

QUESTIONNAIRE SURVEY METHODOLOGY IN EDUCATIONAL AND SOCIAL SCIENCE STUDIES

Reuben Bihu

ABSTRACT: *Questionnaire surveys dominate the methodological designs in educational and social studies, particularly in developing countries. It is conceived practically easy to adopt study protocols under the design, and off course most researchers would prefer using questionnaire survey studies designed in most simple ways. Such studies provide the most economical way of gathering information from representations giving data that apply to general populations. As such, the desire for cost management and redeeming time have caused many researchers to adapt, adopt and apply the designs, even where it doesn't qualify for the study contexts. Consequently, they are confronted by managing consistent errors and mismatching study protocols and methodologies. However, the benefits of using the design are real and popular even to the novice researchers and the problems arising from the design are always easily address by experienced researchers.*

KEY WORDS: survey research, questionnaire survey, sampling procedure, sample size, questionnaire survey data, benefits of questionnaire survey, problems with questionnaire survey.

INTRODUCTION

The aim of this literature research was to describe how questionnaire surveys can be used and related implications of design applicability in educational and social studies. The need for this review arose from the eager to add value to an understanding of questionnaire surveys. It is an endeavor to also catalyze educational and social researchers to think more, weigh, choose, improve and apply in their survey scenarios the appropriate sampling techniques and the sample sizes. Therefore, the content of this paper includes the methodology used for review, comprehending survey research, questionnaire survey, sampling techniques, decision on sample size, the benefits and problems of using questionnaire survey studies, and the conclusion.

LITERATURE REVIEW METHODOLOGY

I adopted with improvements the approach used in Bihu (2020) to review and use cited and referenced sources. The guiding questions were: (1) What does survey research in educational and social context entail? (2) How can a questionnaire survey study be effectively

conducted? (3) How can a decision be made on the appropriate sample size? (4) What are the determinants of the appropriate sample size? (5) What are the benefits of using questionnaire survey design in positivist paradigms? (6) What are the problems associated with using questionnaire survey methods in positivist paradigms? I used Zotero software to collect soft documents reviewed which were co-analyzed with MAXQDA 2020 software.¹ I actually used key words “Survey research”, “Questionnaire survey”, “Sampling procedure”, “Sample size”, “Questionnaire survey data”, “Benefits of questionnaire survey”, “Problems with questionnaire survey” to generate information from the soft documents in online sites. Collection of review documents took about five days with 40 documents listed in analytical software. Documents were screened against the criteria for validity, reliability and data quality so that only twenty seven (27) documents were used. A document could be used in more than one sections of the review. Document coding took about two days.

Summary of the review

Themes reviewed	Citations
Survey research design	12
Questionnaire survey method	12
Sampling of respondents	09
Decision on the sample size	08
Determinants of the sample size	09
Benefits with survey research	09
Problems with survey research	10

Survey Research Design

A survey design means a data collection way used to carry out a survey research (Glasow, 2005). It involves a closely look, examination and description of a particular issue, subject or an object². It specifically refers to the design under which a particular issue or problem is studied, in a positivistic context³. Survey research in educational and social contexts refers to “the collection of information from a sample of individuals through their responses to questions” (Check & Schutt, 2012 cited in Ponto, 2015). The two primary features of a survey research are the use of pre-defined samples and well-structured instruments.

Survey research essentially entails methods which emphasize research practices leading to quantitative analyses such as practices involving questionnaires, structured interviews, or data from published statistics that are founded in positivistic paradigm (Gabble, 1994; Nayak & Narayan, 2019). A positivistic paradigm is a philosophical thinking which rationalize that truth and reality is available to be tested from the outside world of the respondent or subject

¹ MAXQDA, the Maximum Qualitative Data Analysis software allows documentary review through thematic coding and analysis, for qualitative and quantitative purposes

² Surveys originated in 1930s being applied by the ‘Gallup’ and ‘Roper’ polls in public opinion research, USA. Source: Walden (1996) and Zutshi, Paris and Creed (n.d).

³ Other designs under positivism include experiments (Punch, 2009; Thomas, 2009), quasi-experiments (Bryman, 2008), longitudinal studies (Thomas, 2009) and cross-sectional designs (Bryman, 2008)

(Bryman, 2008). Implicitly, the paradigm prefers and deploys techniques other than subjects' voices to infer on the cases or problems under investigation. Being a quantitative design, a survey research is concerned with measures of causality between dependent and independent variables (Cresswell, 2014). Particularly, researchers opting for the design apply it in studies resulting into some statistical inferences⁴. In the process, surveyors use statistical techniques to analyze the self-reported data from sampled respondents and use data for generalizations to wider contexts. In this case, independent and dependent variables provide the context and premise to define the scope of study (Glasow, 2005), in terms of what questions are answered by the methodology. In particular, most of the surveys use structured instruments such as questionnaires with standardized questions aiding to such analyses. In fact, the surveyors utilize different statistical skills with competencies and profound cognitive ability sufficient to make the survey process successful through proper inclusion of content and reliable responses for research work with good quality findings (Ha et al., 2015).

Questionnaire Survey Methodology

The findings from a survey research should involve positive consequences of multiple interactions of involved parties, though primarily upheld as the responsibility of the researcher. Actually, high quality survey study design is achieved through accommodating inputs from targeted respondents, users of the study findings and the expatriates carrying out respective surveys (Glasow, 2005). In practice, the use of questionnaires in surveys allows for the achievement of generating data in the most cost-effective and efficient way. In particular circumstances, such surveys allow the instruments to be easily piloted, implemented, data analyzed and findings shared to stakeholders. As such, questionnaire surveys form a popular data collection method for academic and marketing research (Regmi et al., 2016). The questionnaire survey method involves a several phases' process including defining the research aim, identifying the study population, deciding the sample size and data collection technique, designing the questionnaires, piloting the instruments, implementing the survey, and analyzing the data (Burgess, 2001; Gray, 2004). The overall protocol can be comprehended in Figure 3:1.

