

**CHALLENGES AND STRATEGIES FOR RECRUITING  
REPRESENTATIVE SAMPLES IN RESEARCH AND  
IMPLICATIONS FOR GENERALIZABILITY**

**Suleiman Aliyu Abubakar<sup>1\*</sup>, Abubakar Musa Tafida<sup>2</sup>, Shitu Shuaibu Umar<sup>3</sup>,  
Ishaq Tijjani Usman<sup>4</sup>, Orume Ibrahim Joseph<sup>5</sup> & Kemi Yusuf Agana<sup>6</sup>**

<sup>1</sup>Department of Psychology, Federal University of Lafia, Nigeria

<sup>2,3</sup>Department of Psychology, Nasarawa State University Keffi, Nigeria

<sup>4</sup>Department of Crime Management, Federal Polytechnic Nasarawa, Nigeria

<sup>5</sup>Nigerian Air Force Reference Hospital Port Harcourt. Rivers State, Nigeria

<sup>6</sup>Area Council Services Secretariat, Federal Capital Territory Administration, Abuja, Nigeria

\*suleiman220@gmail.com

**ABSTRACT:** This paper examines the challenges in recruiting representative samples, the impact on generalizability, and practical strategies for improving recruitment in diverse populations. It starts by explaining the concept of representative samples and then explains the generalization of findings. The paper further defines recruiting of representative samples as a process whereby research employs a subset of the population in order to collect data so as to generalize the information gotten from the sample to the general population. The paper further explains the consequences of recruiting the samples. Also, it discusses the factors that influence generalization, importance of representative samples, it also explains the challenges in recruiting representative sample in research. Furthermore, the seminar looks at the approaches to mitigate challenges in recruiting representative samples which include; designed-based approaches, recruitment-based approaches, data collection approaches, analytical approaches and partnership-based approaches. Finally, the paper gave some recommendations for the best practices in recruiting representative samples and generalizability of findings.

**Keywords:** Representative Samples, Generalizability, Recruitment Challenges, Data Collection Approaches, Mitigation Strategies

## **INTRODUCTION**

Recruiting representative samples and generalizability of findings are perennial challenges in psychological research, threatening the validity and applicability of study results (Henrich et al., 2010; Sears, 1986). The reliance on convenience samples, such as college students and online participants, can lead to a lack of diversity and representation, resulting in findings that may not accurately reflect the broader population (Peterson, 2001). Moreover, the increasing use of online data collection methods has introduced new biases and limitations, such as self-selection bias and technical issues (Gosling & Mason, 2015). Recruitment is the dialogue which takes place between an investigator and a potential participant prior to the initiation of the consent process. It begins with the identification, targeting and enlistment of participants (volunteer patients or controls) for a research study. It involves providing information to the

potential participants and generating their interest in the proposed study. There are two main goals of recruitment:

- i. to recruit a sample that adequately represents the target population;
- ii. to recruit sufficient participants to meet the sample size and power requirements of the study (Hulley et al., 2001).

A representative sample is a subset of a population that seeks to accurately reflect the characteristics of the larger group. For example, a classroom of 30 students with 15 males and 15 females could generate a representative sample that might include six students: three males and three females. Samples are useful in statistical analysis when population sizes are large because they contain smaller, manageable versions of the larger group. It is one technique that can be used for obtaining insights and observations about a targeted population group.

Representative sample is a small subset group that seeks to reflect specified characteristics exemplified in a target population proportionally. For extensive surveys, researchers may divide the population into strata according to ethnic markers, gender characteristics, age groups, income, or geographical locations to ensure that they are using a representative sample (Ashery et al., 1992).

Recruiting representative samples challenges refer to the difficulties in selecting and recruiting participants who accurately reflect the characteristics of the larger population, ensuring the generalizability and validity of research findings (Kennis et al., 2020). According to (Lilienfeld, 2017), recruiting representative samples is a cornerstone of rigorous research, yet it poses significant challenges. Ensuring that study participants accurately reflect the larger population is crucial for generalizability, validity, and reliability. However, various obstacles hinder the recruitment process, threatening the accuracy and applicability of research findings.

