Application of Science and Technology to Millennium Challenge via Appropriate Methodology

Ajibade, A.D [1]; Babarinde, S.N [2]; Odusina, M.T[1]

1. *Department of Mathematics and Statistics, The Polytechnic Ibadan, Eruwa Campus, Oyo State, Nigeria.*
2. *Department of Business Administration and Management, The Polytechnic Ibadan, Eruwa Campus, Oyo State, Nigeria*

Abstract: *Writing a standard scientific researchable paper be it project, technical report, dissertation or journal that reveal the application of science and technology to millennium challenges seems to be an herculean task among some scholars and students. Based on the above assertion, this paper critically examined the core aspect of a standard paper that is, methodology. Methodology is the framework and live-wire of any scientific paper. Hence, this paper takes a step to x- ray the components of methodology such as research design, study population, operational variables, sampling technique, and sample size determination, method of date collection and method of data processing/analysis with a view to harnessing the components together in their application for science and technology to millennium challenges.*
Key words: *research design, study population, sample size, data collection, variable(s) and methodology.*

I.   Introduction

In life, challenges vary from one category to another one. The greatest challenge facing developing countries such as Nigeria today is insecurity which is emanated from high rate of unemployment of the educated youths of the country. Other challenges are: food scarcity, unstable power supply, infant mortality rate, spreading of communicable diseases and falling standard of education and to tackle these challenges calls for appropriate scientific and technological approach or technique.

One of the methods of tackling any challenge/problem is to carry out research to know the causes of such problem and try to provide likely solution(s) to such problem via the publication of the report of such research for present and future needs.

When carrying out a research work, one must know the appropriate methods and theories applicable to such research work. In many cases, the application of wrong method can put the theory at risk and account for the special conditions or the use of additional theories which may enlarge the scope of the study, increase the number of research questions and suddenly turn the research work to a strenuous one. However, it is a herculean task to produce an acceptable scientific paper.

In many papers published today, many authors adopted methods without taking cognizance of theory that is relevant to their research topics. They also neglect some of the components of methodology. However, this paper therefore explores the components of methodology in writing the standard paper/journal that is capable of showing or presenting some of the possible solutions to any millennium challenge.

What is methodology? Methodology can be considered to be the theory of correct scientific decisions (Karfman as cited in Mouton & Marais 1996:16). According to Henning (2004:36), it can be defined as coherent group of methods that complement one another and that have the ability to fit to deliver data and findings that will reflect the research questions and suit the research purpose. But, Holloway (2005:293), described methodology as a framework of theories and principles on which methods and procedures are based.

Literally, there are many types of research, these include: Historical, Comparative, Descriptive, Correlation, Experimental, Evaluation, Action, Feminist/identity, politics and cultural. The above listed types of research can be categorized under two groups in terms of research design, namely: qualitative design and quantitative design.

II.   Choice Of Research Design

The choice of research design is based on the type of research work that a researcher intends to embark on. Basically we have two types of design: qualitative design and quantitative design.

Qualitative Design

According to: De Vos (2002:360), Qualitative design is dialectic and interpretive in nature, because during the interaction between the researcher and the research participants, the participants' responses are

observed and interpreted by means of qualitative method. This design is appropriate when a researcher is conducting qualitative research, and this type of research is inductive, holistic and subjective

whish require process-orientated methods to understand, interpret, describe and develop a theory on a phenomena. Also, qualitative design is mostly associated with words, language experiences, perspectives, opinions rather than measurements and statistics. It involves systematic collection and analysis by subjective narrative data in an organized and intuitive fashion to identify the characteristics and the significance of human experience (Holloway 2005:47-51).

Quantitative Design

According to Creswell J.W. (2003), in quantitative methodology, the reality is objective and singular, because the relationship between researcher and researched is independent. Also, the process of research is deductive because the generalizations leading to predictions, explanation and understanding, is valid and reliable design, because it accommodates the use of statistical hypothesis testing.

In quantitative design, a researcher comfort with the methodological assumptions of the quantitative data. The technical writing, computer statistical skills (data mining) and library skill are allowed. Besides, a researcher chooses quantitative design if the research work embarked on has been done by other researchers. This implies that a body of literature exists, is known along with the variables and existing theories. However, the completion time of such research work is always shorter than a research with qualitative methods.

