias: Measurement bias occurs when information collected for use as a study variable is inaccurate. The incorrectly measured variable can be either a disease outcome or an exposure. Measurement bias can be further divided into random or non-random misclassification. We are more concerned with non-random misclassification, as this can spuriously inflate or reduce estimates of effect. Non-random misclassification can itself be divided into subtypes, including observer bias and recall bias (Hox, 1997).

Examples of studies that encountered measurement bias:

Popogbe and Adeosun (2022) researched brain drain among university lecturers in Nigeria, which continued to be a critical issue in 2023. This study analysed factors influencing academic migration, like poor work environments and compensation. The researchers noted potential measurement bias due to self-reporting methods, as lecturers may have exaggerated challenges to justify migration intentions. The study highlighted infrastructure and economic challenges as major motivations for brain drain but faced bias issues due to reliance on subjective reporting and limited representativeness of respondents across diverse Nigerian regions (Popogbe & Adeosun, 2022).

Anunike et al. (2023) investigated the impact of Nigeria's naira redesign on vote-buying in Anambra State during the 2023 presidential election. Using a questionnaire to gauge voter perceptions, this study encountered measurement bias due to respondents' social desirability bias; participants may have altered their answers to align with socially acceptable views against vote-buying. This bias posed challenges in accurately assessing attitudes toward the naira redesign's influence on electoral behaviour

Systematic Review of Food Insecurity Studies in Nigeria (2023) aimed to consolidate research on food insecurity in Nigeria but faced measurement bias in interpreting data from varied metrics and scales. Many studies in this review used inconsistent food security measures, resulting in difficulties comparing data across regions. This inconsistency introduced bias in assessing food insecurity prevalence, as findings varied significantly based on the measurement tools employed in different local studies (MDPI, 2023).

4. Under coverage bias: Under coverage bias occurs when some members of your population are not represented in the sample. It is common in convenience sampling, where you recruit a sample that's easy to obtain. This occurs when a part of the population is excluded from your sample. As a result, the sample is no longer representative of the target population. Non-probability sampling designs are susceptible to this type of research bias.

Examples: You are conducting research by randomly calling landline numbers. Because of your sampling method, individuals who only have mobile phones are not sampled. In this case, they are not covered at all. Online surveys also exclude people who don't have internet access. Previous research shows that internet access also relates to demographics like socioeconomic status and age.

Online surveys exclude people who don't have internet access. Previous research shows that internet access also relates to demographics like socioeconomic status and age. Ideally, researchers should draw a sample that, like a snapshot, adequately captures characteristics present in the target population and relevant to the research. In other words, researchers aim to collect a representative sample. In some cases, researchers may sample too few units from a specific segment of the population. If the segment is small in comparison to others in the population, this may not impact the research findings much. However, if the segment is larger, it can lead to a sample that does not accurately capture the characteristics of the population. In more extreme cases, researchers may completely fail to include a part of the population, which can distort the findings completely (Lessler & Kalsbeek, 1992).

Causes of Sampling Bias:

Poor sampling design Poor sampling size refers to selecting a sample that is too small or too large, which can lead to inaccurate or unreliable conclusions (Cochran, 1977). Example of poor sampling size:

1. Surveying College Students:

- Population: 10,000 college students

- Sample size: 20 students

-Method: Convenience sampling (surveying friends and classmates)

Problems:

- Small sample size (0.2% of population)

- Biased sampling method (convenience sampling)

- Results may not generalize to entire population

2. Inadequate sample size: Inadequate sample size refers to selecting a sample that is too small to accurately represent the population (Krejcie & Morgan, 1970). Example: Surveying 50 students to understand the attitudes of an entire university population (10,000+ students).

Another example is conducting a study on job satisfaction with a sample of 20 employees from a single company.

3. Biased sampling frame: A biased sampling frame does not accurately represent the population, leading to biased or unrepresentative samples (Kish, 1965).

Bias sampling frame examples:

- Non-representative population:

Surveying only urban residents about transportation preferences, ignoring rural residents.

- Incomplete or outdated data:

Using voter registration records from 2010 to sample eligible voters in 2022.

- Selection bias:

Recruiting participants through social media ads, targeting only specific demographics.

- Exclusion criteria:

Conducting a study on "average" consumers, excluding low-income or minority groups.

- Geographic bias:

Sampling only participants from coastal regions, ignoring inland areas.

4. Non-response rates: Non-response rates refer to the percentage of individuals or organisations that do not respond to a survey, questionnaire, or other data collection method (Lindsay & Ehrenberg, 1993).

Examples of non-response rates:

i. Customer Satisfaction Survey:

- Total sample size: 1,000

- Non-response rate: 40% (400 did not respond)

- Response rate: 60% (600 responded)

ii. Employee Engagement Survey:

- Total sample size: 500

- Non-response rate: 25% (125 did not respond)

- Response rate: 75% (375 responded)

5. Social desirability bias: Social desirability bias (SDB) refers to the tendency of individuals to provide responses that are deemed socially acceptable or desirable rather than their true beliefs or behaviours (Nederhof, 1985).