During recruitment, the sampling process can suffer from associated problems of non-response and the resultant selection bias. The proportion of eligible participants who agree to enter the study (the response rate) influences the validity of the inference that the sample represents the population of interest (Woods, et al., 2001). People who are difficult to reach and those who refuse to participate once they have been contacted tend to be different from people who do not enrol (Armstrong et al., 1992).

The retention of participants is an important aspect of recruitment. Poor retention is costly both financially and in terms of time. Early retention techniques should be incorporated into recruitment strategies during the planning phase of the study (Nishimoto, 1998). Retention also involves building relationships with participants to encourage their continuing participation.

The consequences of these challenges are far-reaching, as they can lead to:

1. Overestimation or underestimation of effects (Kennis et al., 2020).
2. Inaccurate predictions and recommendations (Yarkoni, 2010).
3. Limited applicability to real-world settings (Lilienfeld, 2017).

Generalization of findings refers to the extent to which research results can be applied to broader populations, contexts and settings beyond the specific sample or study conditions

(Simons et al., 2017). It involves inferring that the observed effects or relationships will hold true for other individuals, groups, or situations (Kennis et al., 2020).

**Factors influencing generalisation of findings:**

**a. Sample representativeness:** According to (Henrich et al., 2010) a representative sample is a sample from a larger group that accurately represents the characteristics of a larger population. It's known as a representative sample because the answers obtained from it accurately reflect the results you would achieve by interviewing the entire population. For example, in a warehouse with a sample of 1,000 people split equally into 500 males and 500 females, a smaller group of 100 males and 100 females could generate a representative sample of the larger group.

**b. Study design and methods:** Study design refers to the methods and methodologies used in research to gather the data needed to explore a specific question. Some research questions are best approached by statistical analysis of data. This is quantitative research. Others are better answered by looking for patterns, features, or themes in the data, which is qualitative research (Simons et al., 2017).

**c. Contextual factors:** In the social sciences, a contextual factor refers to any external variable or condition that is considered in research to understand its impact on the outcome being studied. These factors can include demographic characteristics such as age, gender, race/ethnicity, education, and employment-related factors, among others. They are often adjusted for or analysed through interaction terms in regression models to better understand their influence on the research findings (Gosling & Mason, 2015).

**Importance of Representative Samples:**

**1. Enhance generalizability:** Enhancing generalizability refers to increasing the extent to which research findings can be applied to broader populations, settings, and contexts. For example: a study on voter behaviour selects participants from 10 states, ensuring representation from urban, rural, and suburban areas to enhance generalization (Simons et al., 2017).

**2. Increase validity:** Increasing validity refers to enhancing the accuracy, reliability, and generalizability of research findings, ensuring they accurately reflect the phenomenon being studied. A study on the effectiveness of a new antidepressant medication randomly assigns 100 participants to treatment or placebo groups (Gosling & Mason, 2015).

**3. Improve reliability:** This means enhancing the consistency, dependability, and accuracy of research measurements, methods, and findings. To achieve this, the researcher will conduct a pilot study to test the instrument to see if it can measure what it intends to measure. (Kennis et al., 2020).

**4. Support external validity:** According to Lilienfeld (2017), Supporting external validity in research refers to the extent to which findings can be generalized to populations, settings, and contexts beyond the study sample. For example, a study to evaluate the effectiveness of a new exercise program for adults randomly selected 1,000 adults from a national database. External validity will be high as the sample represents the broader adult population.

**5. Facilitate meta-analyses:** This refers to the process of making research studies more suitable for inclusion in a meta-analysis, which is a statistical method for combining the results of multiple studies to draw more general conclusions. Example: Evaluating the effectiveness of cognitive-behavioural therapy (CBT) for depression. Means, standard deviations, and sample sizes for treatment and control groups will be assessed and analysed (Peterson, 2001).

**6. Promote replicability** (Simons et al., 2017) explained that promoting replicability is the practice and procedure that ensures that research findings can be duplicated or verified by others, increasing confidence in the results and advancing scientific knowledge.

**7. Inform evidence-based practices:** This involves the process of using research findings, data, and scientific evidence to guide decision-making and improve outcomes in various fields, such as: Healthcare, Education, Social work, Business, Policy-making and public health (Lilienfeld, 2017).

**8. Advance psychological theory:** Advancing psychological theory refers to the process of developing, refining, and expanding existing theories in psychology to better understand human behaviour, cognition, and emotion (Henrich et al., 2010).