STUDY POPULATION

At an early stage in the planning of any investigation, decisions must be made concerning the study population. That is, concerning the population of individual units (whether they are persons, households, animals, materials etc.) to be investigated. The population under consideration should be clearly and explicitly defined in terms of place, time, and other relevant criteria. For example, if the study population comprises cases of a disease, the procedures to be used for case identification should be stated and if the controls are to be considered their method of selection should be stated.

The selection of study population on the basis of suitability usually affects the validity of subsequent generalizations from the findings. This situation requires a close attention at the early stage of the given study. For example, If one is conducting a research on the national insecurity of a particular country which is one of the global and millennium challenges, at the initial stage of the research work one must clearly describe the affected population or area so as to know the variables involved and design an effective method or approach to combat the problem.

OPERATIONAL VARIABLES

A variable is a characteristic of a person, object, or phenomenon that can take on different values. A simple example of a variable is a person's age. This variable can take on different values, such as, 20 years old, 30 years old, and so on. Other examples of variables are: weight in kilograms, height in centimeters, and monthly income of a firm, marital status (single, married, divorced and widowed).

It is important to make a distinction between **dependent and independent** variables. Both the dependent and independent variables together with their operational definitions (when necessary) should be stated.

The variable that is used to describe or measure the problem under study is called the dependent variable. The variables that are used to describe or measure the factors that are assumed to influence (or cause) the problems are called independent variables. For example, in a study of relationship between smoking and lung cancer, "suffering from lung cancer" (with the values yes, no) would be the **dependent** variable and "smoking" (with the values no, less than a packet per day, 1 to 2 packets per day, more than 2 packets per day) would be the **independent** variable. Other variables are:

Background variables - In almost every study involving human subjects, background variables, such as, age, sex, educational status, monthly family income, marital status and religion will be included. These background variables are often related to a number of independent variables, so that they influence the problem indirectly. Hence they are called background variables or background characteristics.

Confounding variable - A variable that is associated with the problem and with a possible cause of the problem is a potential confounding variable. This type of variable may either strengthen or weaken the apparent relationship between the problem and a possible cause.

Composite variable - A variable based on two or more other variables may be termed a composite variable. Incidence and prevalence rates, sex ratios, (rates and ratios are composite variables). However, the variables to be studied are selected on the basis of their relevance to the objectives of the investigation.

SAMPLING TECHNIQUES

Sampling involves the selection of a number of study units from a defined study population. A time the population is too large for one to be considered. Instead, we select a sample of individuals hoping that the sample is representative of the population.

When taking a sample, we will be confronted with the following questions:
a)   What is the group of people from which we want to draw a sample?
b)   How many people do we need in our sample?
c)   How will these people be selected?

Sampling methods

An important issue influencing the choice of the most appropriate sampling method is whether a sampling frame is available, that is, a listing of all the units that compose the study population. Basically, there are two types of sampling methods: probability sampling technique (such as simple random sampling, systematic sampling, stratified sampling, cluster sampling, multistage sampling) and non-probability sampling (e.g. convenience sampling, quota sampling, purposeful sampling and accidental sampling). The methods highlighted above are not discussed in this paper but an appropriate and effective method should be adopted by any intending researcher so as to have representative samples from study population.

SAMPLE SIZE DETERMINATION

In planning any investigation, we must decide how many people, materials, items and subjects need to be studied in order to answer the study objectives. If the study population is too small we may fail to detect important effects, or may estimate effects too imprecisely. Also, if the study population is too large then we will waste resources.

In general, it is much better to increase the accuracy of data collection than to increase the sample size. The eventual sample size is usually a compromise between what is desirable and what is feasible.

More so, the feasible sample size is determined by the availability of resources and it is important to remember that resources are not only needed to collect the information, but also to analyze it so as to use the conclusion drawn from selected samples to infer the entire population

In order to calculate the required sample size, you need to know the following facts:
a) The reasonable estimate of the key proportion to be studied. If one
Cannot guess the proportion, 50% of the population should be considered.
b) The size of the population that the sample is to represent: If it is more than 10,000 the precise magnitude is not likely to be very important; but if the population is less than 10,000 then a smaller sample size may be required.
c) A degree of accuracy is required.
 d) The confidence level required, usually specified as 95%.