Survey Examples:

i. A survey on exercise habits:

- Question: "How often do you exercise per week?"

- Biased response: 75% of respondents claim to exercise 3-4 times/week (actual: 40%)

ii. A survey on dietary habits:

- Question: "Do you follow a healthy diet?"

- Biased response: 90% of respondents claim to follow a healthy diet (actual: 60%)

Implications for Research Validity:

Sampling bias has many implications for research validity

1. Internal validity threats: Internal validity refers specifically to whether an experimental treatment/condition makes a difference or not, and whether there is sufficient evidence to support the claim (Shadish et al., 2002).

Examples of Internal validity threats

i. History: These are events which occur between the first and second measurement. Example: Studying the effect of a new educational program:

- Pre-test: Measure students' knowledge before program implementation

- Event: A major educational reform is announced mid-study

- Post-test: Measure students' knowledge after program implementation.

History threat: The reform may influence students' knowledge, confounding program effects.

ii. Maturation- the processes within subjects which act as a function of the passage of time. i.e. if the project lasts a few years, most participants may improve their performance regardless of treatment. Physical growth: Studying cognitive development in children.

- Pre-test: Measure cognitive abilities at age 8

- Treatment: Implement educational program

- Post-test: Measure cognitive abilities at age 10

Maturation threat: Natural cognitive development may occur regardless of program.

iii. Instrumentation- the changes in the instrument, observers, or scorers which may produce changes in outcomes.

Example:

Observer bias

- Pre-test: Measure behaviour with Observer A

- Treatment: Implement intervention

- Post-test: Measure behaviour with Observer B

Instrumentation threat: Different observers may rate behaviour differently.

iv. Statistical regression- it is also known as regression to the mean. This threat is caused by the selection of subjects on the basis of extreme scores or characteristics. Give me forty worst students and I guarantee that they will show immediate improvement right after my treatment.

v. **Selection of subjects**- the biases which may result in selection of comparison groups. Randomization (Random assignment) of group membership is a counter-attack against this threat. However, when the sample size is small, randomization may lead to Simpson Paradox, which has been discussed in an earlier lesson.

vi. **Experimental mortality-** the loss of subjects. For example, when a participant dies while participating in an experiment.

vii. Selection-maturation interaction- the selection of comparison groups and maturation interacting which may lead to confounding outcomes, and erroneous interpretation that the treatment caused the effect.

2. Biased effect size estimates: This means that in addition to the bias that is due to the fact that the sampling variance is not known but is estimated using the observed effect size, the estimate is biased because of the true between-studies variance (Hedges & Vevea, 1998)

3. Incorrect conclusions: An incorrect conclusion is where all given reasons and evidence point to a given conclusion, but due to the omission, incorrect assumption, lie or missing piece of information required, the individual arrives at a false conclusion. There are two types of false conclusion: Valid false conclusion (Berk, 1983)

4. Limited generalizability: If the results of a study are broadly applicable to many different types of people or situations, the study is said to have good generalizability. If the results can only be applied to a very narrow population or in a very specific situation, the results have poor generalizability (Cohen et al., 2013)

Approaches to Mitigate Sampling bias

1. Probability sampling: Probability sampling is a technique in which the researcher chooses samples from a larger population using a method based on probability theory. For a participant to be considered as a probability sample, he/she must be selected using a random selection.

This statistical method used to select a sample from a population in such a way that each member of the population has a known, non-zero chance of being selected. The most critical requirement of probability sampling is that everyone in your population has a known and equal chance of getting selected.

Probability sampling uses statistical theory to randomly select a small group of people (sample) from an existing large population and then predict that all their responses will match the overall population (Cochran, 1977).

Types of probability sampling:

i. Simple Random Sampling: This method involves randomly selecting a sample from the population without any bias. It's the most basic and straightforward form of probability sampling.

ii. Stratified random Sampling: This method involves dividing the population into subgroups or strata and selecting a random sample from each stratum. This technique is useful when the population is heterogeneous and you want to ensure that the sample is representative of different subgroups.

iii. Cluster Sampling: This method involves dividing the population into groups or clusters and then randomly selecting some of those clusters. This technique is useful when the population is spread out over a large geographical area. But It is not possible or practical to survey everyone.

iv. Systematic Sampling: This method involves selecting every nth member of the population after a random starting point is chosen.

2. Weighting and adjustments: Weighting is a statistical technique in which datasets are manipulated through calculations in order to bring them more in line with the population being studied. The key difference between the initial sample composition and weighting is that weights are applied after data is collected, and allow researchers to correct for issues that occurred during data collection. For this reason, weighting is also known as post-stratification, as it takes place after the sample has been selected, as opposed to pre-stratification, which is used to balance a sample before data has been collected.

Researchers applying weights most often weight demographic characteristics such as age, gender, location, and education, but weighting can also account for the differences between those who participate or do not participate in research studies (known as self-selection bias). Weights can also minimize any effects the survey design or data collection mode may have on the sample makeup and resulting data.