By using representative samples, researchers can increase the accuracy, reliability, and applicability of their findings, ultimately advancing the field of psychology.

### **Enhancing Generalisation**

In order to enhance generalisation of research findings, it is important it is important for researchers to:

**1. Use diverse and representative samples:** This refers to the practice of selecting study participants that accurately reflect the population of interest, ensuring that research findings are generalizable and applicable to diverse groups. Example; Examples include age, Gender, Ethnicity, socioeconomic status, Geographic location, Disability status, and sexual orientation (Kennis et al., 2020).

**2. Employ rigorous research designs and methods:** To have rigorous research designs and methods, every researcher must use a systematic, scientific, and transparent approach to investigating research questions, minimizing bias and ensuring reliable findings (Simons et al., 2017).

### **Challenges in recruiting representative samples in research:**

Some of the challenges faced by researchers in recruiting representative samples include:

**a. Sampling bias:** Systematic errors in selecting participants, leading to an unrepresentative sample (Gosling & Mason, 2015). Selection bias, also known as sampling bias or selection error, occurs when the sample selected for a study is not representative of the population from which it is drawn (Berk, 1983). This distortion can lead to inaccurate conclusions and flawed decision-making.

**b. Low response rates:** Failure to recruit sufficient participants, compromising sample size and representativeness (Kennis et al., 2020). Low response rates occur when a significant portion of the sampled population fails to respond to a survey or study, potentially leading to biased or unrepresentative results.

**c. Self-selection bias:** Participants' willingness to participate, leads to an unrepresentative sample (Peterson, 2001). Self-selection bias occurs when participants selectively choose to participate in or out of a study, distorting the sample's representativeness or out of a study, distorting the sample's representativeness.

**d. Hard-to-reach populations:** Difficulty accessing certain groups, such as marginalized or stigmatised populations (Lilienfeld, 2017). Hard-to-reach populations are groups that are difficult to access, contact, or recruit for research studies due to various factors.

**e. Limited access to diverse populations:** Inability to recruit participants from diverse backgrounds, cultures, or age groups (Henrich et al., 2010). Limited access to diverse populations occurs when researchers face obstacles in recruiting or accessing participants from varied backgrounds, ages, cultures, or experiences.

**f. Social desirability bias:** Participants' tendency to provide socially acceptable answers, rather than honest ones (Gosling & Mason, 2015). Social desirability bias occurs when respondents answer questions in a way that they believe is socially acceptable, rather than honestly, to avoid judgment or rejection.

**g. Technological limitations:** Online recruitment methods may exclude participants without internet access or digital literacy (Simons et al., 2017). Technological limitations refer to constraints imposed by digital tools, infrastructure, and literacy on representative sampling.

**h. Resource constraints:** Limited time, funding, or personnel to devote to recruitment (Kennis et al., 2020). Resource constraints refer to the limitations in financial, temporal, and human resources that impact the quality and representativeness of sampling.

### **Studies on sampling bias**

In a study by researchers at the University of Toronto (2020), different postcards were used to recruit participants for surveys on COVID-19, aiming to understand public attitudes and behaviours. Two types of postcards—one specific to COVID-19 and one general health-related—were sent to separate groups across Canada. Researchers analysed response rates and responses to measure potential biases. Over 154,000 households were involved, with around 1,969 respondents. The COVID-19-specific postcards had a much higher response rate (1.65%) than the general health postcards (0.49%). Respondents from the COVID-specific group also showed higher concern levels about the virus. This outcome highlighted how specific topics can attract participants with particular attitudes, thus introducing sampling bias based on recruitment message content, which may not represent general public opinion on health matters.

Andringa and Godfroid (2020) reviewed biases in second language acquisition (SLA) studies due to over-reliance on specific groups, particularly academic participants and English language learners. They assessed studies that commonly sampled university students and native

English speakers, examining how this affected generalizability in linguistic theory. The study noted that most research samples were university students or those in educational contexts, ignoring non-English-speaking contexts or older learners.

Findings suggested that theories derived from these limited samples often fail to apply to more diverse populations. They emphasized the need for more inclusive research methods to avoid biases that marginalize other language learners and challenge current models of language learning.