⁴ Made based on relevant statistical tests such as Spearman's correlations, Chi-Square test, etc.

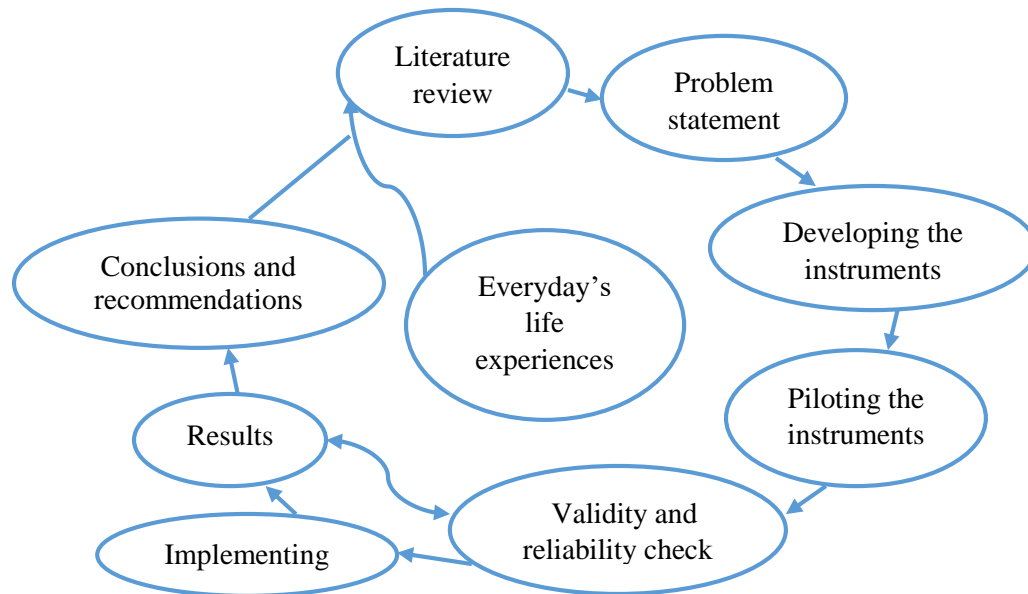


Figure 3:1 Functional protocol in questionnaire survey methodology. **Source:** Adopted from Burgess (2001)

Essentially, problems researched are conceived and originated by researchers from everyday life experiences in human societies, specifically, in business, academic and political settings, and daily life activities. Such problems create the aims and purposes to pursue for their solutions through research endeavors, resulting to empirical knowledge and theories (Cresswell, 2014) positioning tested solutions to problems. Such problems have had justifiably been stated and restated in advance by the researchers through intensive literature reviews for validation. In particular, literature reviews function to assure researchers on the emerging research problems as gaps or needs which require bridging of either knowledge content or solutions to pre-existing problems that required solutions or awaited for novelty in pursuit of quality situations or interventions in life endeavors. In fact, the knowledge of the subject matter content around the literature gap enables researchers to define the problems clearly. The attributes of the studies are outweighed by the quantitative findings leading to generalizations. Within the framework, a research problem becomes essentially a question of attention to the researcher, which he or she must importantly answer through appropriately designed method

Questionnaire survey data

Structured survey instruments lead to generation of research data in one or more of the four forms. The first form is Data on a Ratio Scale (DRS). Such data have constant interval size between successive units, and true zero point which make it possible to establish a ratio constituting actual amounts of variables (Kothari, 2004). Examples of DRS include temperature on a Kelvin scale, years married, income, and reaction time (Matthews, 2018). Using the DRS researchers should be aware that the absolute zero of some data such as time and length can be measured while for some case like that of the temperature on Kelvin scale cannot, and therefore it remains theoretically conceptualized by the researcher (Kothari, 2018). The second form of data is the Data on an Interval Scale (DIS). The DIS have constant interval size between successive units but lack true zero point, and it is not possible to establish a ratio. Examples of such data include temperature on Celsius scale. The intervals enable researchers to do basic mathematical operations such as addition and subtraction with ordinal measurements (Matthews, 2018). It can be used with mean⁵ as the appropriate measure of central tendency and standard deviation⁶ as the measure of dispersion. The DIS can also use product moment correlation techniques and tests for statistical significance including 't' test⁷ and 'F' test⁸ (Kothari, 2004:72). The third form is the Ordinal Scale Data (OSD). The OSD involves observations or data which can be put in order from the lowest to highest but which do not have constant interval between them. The example of the OSD are ranked data such as general time of day- morning, afternoon, evening, and attitudes or opinions- strongly dislike, moderately like, greater than, and less than (Kothari, 2014; Matthews, 2018). Such data only give the relative differences among the values and order them appropriately. The fourth form is the Nominal Scale Data (NSD). The NSD are the data that classify variables by some quantities or attributes. Variables are not measured but put in certain categories. The examples of NSD include classifications of counted data according to social class, gender, political party affiliation, religion and education (Kothari, 2014; Matthews, 2018). Moreover, it seems that overreliance on such data classification can be misleading in some contexts due to a sense that good data analysis does not necessarily assume data types (Matthews, 2018). Accordingly, in many cases categories may not necessarily describe fixed attributes.