Estimating a proportion
• Estimate how big the proportion might be (P)
• Choose the margin of error you will allow in the estimate of the proportion (say ± w)
• Choose the level of confidence that the proportion in the whole population is indeed between (p-w) and (p+w). We can never be 100% sure. Do you want to be 95% sure?
• The minimum sample size required, for a very large population (N>10,000) is:
$n = 2 Z p (1-p) / w^2$. For example,
(1). A hospital administrator wishes to know what proportion of discharged patients is unhappy with the care received during hospitalization. If 95% Confidence interval is desired to estimate the proportion within 5%, how large a sample should be drawn?
$n = Z^2 p (1-p)/w^2 = (1.96)^2 (0.5 \times 0.5)/ (0.05)^2 = 384.2 \approx 385$ patients
Note: If you don't have any information about P, take it as 50% and get the maximum value of PQ which is 1/4 (i.e., 25%).

Comparison of two proportions
N (in each region) = $(p_1 q_1 + p_2 q_2) (f(\alpha,\beta)) / ((p_1 - p_2)^2$

α = type I error (level of significance)
β = type II error (1-β = power of the study)
Power = the probability of getting a significant result
F (α, β) =10.5, when the power = 90% and the level of significance = 5%
For example,
(2) The proportion of patients severing from cancer is compared between two cohorts. In one cohort, 30% of the patients are estimated within 3 years of prevalence. In other cohort it is probably 15%.
The required sample to show, with 90 % likelihood (power), that the percentage of patients is different in these two cohorts would be: (assume a confidence level of 95%) n = (1.28+1.96)2 ((.3×.7) + (.15 ×.85)) / (.30 - .15)2 = **158.**

III.   Method Of Data Collection

Having decided on how to design the research study, the next methodological design is how to collect information. The most commonly used methods of collecting information (quantitative data) are the use of documentary sources, interviews and self-administered questionnaires.

The choice of methods of data collection
        The choice of method of data collection based on the followings:
• The accuracy of information they will yield
• Practical considerations, such as, the need for personnel, time, equipment and other facilities, in relation to what is available. Accuracy and *"practicability"* are often inversely correlated. A method providing more satisfactory information will often be a more expensive or inconvenient one. Therefore, accuracy must be balanced against practical considerations (resources and other practical limitations).
        In statistics, there are many methods of data collection but few are stated in this paper: personal interview, direct observation, questionnaire method, experimental and registration method. However, the effective and applicable method should be adopted in writing a report of any scientific and technological research work.

DATA PROCESSING AND ANALYSIS
Data processing and analysis should start in the field, with checking for completeness of the data and performing quality control checks, while sorting the data by instrument used and by group of informants. Data of small samples may even be processed and analyzed as soon as it is collected.

Why is it necessary to prepare a plan for processing and analysis of data?
The plan for data processing and analysis must be made after careful consideration of the objectives of the study as well as of the tools developed to meet the objectives. The procedures for the analysis of data collected through qualitative and quantitative techniques are quite different.

• For **quantitative data** the starting point in analysis is usually a description of the data *for each variable* for all the study units included in the sample. Processing of data may take place during data collection or when all data has been collected. Description and analysis are usually carried out *after* the fieldwork has been completed.

• For **qualitative data** it is more a matter of describing, summarizing and interpreting the data obtained *for each study unit* (or for each group of study units). Here the researcher starts analyzing *while* collecting the data so that questions that remains unanswered (or new questions which come up) can be addressed before data collection is over.
Preparation of a plan for data processing and analysis will provide you with better insight into the feasibility of the analysis to be performed as well as the resources that are required. It also provides an *important review of the appropriateness of the data collection tools* for collecting the data you need. That is why you have to plan for data analysis *before* the pretest.
When you process and analyze the data you collected during the pre-test you will spot gaps and overlaps which require changes in the data collection tools before it is too late.

What should the plan include?
When making a plan for data processing and analysis the following procedures should be considered:
 a) Sorting of data
 b) Processing of data
 c) Data analysis

Sorting data

The reason for sorting the data is to facilitate the subsequent processing and analysis of data. For example, if we have different study populations (such as village health workers, village health committees and the general population), the questionnaires used to obtain the data must be *separated*.