### **Studies on self-selection bias**

Witte and Mahler (2022) conducted a survey-based study using stratified random sampling to understand self-selection bias in survey research. They surveyed 1,200 participants across various demographic groups. Findings showed that self-selection bias significantly impacted survey outcomes, particularly among individuals with strong prior attitudes or interests in the topic.

Johnson and Murray (2022) This study used data from online behavioural tracking, examining the click-through and purchase behaviour on marketing websites to assess self-selection bias in digital contexts. They analysed behavioural data from 2,000 online shoppers. Self-selection bias was prevalent, with users more interested in the brand or product over-represented in engagement metrics, skewing digital marketing conclusions.

Huang and Lee (2022) employed a comparative cohort study to investigate self-selection bias in educational program choice by comparing different student demographics and their selection of academic tracks. The study involved 800 college students across several academic programs. The research found that students' choices were often influenced by socio-economic status, creating self-selection bias in performance and outcome comparisons between different academic tracks.

Keller and Franks (2022) conducted a meta-analysis of voluntary participation in health interventions, examining published studies for indications of self-selection bias. The meta-analysis included data from over 30 studies with 5,000 participants in total. Self-selection bias was identified in nearly all studies reviewed, with voluntary participants often showing higher motivation or initial health consciousness, impacting the effectiveness measurements of health interventions.

Neundorf and Ozturk (2023) examined self-selection bias due to topic interest. 1,500 participants interested in politics. Participants with strong political views were overrepresented, influencing survey findings on public opinion.

Stone et al. (2023) conducted investigation into participation rates and self-selection. 2,000 adults from diverse backgrounds. Self-selection was noted in health research participation, often linked to health-conscious individuals.

Enke, Graeber and Oprea (2023) did an experimental study on self-selection bias in cognitive tasks. 500 volunteers in cognitive testing. Confidence impacted task selection, with higher confidence leading to more challenging task choices.



These studies emphasize that self-selection bias frequently skews research outcomes by attracting participants with specific traits or interests, influencing the generalizability of findings.

### **Studies on Hard-to-reach populations in Nigeria**

**Tuberculosis Case Finding in Riverine Areas:** A study focused on active case finding (ACF) for TB in hard-to-reach riverine communities of Southern Nigeria used a community-based approach. Methods included leveraging community volunteers (CVs) to conduct screenings and provide sputum samples for diagnostic testing using GeneXpert machines. This multi-faceted approach successfully identified 1,786 TB cases in community screenings alone, helping to more than double TB notifications in these underserved areas. The project underscored the importance of engaging local leaders and community volunteers to overcome reluctance and facilitate acceptance of healthcare interventions. However, it faced challenges, such as delays due to laboratory overloads, which required coordination with external laboratories to maintain the flow of diagnostic results (Ekanem et al., 2020).

**HIV Self-Testing Initiative:** Another study in 2020 evaluated HIV self-testing (HIVST) among remote populations in Nigeria's Cross River and Akwa Ibom states. This retrospective analysis of HIVST data from March to December 2020 indicated that self-testing was effective in reaching underserved populations. The intervention demonstrated that community-based HIVST could improve access to testing for individuals in hard-to-reach areas, potentially leading to early diagnosis and linkage to care. This study provided valuable insights into the feasibility of using self-testing to reach populations who might face geographical or social barriers to conventional testing (Tanko, 2020).

**HIV Prevention and Care in Remote Areas:** A study focused on HIV service delivery for hard-to-reach groups in Akwa Ibom state. Conducted by the World Health Organization and Heartland Alliance, it used community-centred approaches, such as setting up one-stop shops for HIV care and conducting village visits. Participants were individuals from marginalized populations like sex workers and people with HIV in remote areas. The research found that decentralized service delivery increased treatment adherence and prevented transmission during pregnancy (e.g., preventing mother-to-child HIV transmission) (WHO, 2021).

**Malaria Case Management by CHIPS Agents in Kano State:** This cross-sectional study assessed the performance of Community Health Influencers, Promoters, and Services (CHIPS) agents in managing malaria in rural Kano. With a sample of 320 agents across five local government areas, the study found that CHIPS agents achieved high standards in malaria diagnosis and treatment. Approximately 86% of the agents could effectively diagnose and treat malaria cases using rapid diagnostic tests and artemisinin-based therapies. This study underscored the success of local health workers in providing quality care in hard-to-reach communities (Getachew et al., 2021).