Questionnaire survey data patterns have two important attributes that a survey researcher has to put at the heart of the analytical endeavor- data accuracy and data precision. This is particularly why the researcher has to determine the measures of central tendencies and dispersions for a particular research (Ali & Bhaskar, 2016). Accuracy refers to the closeness of a measurement to the true value of the variable being measured, and hence, relates to validity of findings (Bryman, 2008). Precision refers to the closeness of repeated measurements of the same quantity, and hence, relates to reliability of findings (Kothari,

⁵ The arithmetic average (Ali & Bhaskar, 2016:663)

⁶ The square root of variance of measures of a parameter in a population (ibid:664)

⁷ When sample size is 30 or less and population variance is not known (Kothari, 2004:160)

⁸ Based on F-Distribution and used to compare statistical analyses (Kothari, 2004:196), others are χ^2 -test, z-test (ibid:195)

2004). The quality of findings from which researchers draw conclusions and make recommendations are validity and reliability dependent. A good data must be both valid and reliable. This can be confirmed through repeated procedures attempting to generate findings on the same case in different trials (Cresswell, 2014). This matter has caused a custom of testing and retesting of methodologies and relevant findings. Mostly, life experiences tend to generate research demands arising as gaps of empirical knowledge from such previously published research findings, so much so that the whole human life settings become immersed in a research driven social, economic and cultural environments.

Questionnaire development and implementation plan

The requirements for a good quality questionnaire survey method dictate for an appropriate and standard instrument, carefully made and used to answer the question at hand. If the researcher is not in capacity, non-conversant or inexperienced with developing questionnaires, would be advised to adopt or adapt the standard questionnaires that have been developed and tested already in the respective research discipline. This approach to the design responds to the compelling high quality need which requires that the instrument be developed, tested, retested and standardized before being implemented. Nevertheless, prior to developing the questionnaire, the researcher must indeed consider the kind of the population and representative sample to participate in the research, and how the responses should be coded and analyzed. In fact, for data quality assurance, the questionnaire must be piloted before it is used in the main survey (Burgess, 2001). This requirement should be non-waived to ensure maximum validity and reliability. Much the merits piloting has, gives a picture in advance of what kind of data and the extent of valid and reliable data can be expected (Rattray & Jones, 2007). It also determines the appropriateness and relevance of sampling technique planned for implementation of the survey (Ponto, 2015). Viewing on the merits of questionnaire piloting, Bryman et al. (2008) maintain that data validity and reliability, and sampling technique appropriateness and relevance are very important in the same way as the explicitness and transparency of the procedural guidelines used to offer items of the instruments do. Indeed, a questionnaire has to be piloted and standardized for increasing validity and reliability of findings, and explicitness and transparency of its items' procedures provided to respondents.

In spite of that, researchers have many options available for choice on implementation of the instruments especially the mechanisms of administering. The decision on how the survey instrument has to be administered, be it face-to-face, by post, mails or online may be dictated by the nature of respondents, sample size and response and return rate the researcher plans for implementation of the study (Glasow, 2005). In this case, the methodological coherence demands researchers to have the implementation plan clearly thought earlier at the start of the design process. In a real sense, during the implementation, the researcher is obliged to monitor the appropriateness and relevance of the implementation plan to the espoused design to enhance expected end-results of the process.

Sampling Procedure

A survey researcher finds it imperative to consider the units of analysis for the study (Aldhein, 2020). Particularly, considerations are made on whether the prospective respondents could be individuals, offices, departments, entire firms or organizations (Glasow, 2005). The choice of the functional units can be influenced by the nature of the study, the instruments to be used and the end-results forecasted. In fact, the instrument specifying particular individuals within the firm or organization can differ from those suited to cover offices, departments and firms or organizations in general. Such kind of rationality can also guide researchers in establishing whether the study should use informed individuals, that is, experts in the area investigated (Hasson, Keeney & McKenna, 2000). As such, with sampling frame determined there are many options for sampling techniques which survey researchers use to obtain the representative samples, that can be used singly or in combinations within a study. Mandatorily across all the techniques, is the requirement that the frame should constitute all members of the research population.

Sampling techniques

Normally, the sampling technique used depends on whether the sampling frame is determinate or indeterminate. Notwithstanding, across all the techniques, the frame should constitute all members of the research population. In short presentation on one hand, the determined sampling frame allow applicability of *Simple random sampling technique* which gives equal weight of being selected to all subjects within a target population. Normally, random samples can be obtained by lottery, computer and tables of random numbers (Bird, 2009). Secondly, *Cluster sampling technique* can also be used. In this technique, first groups of clusters are selected and then individual participants are selected from these groups (Bird, 2009; Dornyei, 2012). Thirdly, the researcher may use *Longitudinal sampling technique*. In this technique, the same participants from an original sample are studied on more than one occasion (Bird, 2009). Fourthly, *Spatial sampling technique* may be used which is about sampling people who have temporarily congregated in a specific space (Bird, 2009; Dornyei, 2012). Also, *Purposive sampling technique* guide researchers to purposively select participants who are thought to be relevant to the research (Bird, 2009). Moreover, the *Census technique* requires that all prospective individuals of the population be included in the study. However, Lupu and Michelitch (2018) caution on the common challenge in using census techniques that their data may be unreliable, outdated, or unavailable in some aspects. These are basically the common used probability sampling techniques. On the other hand, if the sampling frame is undetermined convenience, snowball and quota sampling techniques can be used. In *Convenience sampling technique* all subjects that the researcher meets in a particular period of time are considered and involved in the research as respondents (Bird, 2009). Its alternative is the *Head count technique* in which a particular social group is counted and involved at once in a specific time and location, for example, street children count in a

specific date at Ubungo⁹ area. Convenience samples are usually purposively obtained (Dornyei, 2012). The *Snowball sampling technique* allows for first participant to recommend other participants known to him or her who meet the research criteria (Bird, 2009). In *Quota sampling technique*, a 'quota' of participants to be chosen from a specific group in the studied population is predetermined (Bird, 2009).