In addition to weighting on common demographic variables, studies have found that weighting based on other variables such as internet usage and political affiliation can further reduce bias in some cases. If conducting a phone survey, for example, weights can be applied based on mobile versus landline phone users (Hox, 1997).

3. Response rate optimization: Response Rate Optimization (RRO) refers to strategies and techniques aimed at maximising the number of participants who respond to a survey, questionnaire, or other research instrument (Lindsay & Ehrenberg, 1993).

4. Pilot testing: Pilot testing is a preliminary test or study conducted before a larger-scale study. A pilot study can present key information that can help guide the direction of the larger study or research project, including providing insights into the ultimate cost of the study, its overall feasibility, and any challenges that the actual study may face once it gets off the ground (Krejcie & Morgan, 1970).

Pilot testing is important because it goes a long way toward providing more information and deeper insights into your future study.

Recommendations for best practices in sampling from a population in order to avoid sampling bias:

To ensure adequate sampling bias of future research, the following recommendations and best practices can be implemented:

a. Clear sampling objectives: These refer to the specific goals and purposes of selecting a sample from a population and guide the sampling design and methodology.

Types of Sampling Objectives:

1. Descriptive: To describe the characteristics of the population.

2. Inferential: To make inferences about the population based on sample data.

3. Exploratory: To explore relationships or phenomena.

4. Predictive: To predict outcomes or behaviours.

b. Documented sampling procedures: Documented sampling procedures refer to the systematic recording and documentation of every step involved in selecting and collecting data from a sample, ensuring transparency, reproducibility, and accuracy.

Importance of Documented Sampling:

1. Ensures research transparency and accountability

2. Facilitates reproducibility and verification

3. Enhances data quality and validity

4. Supports research reliability and generalizability

5. Helps in addressing sampling bias and errors

c. Representative sampling frames: A representative sampling frame is a list or database that accurately reflects the characteristics of the target population, ensuring that every individual or unit has an equal chance of being selected. It is important because it ensures sample representativeness, reduces sampling bias, increases data quality and validity, supports reliable conclusions, and enhances generalizability.

d. Adequate Sample Size: An adequate sample size is the minimum number of participants or observations required to ensure that the research findings are reliable, precise, and generalizable to the population. It is also important because it ensures statistical power and precision, reduces sampling error and bias increases confidence in research findings, supports reliable conclusions and decisions and also enhances research validity and credibility

e. Response rate monitoring: Response rate monitoring refers to the systematic tracking and analysis of the percentage of participants who respond to a survey, questionnaire, or other data collection instrument.

Conclusion

Sampling bias poses a significant threat to research validity, compromising the accuracy and generalizability of findings. It occurs when the sample selected for a study is not representative of the population, leading to distorted or incomplete data. It significantly impacts research validity, affecting the accuracy, reliability, and generalizability of findings. It can lead to distorted estimates, limited generalizability, invalid statistical inferences and misinformed decision-making.

REFERENCES

Atkinson, R., et al. (2021). Location Bias in Fitness Research.

Berk, R. A. (1983). An introduction to structural equation modeling with an application to the study of dropout rates. *Journal of Educational Statistics*, 8(2), 137-161.

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. American Sociological Review, 48(3), 386-398.

BMJ Open. (2022). "COVID-19 self-testing acceptance in Nigeria."

Bryman, A. (2020). Referral Bias in Cancer Screening Studies.

Cochran, W. G. (1977). Sampling techniques (3rd ed.). John Wiley & Sons.

Groves, R. M. (2006). Undermining the validity and reliability of surveys: Nonresponse. In Survey nonresponse (pp. 1-17). Wiley.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in the measurement of weighted mean effect size. Psychological Methods, 3(4), 393-406.

Hox, J. J. (1997). From theoretical concept to survey question: Issues in the study of observer and recall biases. Bulletin de Méthodologie Sociologique, 55(1), 68-93.

Kish, L. (1965). Survey sampling. John Wiley & Sons.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30(3), 607-610.

Lessler, J. T., & Kalsbeek, W. D. (1992). Nonsampling error in surveys. Wiley.

Liamputtong, P. (2020). Selection Bias in Health Awareness and COVID-19 Testing.

Lindsay, R. M., & Ehrenberg, A. S. C. (1993). How businesses can guard against nonresponse bias. *Journal of Business Research, 26*(2), 161-174.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. European *Journal of Social Psychology,* 15(3), 263-280.

Oyedeji, R. (2022). "Selection biases in tenant selection in Enugu."

PLOS ONE. (2022). "Gender inequality in medical and dental institutions in Nigeria."

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002) identified several threats to internal validity in experimental and quasi-experimental designs.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Tourangeau, R. (2014). Defining hard-to-survey populations. In R. Tourangeau, B. Edwards, & T. P. Johnson (Eds.), Hard-to-survey populations (pp. 3-22). Cambridge University Press.

Tourangeau, R. (2014). Defining hard-to-survey populations. In R. Tourangeau, B. Edwards, T. P. Johnson, K. M. Wolter, & A. M. Bates (Eds.), Hard-to-survey populations (pp. 3-24). Cambridge University Press.