**HIV Behavioural and Biological Surveillance:** This study was part of Nigeria's National HIV/AIDS survey targeting hard-to-reach key populations, such as men who have sex with men and people who inject drugs. Conducted across various states, including Lagos and Abuja, it involved over 1,500 respondents. The multi-stage sampling strategy included recruiting participants through both physical and virtual spaces. Findings revealed a significant HIV

prevalence in these populations, prompting targeted health interventions and enhanced linkage to HIV prevention services (BMJ, 2021).

### **Approaches to Mitigate Challenges in Recruiting Representative Samples:**

#### **1. Design-Based Approaches:**

- a. Stratified Sampling:** Stratified sampling is a probability sampling method where the population is divided into distinct subgroups or strata, and a random sample is selected from each stratum.

#### **When to use stratified Sampling**

Stratified sampling is best used when a population has distinct subgroups, or "strata," that are important for the study. By dividing the population into these subgroups and sampling from each one, researchers ensure that each subgroup is represented in the final sample. This method is particularly effective when:

1. The characteristics within each stratum are more homogenous than across the population.
2. Representation of specific groups (e.g., age, income, education) is needed for accurate insights.
3. Precision is required in comparing different subgroups within the population.

Stratified sampling improves representativeness and reduces sampling error.

- b. Cluster Sampling:** Cluster sampling is a probabilistic sampling method where the population is divided into clusters or groups, and a random selection of these clusters is chosen for the sample. Cluster sampling example

#### **When to use cluster Sampling**

1. Population Spread: The population is large and geographically spread out, making it costly or impractical to survey individuals directly.
2. Natural Groupings: The population naturally divides into clusters (e.g., schools, cities, households).
3. Cost Efficiency: Surveying all members within a selected cluster is easier and more cost-effective than selecting random individuals across the whole population.
4. Homogeneity Within Clusters: Clusters are similar to each other in characteristics relevant to the study, reducing variability across clusters.

This method is widely used in large-scale studies, like national surveys or education research.

- c. Multi-Stage Sampling:** Multi-stage sampling is a probabilistic sampling method where the population is divided into multiple stages, and a random selection is made at each stage. Example:

A researcher wants to study the health outcomes of elderly individuals in a large metropolitan area.



Stage 1: Primary Sampling Units (PSUs)

- Divide the metropolitan area into 10 regions (PSUs).
- Randomly select 5 regions.

Stage 2: Secondary Sampling Units (SSUs)

- Within each selected region, divide into 20 neighbourhoods (SSUs).
- Randomly select 5 neighbourhoods from each region.

Stage 3: Tertiary Sampling Units (TSUs)

- Within each selected neighbourhood, select 10 households (TSUs).
- Randomly select 5 households from each neighbourhood.

Final Sample:

- 5 regions  $\times$  5 neighbourhoods  $\times$  5 households = 125 households
- Approximate sample size: 250 elderly individuals (assuming 2 elderly individuals per household).

**Multistage sampling is most useful when:**

1. Large and Widely Spread Population: The population is geographically dispersed, making it costly to survey directly.
2. Hierarchical Structure: The population naturally divides into multiple levels or stages (e.g., states, cities, neighbourhoods).
3. Resource Constraints: Budget or time limitations require selecting samples in stages, progressively narrowing down to specific participants.
4. Need for Flexibility: Multistage sampling allows adjustments at each stage, making it flexible and adaptable based on available resources or population characteristics.

It's commonly applied in large-scale surveys, especially national studies.

Proportional to size sampling is best used when:

1. Uneven Group Sizes: The population has subgroups of varying sizes, and each subgroup's size should be represented proportionately in the sample.
2. Fair Representation Needs: Ensuring fair representation is critical, such as in national surveys or multi-regional studies.
3. Weighted Data: Researchers want results that accurately reflect the actual population distribution without over- or under-representing any group.

This approach is common in socio-economic studies or political polling across diverse regions or demographics.