Summary of sampling techniques

Frame	Sampling Technique
Determined	Simple random sampling
	Longitudinal sampling
	Spatial sampling
	Cluster sampling
	Purposive sampling
	Census sampling
Undetermined	Convenience sampling
	Snowball sampling
	Quota sampling

In fact, simple random and purposive sampling techniques are the most frequently used in educational research methodologies. Significantly, the techniques enable educational and social researchers to effectively obtain the appropriate participants for research topics (Kothari, 2004; Jupp, 2006). Nevertheless, the sampling techniques used in researches may also be context specific and may be influenced by researchers' decisions on the size of the sample estimate to achieve an adequate representation of the research population. Also, the choice of the media through which the survey will be administered may influence researcher's decision on the sampling technique and the sample size (Glasow, 2005). In fact, the decision on whether the survey has to be administered face-to-face, by post, mails or online may dictate the kind of sampling technique the researcher has to adopt as well as the sample size the researcher has to cover. The best approach is to adhere to the principle of parsimony that the most cost effective and efficient way of collecting accurate and reliable data in a short period of time should be the most preferred. For example, most of the online administered survey tools may involve a greater number of respondents and cover larger study area in a short time with little financial resource input than face-to-face and telephone interviews.

Deciding on Appropriate Sample Size

The critical concern of the discussion on this matter results from the fact that survey researchers may use misguided approaches to set the sample sizes. For example, Cresswell (2014) notices that a researcher may select a sample size that is unusual, typical based on

⁹ In the United Republic of Tanzania (URT)

past studies, or base the sample size simply on the margin of error e they are willing to tolerate, or just a fraction of population, say 10%, without putting forth the analysis plan for their studies. In another case, Memon et al. (2020) explains that researchers should use sample-to-item and sample-to-variable ratios, and that the least ratio should be at least 5:1 and the most preferred should be around 20:1. This means, if a research put in list 30 items, questions or variables, then, he or she should use 150 participants for the first ratio and 600 participants for the second ratio. In this approach, the logic behind the decision remains void of a scientific speculation as to why. Notwithstanding, sample selection approach should base on technical factors such as research design, the size of the research population, its nature of homogeneity, the sample administration media, related cost, and the required degree of precision (Glasow, 2005). So, during the sampling process, a researcher has to put into account all such factors in order to ensure good sample estimation (UNODOC, 2018). Then, he or she may decide to use “census” for small populations, “sample size of a similar study” already done, “published tables”¹⁰ or “using formulas to calculate” it (Singh & Masuku, 2014: 10-12). Even though errors may exist in sampling processes, the maximum acceptable sampling error is given by the Kothari’s (2004:176) relation:

$$e = \frac{z \cdot \delta_p}{\sqrt{n}}$$

Whereby e is the maximum absolute error that can be tolerated, z is the standard normal variate, n is the sample size and δ_p is the standard deviation of the population¹¹. At a significance level (α) of 5% type 1 error ($p < 0.05$) the value of z (also denoted as z_α) is 1.96. Also, for control studies linking between scenarios, z (also denoted as z_β) is 0.84 for statistical power of 80%, the power being determined by the researcher. However, most studies are conducted at α below 0.05, and therefore z_α of 1.96 is the most frequently applied in the formula¹². On this matter, Cresswell (2014) uses Fowler’s (2009) research guide to insist on three important issues to be observed by the researcher while deciding on the sample size. Firstly, the margin of error e that can be tolerated should be determined, for example say +/- 3% confidence interval, that would represent how accurate the data generated by the sample correlate to responses from the entire population. Secondly, the confidence level α for the latter e should be determined, for example, say 95 out of 100 times, or a α of 5% chance. Thirdly, the researcher has to estimate the percentage or proportion of success p of the sample that will respond in a given way. The p of 50% or 0.5 is the most conservative since subjects could respond either way (50/50).

Then, the researcher may decide on the size of the sample depending on the nature of the

¹⁰ Some tables on samples sizes selection listed by Singh and Masuku (2014: 10-12)

¹¹ The value of standard deviation is usually taken from previous studies or determined through piloting.

¹²The value of z is determined as the ratio of the difference between the presumed probability of the score and the mean score to the standard deviation

population to be studied. Some research populations involve individuals with no boundaries, scattered like stars in the universe, and are therefore considered infinite (Kothari, 2004)¹³. Conversely, some other research populations may be finite, that is, may be determinate universes¹⁴. However, sampling from finite populations should be done without replacement, or otherwise it may turn up to be infinite (Kothari, 2004).

Sample size from infinite population

Firstly, in quantitative studies the sample size (n) from infinite population may be determined by the Kothari (2004:176) relation:

$$n = \frac{z^2 \delta_p^2}{e^2}$$

Implicitly, estimation of the sample size in quantitative measures involving for example exam score rates, is dependent of the statistical speculations and assumptions determined by the researcher as deemed significant to address the issues of data quality. The researcher has to decide on the value of e and establish the δ_p value for the sample size.

Secondly, for qualitative studies involving infinite populations, the sample sizes can be as well established based on the Kothari's (2004:179) relation:

$$n = \frac{z^2 p \cdot q}{e^2}$$

Whereby p is the expected proportion of success based on previous or pilot studies and $q = 1-p$. For example, a decision on the sample size from a population of peasants or laborers in informal works in a theoretical village *Bisheshe* in the year 2020 may be theoretically determined. Firstly, it should be considered that it is not known how many adults eligible for labor are available in the year, in a sense that no on-time data available for that context. Then, consider that based on the previously 2012 published National Bureau of Statistics (NBS) census information¹⁵, the actual number of labor force in the village may not exceed 20% of the population. So, if the study can tolerate the maximum absolute error e of 5% ($e < 0.05$ type 1 error) (confidence interval) at α value of 5% (confidence level) with z value of 1.96, then the sample size would be 246 respondents decided as:

¹³ Infinite populations: e.g. listeners of a specific radio programme, throwing of a dice, etc.

¹⁴ Finite populations: e.g. number of workers in a factory, the population of a Municipality or City, students registered in a University, teachers employed in a district, etc.

¹⁵ Of the URT

$$n = \frac{1.96^2 \times 0.2(1.0 - 0.2)}{0.05^2}$$

But what if the researcher can only tolerate the maximum error e of 3% under similar other conditions? Definitely, she or he could have to work out for the problem with a sample size of about 683 individuals. This means that increasing the sample size is directly related with increase in precision.