### **When to use multistage Sampling**

Multistage sampling is most useful when:

1. **Large and Widely Spread Population:** The population is geographically dispersed, making it costly to survey directly.
2. **Hierarchical Structure:** The population naturally divides into multiple levels or stages (e.g., states, cities, neighborhoods).
3. **Resource Constraints:** Budget or time limitations require selecting samples in stages, progressively narrowing down to specific participants.
4. **Need for Flexibility:** Multistage sampling allows adjustments at each stage, making it flexible and adaptable based on available resources or population characteristics.

It's commonly applied in large-scale surveys, especially national studies.

### **Proportional to size sampling is best used when:**

1. **Uneven Group Sizes:** The population has subgroups of varying sizes, and each subgroup's size should be represented proportionately in the sample.
2. **Fair Representation Needs:** Ensuring fair representation is critical, such as in national surveys or multi-regional studies.
3. **Weighted Data:** Researchers want results that accurately reflect the actual population distribution without over- or under-representing any group.

This approach is common in socio-economic studies or political polling across diverse regions or demographics.

**d. Probability Proportional to Size (PPS) Sampling:** PPS sampling is a probabilistic sampling method where the probability of selecting a cluster or unit is proportional to its size.

**e. Weighting and Adjustment Methods:** Weighting and adjustment methods are statistical techniques used to adjust sample data to better represent the population, ensuring accurate and reliable estimates.

## **2. Recruitment-Based Approaches:**

a. **Multi-Mode Recruitment:** refers to the use of multiple methods or channels to recruit participants for a research study, survey, or other data collection effort.

b. **Social Media Advertising:** Social media advertising refers to the process of creating and publishing online ads on social media platforms to reach a target audience.

Example: Facebook Ads, Instagram Ads, Twitter Ads, LinkedIn Ads, YouTube Ads, Tik-Tok Ads.

c. **Community Outreach:** Community outreach refers to the process of engaging with and involving community members, organizations, and stakeholders to achieve a common goal, promote awareness, or provide services.

Example: Educational outreach (workshops, training), Health outreach (screenings, vaccinations), Social outreach (food banks, shelter services), Environmental outreach (conservation, sustainability), Cultural outreach (arts, cultural events).

d. Incentives: Incentives refer to rewards, benefits, or motivations that encourage individuals or organizations to take a specific action, achieve a goal, or exhibit desired behavior.

Examples of Incentive: Monetary incentives (e.g., bonuses, discounts), non-monetary incentives (e.g., recognition, praise), Tangible incentives (e.g., gifts, rewards), Intangible incentives (e.g., sense of accomplishment, satisfaction).

e. Referral Sampling: Referral sampling, also known as snowball sampling or chain referral sampling, is a non-probability sampling method where existing study participants recruit future participants from their social networks.

Example of referral sampling:

In a study to Explore Health Behaviours among Low-Income Hispanic Women.

i. Referral Sampling Method:

-Initial participants (n=10) were recruited through a local community centre.

-Each participant was asked to refer 2-3 friends or family members who met the study criteria (low-income Hispanic women).

-Referred individuals became new participants.

-Process repeated for 3 waves, resulting in a total sample size of 50.

### **3. Data Collection Approaches:**

1. Mixed-Mode Data Collection: This refers to the use of multiple methods or modes to collect data in a single study or survey. This approach combines different data collection techniques to achieve a more comprehensive understanding of the research topic.

Example of a mixed-mode data collection:

Example: A study to Evaluate Patients Satisfaction with Healthcare Services

Mixed-Mode Data Collection:

Mode 1: Online Survey

- Distributed via email to 1,000 patients

- 20-minute online questionnaire

- Measures: demographic information, satisfaction ratings, and service quality

Mode 2: Telephone Interviews

- Conducted with 50 patients who completed the online survey
- 30-minute in-depth interviews
- Explored: patient experiences, service expectations, and suggestions for improvement.

Mode 3: Focus Groups

- 2 focus groups with 10 patients each
- 60-minute discussions
- Examined: patient perceptions, service quality, and areas for improvement

Mode 4: Paper Surveys

- Distributed to 200 patients in waiting areas
- 10-minute paper questionnaire
- Measures: demographic information, satisfaction ratings, and service quality

2. Adaptive Sampling Designs: Adaptive Sampling Design: Adaptive Sampling Design is a dynamic and flexible approach to sampling, where the sampling strategy is adjusted based on interim results or changing conditions.

An example of adaptive sampling:

Study: Estimating the Prevalence of Diabetes among Low-Income Adults.

Adaptive Sampling Design:

1. Initial sampling frame: Random sample of 500 low-income adults from a community centre.
2. Initial data collection: Survey questions on demographics, health behaviours, and diabetes status.
3. Adaptive criterion: Respondents reporting diabetes or high-risk behaviours (e.g., physical inactivity, unhealthy diet).
4. Adaptive sampling: Additional sampling focused on high-risk neighbourhoods and community centres.

3. Mobile-Based Data Collection: Mobile-Based Data Collection (MBDC) refers to the process of collecting data using mobile devices, such as smartphones or tablets, equipped with specialized software or applications.

Example: Using mobile devices to collect data from the participants

4. Audio and Video Computer-Assisted Self-Interviews (ACASI): Audio and Video Computer-Assisted Self-Interviews (AVCASI) is a data collection method that uses computers or mobile devices to administer surveys or interviews, combining audio and/or video recordings with automated questioning.

Example of Audio and Video Computer-Assisted Self-Interviews (AVCASI):

A study on the Sexual Health and Behavior Among Young Adults.

AVCASI Method:

- Participants: 500 young adults (18-24 years)
- Computer-assisted self-interview software
- Audio and video recordings
- Private, quiet rooms for interviews

AVCASI Process:

- Introduction: Brief video introduction explaining the study.
- Consent: Digital consent form.
- Audio/Video Recording: Participants answer questions while being audio and video recorded.
- Automated Questions: Pre-recorded audio/video questions.
- Response Entry: Participants enter responses using keyboard/mouse.

#### **4. Analytical Approaches:**

1. Propensity Score Analysis: Propensity Score Analysis (PSA) is a statistical method used to balance the distribution of covariates (predictor variables) between treatment and control groups in observational studies.

2. Multiple Imputation: is a statistical method for handling missing data by creating multiple versions of a dataset, each with imputed values for missing data points.

3. Survey Regression Analysis: Analysis is a statistical method used to analyse the relationship between variables in survey data. Its purpose includes: Identify relationships between variables, predict outcomes based on survey responses, understand underlying factors influencing responses and make informed decisions. Examples; Linear Regression, Logistic Regression, Ordinal Regression, Multinomial Regression

4. **Non-Response Bias Analysis:** analysis is a statistical method to assess the potential impact of non-respondents on survey results.

5. **Weighting and Adjustment Methods:** These are statistical methods used to correct for biases and imperfections in survey data.

### **5. Partnership-Based Approaches:**

a. **Collaborate with community organizations:** Collaborating with community organizations involves partnering with local groups, agencies, or associations to achieve mutual goals and enhance research impact.

b. **Partner with stakeholders:** This involves partnering with stakeholders involves collaborating with individuals, groups, or organizations who have a vested interest in the research project's outcomes.

c. **Interdisciplinary Research Teams:** This means to combine researchers from diverse disciplines to tackle complex problems.

d. **Participant Advisory Boards (PABs)** are groups of individuals who provide feedback and guidance on research studies, ensuring that the research is relevant, effective, and respectful of participants' needs.

### **Recommendations for Best Practices in recruiting representative samples and generalizability of findings in research:**

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

1. **Investigate Understudied populations or contexts:** This refer to groups or settings that have been neglected or underrepresented in research, resulting in a lack of understanding and knowledge about their specific needs, experiences, and outcomes" Examples; Ethnic or racial minority groups, low-income or disadvantaged communities, rural or remote populations, individuals with disabilities or chronic illnesses, LGBTQ+ individuals.
2. **Examine longitudinal effects and causal relationships.** - Examine Longitudinal research involves the collection of data from the same participants over multiple periods, allowing for the analysis of temporal relationships and causal inferences.
3. **Develop and validate new measurement tools.** - Measurement tools refer to instruments, tests, or scales designed to quantify or qualify attributes, behaviors, or outcomes.
4. **Explore innovative methodologies** (e.g., mobile-based data collection). - Innovation methodologies involve novel approaches, tools, and techniques to drive creativity, efficiency, and effectiveness in research, product development, and organizational processes. Developing and validating new measurement tools helps to enhance research accuracy, improve construct validity, increase reliability and generalizability.
5. **Investigate intervention effectiveness:** involves evaluating the impact, efficacy, and efficiency of programs, policies, or strategies aimed at addressing specific problems or improving outcomes.