Moreover, the sample size for case-control qualitative studies can be obtained according to the Sharma et al.'s (2020) relation:

$$n = \frac{(r+1)(p^*)(q^*)(z_\beta + z_\alpha)^2}{r(p_1 - p_2)^2}$$

Whereby, r is the ratio of number of control to cases, p^* is the average proportion exposed, $q^* = 1 - p^*$, p_1 and p_2 are proportions in cases and control respectively, z_β is the standard normal variate for power¹⁶ and z_α is the standard normal variate for significance level¹⁷. Also, it is determined by Sharma et al. (2020) that:

$$p^* = \frac{p_1 + p_2}{2}$$

For example, assuming that a social researcher wants to know the link between investors' expectations of future outcome before entry into the business and achievement perceptions after some time. The researcher can categorize investors who perceive higher achievement and non-achievers. The comments and opinions from both groups may be studied, observations on influences in both groups may be made, and the odds ratio r can be determined. If the researcher fixes the power of the study at 80% while assuming proportions in case and control based on previous studies to be 0.38 and 0.25 respectively, and considering that there is an equal number of case and control, then, the value of p^* is 0.315 obtained as $(0.38 + 0.25)/2$ and the sample size n for the study would be 200 respondents decided as:

$$n = \frac{(1+1)(0.315)(1-0.315)(0.84+1.96)^2}{1(0.38-0.25)^2}$$

Sample size from finite population

The proportion of the sample size for research topics with finite populations can be obtained by using several approaches. One of the most common approach is the use of the Kothari's

¹⁶The common z_β values are 0.84, 1.28, 1.65 and 2.33 for the statistical powers of 80%, 90%, 95% and 99% respectively (Sharma et al., 2020)

¹⁷ Common z_α values are 1.96, 2.58 and 3.29 for α of 5%, 1% and 0.1% respectively (ibid; Cresswell, 2014)

(2004: 179&176) relations:

$$n = \frac{z^2 \cdot p \cdot q \cdot N}{e^2(N-1) + z^2 \cdot p \cdot q}$$

And an alternative would be the use of standard deviation if it is predetermined:

$$n = \frac{z^2 \cdot N \cdot \delta_p^2}{e^2(N-1) + z^2 \cdot \delta_p^2}$$

Whereby, N is the population size. The determined e value compels the researcher to state the agreement or benchmark performance, that is, α value above which no more sampling mistakes can be committed. It is from such values of α and β the standard variate z_α and Z_β are determined, respectively. The commonly used α value in educational and social studies is 5% and may be as low as 0.1% and below in biological system's research.

Moreover, Memon et al. (2020) suggest for use of online calculators as an alternative to determine the minimum sample size required especially with the values for α , e and N available.

Determinants of sample size

Survey researcher's decision on the sample size may be influenced by one or more of several factors. Firstly, it is important for the researcher to think in advance about the type and nature of the respondents he or she is about to involve in research. Specifically, some respondents may be difficult to acquire or may logistically demand a long period of time to have them responded to the research instrument. Examples of such respondents would include busy key respondents like officers, and vulnerable populations like the handicapped requiring some structural support to respond to the research instruments (Bihu, 2020). In some other special cases where researchers would be interested at researching the experiences felt by unique subjects such as drug addicts, victims of rape, assaults or violence in homes and bullying in offices, deciding on the sample size for survey research would be more challenging. This in particular, is a point where researchers faultily decide to apply surveys in contexts suited for qualitative case studies and relevant sampling techniques.

Secondly, the researcher should consider whether the research universe is finite or infinite, and homogeneous or forming different categories or sub-groups (Cresswell, 2014; Kothari, 2004). In fact, the homogeneous populations are subjected to the rules of sample decision different from the rules guiding determination of sample size from finite populations. In spite of that, if the research population does not constitute homogeneous group, and strata present in the population have to be all involved in the study, then the researcher has to decide which

stratum should take which proportion and why. However, if the strata have to be involved all randomly equitably, the proportional allocation technique suggested by the Kothari's (2004:63) relation is relevant practically for determining the proportion of the representation in each stratum:

$$P_i = n \cdot \frac{P}{N}$$

Whereby P_i is the sample proportion in the stratum, n is the working sample size, P is the size of the individuals in the stratum i , and N is the total population size. Each stratum contributes to the operational sample size according to the proportion of its size in the total population. For example, a study of the school district may constitute educational officers, teachers and students in a bigger population. In this case, a researcher may have definite sub-populations for students, teachers and officers.

Thirdly, the desired degree of precision dictates the decision on the sample size for the research. It is increased by increasing the sample size especially in a case of larger variance (Kothari, 2004). However, increasing the sample size has cost implications on cost sensitive research and may enhance systematic bias. As Kothari (2004) has suggested, using a better sampling design with a smaller sampling error for a given sample size at a given cost would be the effective way to increase precision.

Fourthly, the type and level of data analysis plan may dictate the sample size to be involved. Quantitative analysis is concerned with grouping numbers and testing the associations and correlations among groups of responses, and the causality among variables (Cohen, Mannion & Morrison, 2007; Kothari, 2004). As such, the analyses would require a particular sample size to achieve the analytical goals. In this context, Cohen, Mannion and Morrison (2007:102) have insisted that "survey research should have no fewer than 100 cases in each major subgroup and twenty to fifty in each minor subgroup". When the cases are fewer than such theoretical assumptions, the analyses and tests of statistical significances become under severe limitations. On this matter, Burmeister (2012) suggests that for the regression model, the sample size obey the previous 20:1 ratio approach. Additively, Memon et al. (2020) suggest that the sample size should not be less than 50 participants for exploratory factor analysis, at least 200 for Pearson correlation analysis and very small for piloting.

Also, a plan to involve respondents relevant to researched matter may influence decision on the sample size in a particular way. Implicitly, where the true population is not well known or understood, or when respondents are unable or unlikely to respond to written surveys, researchers are inflicted by respondents' representation of populations in questionnaire survey studies. Several times, researchers may choose to reduce the sample sizes or involve irrelevant subjects just to avoid discontinuity of the process. The advice has been to understand well the research population at the planning stage, or opt for other methods such

as for face-to-face or in-depth interviews (Bihu, 2020; Glasow, 2005) to ensure high return rate and completion.

Moreover, time scale and cost of implementation have significant explicit implication on the decision on the sample size for a survey research (Cohen, Mannion & Morrison, 2007). If the time available is too short, the research process would be constrained by reduced possibility of increasing the sample size for the research. On the other hand, extended period of time would mean use of more resources that may increase the cost to exceed the capacity level. Nevertheless, surveys strive to obtain responses from a large number of respondents (Ha et al., 2015) in short time and with fewer resources than would be for similar sample sizes with any other qualitative method.

Summary of determinants of sample size

- i. Ability of the researcher to obtain respondents
 - ii. The degree of population stratification
 - iii. Desired degree of precision
 - iv. Statistical power required
 - v. Relevance of units of analysis
 - vi. Time scale and the cost for the research project
-

Some Benefits of Using Questionnaire Survey Method

Questionnaire survey method shows multiple benefits when adopted with appropriate sampling techniques. The first benefit is that it allows for generalizability of the findings to the wider population studied (Ha et al., 2015). With appropriate representation, a studied sample can describe in details the phenomena pertaining a particular population in question (Ponto, 2015). Implicitly, it does not require researchers to reach all members of the research population as it would have been in census studies, but instead, scientifically decided samples can give estimate information pertaining every member of the population. This representation work to save time and other resources that would be incurred in bigger operations. However, the samples used must adhere to universal research rules of generalizing the findings to the wider population (Delice, 2010).

Secondly, it enables researchers to gather data on a one-shot basis as questionnaires are administered to respondents at once after its validity and reliability check. This also work to save time and other resources that would be incurred with other methods requiring extended time, space and more financial resources. For example, in the case of self-administered questionnaires, researchers may not need to hire the interviewers as in the case of one-to-one interviews. It can also be posted to online channels like websites and other application software to allow multitudes of the population members to put in responses. The use of emails can also effect the same advantage with documents sent in Microsoft Word or other soft document formats enabling the sampled respondents to respond and send back the filled-in copies. Such online mobile tools provide the survey researchers with options to reach

respondents in extended scales geographically, socially and culturally with minimal time and other resource inputs (Evans & Mathur, 2015; Nayak & Narayan, 2019). Strictly, researchers must prepare in advance the elaborate mechanisms for maximum return rates to ensure internal efficiency of the research process. In reality, online approach does not necessarily impose time constraints to the research process, and can result to more truthful data than data generated from respondents under physical presence of the researchers. Therefore, the method provides the most economic and efficient way embracing the positivist principle of parsimony.

Thirdly, the method gives the most effective way of ascertaining correlations and relationships between study variables (Ali & Bhasker, 2016). Variables providing the parameters measured, have influences among themselves which can be well traced using questionnaires with question settings allowing for such kind of statistical inferences on the causality among such variables. The questionnaires provide the most efficient platform for such quantitative analyses. The good relevant trait within the method is its possibility of effectively and efficiently coding and quantifying the responses made in a highly structured way to allow for statistical testing. The ability of undergoing statistical testing means increasing the credibility of the study arising from the opportunity to generate data in the most scientific way as far as the positivistic controversial stance is concerned in a paradigm crisis (Gray, 2004).

Fourthly, it is easy to do analysis and make visualization of data. Survey researches provide descriptive and inferential explanatory information on the studied cases numerically and statistically (Kothari, 2004). Data visualization gives an effective overview of ontological construction aiming at domain and data understanding (Davies, Studer & Warren, 2006). This in turn gives the easiest way of estimating, understanding and describing the findings from research studies. Consequently, large information and big data can be reduced into grouped categories of presentation that is less demanding in terms of time, space and other resources (Muniswamaiah et al. 2020). As a matter of fact research results presented in graphics, tables, charts and relevant theoretical analyses for example, are easily understood through increased visibility that enhance mental captures in the most simple way compared to extended narratives. Technically, survey research always ends up producing vast amount of uncluttered data. Basically, the researcher presents to the respondents the instruments with most conspicuous, screened, weighed and checked items, of which analyses of findings do not raise much concerns of single themes impeded by too many objects, much details or mixed elements. In its construction, a questionnaire involved can cover every important aspect of the study with a reasonable number of questions.

Also, survey research uses more accurate instruments with quality tested and enhanced through piloting and revisions. As a rule, the instruments are tested for construct, content and criterion validity before they are implemented (Marczyk, DeMatteo & Festinger, 2005). This goes in line with checking whether the same instruments can produce reliable data.

Instruments used in surveys therefore assures that data generated are valid and can apply to measure what the researcher intended to measure even in different contexts if the same case is studied under similar methodology.

Moreover, survey research develops data with statistical patterns that can be used to generate a theory. In this case, the researcher may begin with a hypothetical theory or predictions (Marczyk, DeMatteo & Festinger, 2005) which can be written as the research process reveals the theoretical underpinnings and conceptual patterns. Though commonly being undertaken as a theory-first practice using a deductive reasoning method, it can give a room to the emergence of new data patterns that can be used to develop a new theory. Indeed, well thought and practiced survey methodology under skilled researcher can guide the endeavor to the development of general principles or laws that attempt to explain how and why something happens or happened. This emerges as a fundamental truth or proposition that guides for a system of belief, behavior and reasoning.