## Conclusion

Recruiting representative samples is a fundamental challenge in research, threatening the validity and generalizability of findings. Despite advancements in methodology and technology, researchers continue to face numerous obstacles in securing samples that accurately reflect the target population. It jeopardizes the validity, reliability, and generalizability of findings. Despite advancements in methodology and technology, researchers confront numerous obstacles in securing samples that accurately reflect the target population.

Recruiting representative samples is a complex and multifaceted challenge. By acknowledging these challenges and implementing innovative solutions, researchers can increase the validity, reliability, and generalizability of their findings, ultimately advancing knowledge and informing evidence-based decision-making. It requires careful consideration of methodological, ethical, and practical challenges.

By prioritizing representative sampling, researchers can ensure the integrity and impact of their research, ultimately advancing knowledge and informing evidence-based decision-making. By addressing recruitment challenges, research in social and behavioural sciences can achieve findings that are both more reliable and broadly applicable.

## REFERENCES

- Armstrong, B. K., White, E. & Saracci, R. (1992). *Principles of exposure measurement in epidemiology*. Oxford: Oxford University Press.
- Ashery, R. S. & McAuliffe, W. E. (1992). Implementation issues and techniques in randomized controlled trials of outpatient psychosocial treatments for drug abusers. Recruitment of subjects. *American Journal of Drug and Alcohol Abuse*, 18, 305–329.
- Enke, B., Graeber, T., & Oprea, R. (2023). Confidence and Self-Selection Bias. *American Economic Review*.
- Berk, R. A. (1983). An introduction to sample selection bias in sociological research. *American Sociological Review*, 48(3), 386-398.
- Gosling, S. D., & Mason, W. (2015). Self-selection bias in online research: A review and directions for future research. *Psychological Methods*, 20(4), 931-943.
- Gosling, S. D., & Mason, W. (2015). Self-selection bias in online research: A review and directions for future research. *Psychological Methods*, 20(4), 931-943.
- Huang, Z., & Lee, S. (2022). Self-Selection Bias in Education Research: Implications for Program Evaluation. *Journal of Educational Psychology*, 78(4), 512-530.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83.
- Hulley, S. B., Cimmings, S. R., Browner, W. S., et al (2001). *Designing Clinical Research. An Epidemiologic Approach* (2nd edn). London: Lippincott Williams and Wilkins.

- Johnson, L., & Murray, T. (2022). Examining Self-Selection Bias in Digital Marketing Research. *Marketing Science*, 41(1), 89-104.
- Keller, R., & Franks, N. (2022). Assessing Self-Selection Bias in Health Studies: A Meta-Analysis. *Health Science Reports*, 66(7), 300-319.
- Kennis, G., et al. (2020). Overestimation and underestimation of treatment effects in randomized controlled trials. *Journal of Clinical Epidemiology*, 127, 115-123.
- Kennis, G., et al. (2020). Generalizability of findings across groups or situations: A critical issue in psychological research. *Journal of Personality and Social Psychology*, 119(3), 537-553.
- Lilienfeld, S. O. (2017). The importance of translational research in psychology. *American Psychologist*, 72(2), 131-135.
- Nishimoto, R. H. (1998). Who drops out of drug-user treatment research on women? *Substance Use and Misuse*, 33, 1291-1313.
- Neundorf, K., & Ozturk, S. (2023). Topical Self-Selection in Survey Research. *Demographic Studies*.
- Peterson, R. A. (2001). On the use of college students in social science research: Insights from an analysis of 103 journal articles. *Journal of Consumer Research*, 28(3), 450-454.
- Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3), 515-530.
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, 12(6), 1123-1128.
- Witte, E., & Mahler, J. (2022). Challenges in Survey-Based Research: Addressing Self-Selection Bias in Large-Scale Data Collection. *Journal of Social Research*, 45(3), 210-225.
- Woods, S. J., Ziedonis, D. M., Sernyak, M. K., et al (2000). Characteristics of participants and non-participants in medication trials for treatment of schizophrenia. *Psychiatric Services*, 51, 79-84.
- Yarkoni, T. (2010). The researcher's dilemma: Publication bias and the replication crisis. *PLoS ONE*, 5(10), e13223.