Summary of benefits of using survey studies

- i. It allows for generalizability of the findings to the wider population using appropriate samples
 - ii. Gathers data on a one-shot basis to save time and resources
 - iii. Gives the most effective way of ascertaining correlations and relationships between variables
 - iv. Provides descriptive and explanatory information numerically and statistically
 - v. Gives high possibility of presenting data that are uncluttered
 - vi. Uses more accurate instruments with quality enhanced through piloting and revisions
 - vii. Develops statistical patterns that can be used to generate a theory
-

Selected Problems Associated with Survey Studies

Survey methodologies suffer from several problems. Firstly, some attempts exhibit mismatches between paradigms, research questions and instruments (Young, 2016). A good survey research requires that coherence and alignment be ensured from the epistemological and ontological assumptions pre-stated, the design and method used, the nature and type of questions studied to the instruments used (Cresswell, 2014). However, much contradictorily it becomes, several times researchers opt to study topics without causality among variables using the methods of such kind suited for testing causality among variables, and vice versa. Such research topics and hence questions in execution which are well explored under interpretivism and critical theory paradigms may end up with invalid and unreliable findings under survey methodologies. Bryman (2008:15) argues on the survey methodological contradictions that “in fact, research is full of false starts, blind alleys, mistakes, and enforced changes to research plans.” It is therefore imperative to monitor and do some necessary

improvements of the planned methodological undertakings during piloting phase, a stage where success in methodological changes may be granted with higher success.

Secondly, survey research are prone to instrumentation systematic errors in cases structured instruments may not reflect the levels of accuracy and preciseness in enquiries thought-out in certain studies (Marczyk, DeMatteo & Festinger, 2005). In particular, if not well moderated may end up with documenting data resulting from faulty measures. In this case, researcher's experience may affects the relevance of the choice of the research instrument that can appropriately and accurately generate data of a particular type and quality involving categorized respondents.

Thirdly, technicalities involved in data analyses pose limitations to most of researchers opting for doing research topics which demand advanced statistical computations (Kothari, 2004). Very often, researchers would do partial analyses on cases thereby creating the methodological deficiencies pledging for analytical unsaturation in study findings. However, researchers would be advised to train and match with skills required specifically for the level of the analyses planned. This is a matter of researcher's preference and choice, and the nature of research topic. Every researcher could do any kind of statistical analysis if he or she train sufficiently. I suppose, if the mountain was high, the crane could not hike it. Implicitly, no analyses that cannot be performed by relevantly trained research analysts. An optimistic critical thinking leads to the reality that no mountain on planet earth could not be climbed by human species. Indeed, as one of the old sayings suggests, "the powers of the brain are so great that it can conquer even the elements". Researchers' technical cognitions should be continuously developed from novice to professional expertise in order to cope with the fundamental analytical requirements.

Fourthly, the use of questionnaires can be prone to dishonesty in responses supplied especially in the physical presences of the researchers or under a certain context. In some cases, respondents may worry about the way the surveys can be used, their privacy and security issues (Evans & Mathur, 2015). Specifically, when responses are made in such a context and with respondents familiar to the researcher evaluating a certain situation of which interests of both parties are covertly maintained, there are higher chances of skipping truths and jumping to responses which favour such interests. If respondents take an obligatory position of supplying data, may give false and duplicated entries in case of electronic surveys (Nayak & Narayan, 2019). In fact, with impersonal online surveys there is a challenge of probing and eliciting sufficiently right responses from right respondents (Evans & Mathur, 2005).

Also, surveys are most misused and particularly the questionnaire surveys popularly used in too small-scale researches to make generalizations (Young, 2016). This is due to the perceived easiness of using the methodologies especially with novice researchers. However,

it should be comprehended that surveys only provide estimates for the true population but not exact measurements (Glasow, 2005).

Moreover, other problems are low response and return rates hampering the applicability of the methodology, and over-claims for what data from a small sample can really generalize to the whole population in question (Lupu & Michelitch, 2018; Nayak & Narayan, 2019; Young, 2016). In some circumstances, some questions in self-completion questionnaires are always left unanswered. This seems to be the lesser effect burden than the contexts when questionnaires are not returned. The survey researchers face the problem of optimizing responses with missing systems in responses and especially when response, completion and return rates are too low to argue and infer on the cases studied (Nayak & Narayan, 2019). Researchers have to create motivating instruments and favorable environment for the respondents to participate to their maximum. This can increase chances of utilizing data from reasonably and technically compliant samples.

Summary of Problems of using survey studies

- i. May exhibit mismatches between paradigms, research questions and instruments
 - ii. Prone to instrumentation systematic errors
 - iii. Dishonesty responses
 - iv. Difficulty use of statistical computations
 - v. Challenged by appropriateness of the study scope
 - vi. Problems of low response and return rates
-

CONCLUSION

Questionnaire surveys are significantly constituting researches that form the wider part of the positivist paradigm than any other. Its detrimental part for the concerns of methodological evolution of other paradigms in research has been based on critique of rigid nature of instruments which has given rise to consideration of the alternative tools. If a faulty instrument is used, errors are taken consistently throughout the study. Also, researchers using the method have the greater chances of running through processes with misaligned paradigms, research questions and instruments, especially if not technically experienced and careful. In some cases, technical limitations face researchers especially where advanced statistical computations are required to infer to the hypotheses. However, the method has benefits which drive for its popularity. If problems and challenges inherent in its methodological, design, instrumentation and implementation contexts are well addresses, it becomes quick to write a research which parsimoniously give valid, reliable and generalizable results.

References

- Aldhein, M. (2020). Interview versus questionnaire from the perspective of CBE members. *International Journal of Education, Learning and Development*, 8(2): 21-41. Retrieved on November 17, 2020 from <https://www.eajournals.org/wp-content/uploads/Interview-versus-Questionnaire-from-the-Perspective-of-CBE-Members.pdf>
- Ali, Z. and Bhasker, S. B. (2016). Basic statistical tools in research and data analysis. *Indian Journal of Anaesthesia*, 60: 662-669. DOI: 10.4103/0019-5049.190623
- Bihu, R. (2020). Using unstructured interviews in educational and social science research: The process, opportunity and difficulty. *Global Scientific Journals*, 8 (10): 712-721. EOI: 10.11216/gsj.2020.10.44855
- Bird, D. K. (2009). The use of questionnaires for acquiring information on public perception of natural hazards and risk mitigation – a review of current knowledge and practice. *Nat. Hazards Earth Syst. Sci.*, 9: 1307–1325. Retrieved on November 17, 2020 from file:///C:/Users/Reuben%20Bihu/Desktop/nhess-9-1307-2009.pdf
- Bryman, A. (2008). *Social research methods, 4th Edition*. Oxford University Press. Retrieved on November 18, 2020 from file:///C:/Users/Reuben%20Bihu/Desktop/SocialResearchMethods.pdf
- Bryman, A., Becker, S. and Sempik, J. (2008). Quality criteria for quantitative, qualitative and mixed methods research: A view from social policy. *International Journal of Social Research Methodology*, 11(4): 261-276, DOI: 10.1080/13645570701401644
- Burgess, T. F. (2001). *A general introduction to the design of questionnaires for survey research*. University of Leeds. Retrieved on November 17, 2020 from <http://www.each.usp.br/rvicente/questionnaireLeeds.pdf>
- Davies, J. (2006). Semantic web technologies. In J. Davies, R. Studer & Warren, P. (Eds.), *Trends and Research in Ontology-based Systems*. West Sussex: John Wiley & Sons, Inc.
- Delice, A. (2010). The sampling issues in quantitative research. *Educational Sciences: Theory & Practice*, 10(4): 2001-2018. Retrieved 19 December 19, 2020 from <https://files.eric.ed.gov/fulltext/EJ919871.pdf>
- Dorney, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. Lawrence Erlbaum Associates, Publishers. Retrieved on November 18, 2020 from http://www.drromartinez.com/uploads/4/4/8/2/44820161/dorney_questionnaires_2003.pdf
- Evans, J. R. and Mathur, A. (2015). The value of online surveys. *Internet Research*, 15(2):195-219. DOI: 10.1108/10662240510590360
- Gabble, G. G. (1994). Integrating case study and survey research methods: an example in information systems. *European Journal of Information Systems* 3(2):pp. 112-126. Retrieved on November 16, 2020 from <https://eprints.qut.edu.au/5853/1/5853.pdf>

- Glasow, P. A. (2005). *Fundamentals of survey research methodology*. MITRE Washington C3 Center. Retrieved on November 18, 2020 from https://www.mitre.org/sites/default/files/pdf/05_0638.pdf
- Gray, D. E. (2004). *Doing research in the real world*. London, Thousand Oaks, New Delhi: SAGE Publications
- Ha et al. (2015). Use of survey research in top mass communication journals 2001–2010 and the total survey error paradigm. *The Review of Communication*, 15(1):39–59. Retrieved on November 16, 2020 from https://www.winona.edu/communicationstudies/media/2015a_stana_use-of-survey-research-in-top-mass-comm-journals-2001-2010.pdf
- Hasson, F., Keeney, S. and McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4): 1008-1015. Retrieved on November 17, 2020 from file:///C:/Users/Reuben%20Bihu/Desktop/DelphiGuidelines.pdf
- Jupp, V. (2006). *The Sage dictionary of social research methods*. Sage Publications. Retrieved on November 18, 2020 from <http://text-translator.com/wp-content/filesfa/Dic-of-Social-Research.pdf>
- Kothari, C. R. (2004). *Research methodology: Methods and techniques, second edition*. New Age International Publishers. Retrieved on November 17, 2020 from <https://www.modares.ac.ir/uploads/Agr.Oth.Lib.17.pdf>
- Marczyk, G., DeMatteo, D. and Festinger, D. (2005). *Essentials of research design and methodology*. Hoboken: John Wiley & Sons, Inc.
- Matthews, N. L. (2018). Measurements, levels of. Retrieved on November 17, 2020 from <https://www.researchgate.net/publication/328368748>
- Nayak, S. D. P. and Narayan, K. A. (2019). Strengths and Weakness of Online Surveys. *IOSR Journal of Humanities and Social Science*, 24(5): 31-38. DOI: 10.9790/0837-2405053138
- Ponto, J. (2015). Understanding and Evaluating Survey Research. *Journal of Advanced Prac.Oncol.*, 6(2): 168-171. Retrieved on November 15, 2020 from https://pdfs.semanticscholar.org/9a74/8722741fb64ba17422e8d253b6dce99322d7.pdf?_ga=2.5223603.1216141735.1605724382-417724120.1605724382
- Rattray, J. and Jones, M. C. (2005). Essential elements of questionnaire design and development. *Journal of Clinical Nursing*, 16: 234–243. DOI: 10.1111/j.1365-2702.2006.01573.x
- Regmi, et al. (2016). Guide to the design and application of online questionnaire surveys. *Nepal Journal of Epidemiology*, 6(4): 640-644. Retrieved on November 16, 2020 from file:///C:/Users/Reuben%20Bihu/Desktop/Guidetodesignandapplicationofonlinesurvey.pdf
- Sharma, S. K., Mudgal, S. K., Thakur, K. and Gaur, R. (2020). How to calculate sample size for observational and experimental nursing research studies? *National Journal of*

Physiology, Pharmacy and Pharmacology, 10(1): 1-8. DOI: 10.5455/njppp.2020.10.0930717102019

Singh, A. S. and Masuku, M. B. (2014). Sampling techniques & determination of sample size in applied statistics research: an overview. *International Journal of Economics, Commerce and Management*, 2(11): 1-22. Retrieved on November 18, 2020 from <http://ijecm.co.uk/wp-content/uploads/2014/11/21131.pdf>

Muniswamaiah , Manoj , Tilak Agerwala , and Charles C. Tappert . "Approximate query processing for big data in heterogeneous databases ." *2020 IEEE International Conference on Big Data (Big Data)*. IEEE, 2020.