In a comparative study it is best to sort the data right after collection into the two or three groups that you will be comparing during the data analysis. In a study concerning the reasons for low acceptance of family Planning services, for instance, users and non-users would be the basic categories. In a case control Study, obviously the cases are to be compared with the controls and it is useful to number the questionnaires belonging to each of these categories *separately* right after the separation. For instance, the questionnaires administered to users of family planning services could be numbered U1, U2, U3, etc., and those for the non-users N1, N2, N3, etc.

Data processing

Decide whether to process or analyze the data from questionnaires: **Manually**, using data master sheets or manual compilation of the questionnaires. **By computer**, for example, using a micro-computer and existing software or self written programmes for data analysis.

Data processing in both cases involves:
• categorizing the data,
• coding, and
• summarizing the data in data master sheets, manual compilation without master sheets, or data entry and verification by computer.

1. Categorizing

Decisions have to be made concerning how to categorize responses.

For **categorical variables** that are investigated through closed questions or observation, the categories have been decided upon beforehand.

In interviews the answers to open-ended questions (for example**,** 'why do you visit the health centre?') can be pre-categorized to a certain extent, depending on the knowledge of possible answers that may be given. However, there should always be a category called 'Others, specify . . .', which can only be categorized afterwards.

These responses should be listed and placed in categories that are a logical continuation of the categories you already have. Answers that are difficult or impossible to categorize may be put in a separate residual category called 'others', but this category should not contain more than 5% of the answers obtained.

For **numerical variables,** the data are often better collected without any pre- categorization. If you do not exactly know the range and the dispersion of the different values of these variables when you collect your sample (e.g., home-clinic distance for out-patients, or income), decisions concerning how to categorize and code the data at the time you develop your tools may be premature. If you notice during data analysis that your categories had been wrongly chosen you cannot reclassify the data anymore.

2. Coding

If the data will be entered in a computer for subsequent processing and analysis, it is essential to develop a CODING SYSTEM.

For computer analysis, each category of a variable can be coded with a letter, group of letters or word, or be given a number. For example, the answer 'yes' may be coded as 'Y' or 1; 'no' as 'N' or 2 and 'no response' or 'unknown' as 'U' or 9. The codes should be entered on the questionnaires (or checklists) themselves. When finalizing your questionnaire, for each question you should insert a box for the code in the right margin of the page. These boxes should not be used by the interviewer. They are only filled in afterwards during data processing. Take care that you have as many boxes as the number of digits in each code.

Coding conventions

Common responses should have the same code in each question, as this minimizes mistakes by coders.
For example:
Yes (or positive response) code - Y or 1
No (or negative response) code - N or 2
Don't know code - D or 8
No response/unknown code - U or 9

Codes for open-ended questions (in questionnaires) can be done only after examining a sample of (say 20) questionnaires. You may group similar types of responses into single categories, so as to limit their number to at most 6 or 7. If there are too many categories it is difficult to analyze the data.

3. Summarizing the data in data master sheets, manual compilation, or compilation by computer
(a) Data master sheets
If data are processed by hand, it is often most efficient to summarize the raw research data in a so-called **DATA MASTER SHEET**, to facilitate data analysis. On a data master sheet all the answers of individual respondents are entered by hand.

(b) Compilation by hand (without using master sheets)
When the sample is small (say less than 30) and the collected data is limited, it might be more efficient to do the compilation manually.
Certain procedures will help one ensure accuracy and speed, the procedures are:
(1) If only one person is doing the compilation, **manual sorting** is appropriate. But if 2 persons are working together, tally method is applicable.
(2) To do **manual sorting** the basic procedure is to:
- Take one question at a time, for example, 'use of health facility',
- Sort the questionnaires into different piles representing the various responses to the question, e.g., hospital/ health centre/ traditional practitioners) and Count the number in each pile. When you need to sort out subjects who have a certain combination of variables (e.g. females who used each type of health facility) sort the questionnaires into piles according to the first question (gender), then subdivide the piles according to the response to the other question (use of health facility).
(3) To do **tally counting** the basic procedure are:
a) One member of the compiling team reads out the information while the other records it in the form of a tally (e.g., ///representing 3 subjects, ////representing 4. subjects who present a particular answer).
b) Tally count for no more than two variables at one time (e.g., sex plus type of facility used). If it is necessary to obtain information on 3 variables (e.g., sex by time of attending a health facility by diagnosis), do a manual sorting for the first question, then tally count for the other two variables.
c) After tally counting, add the tallies and record the number of subjects in each group.
(4) After doing either manual or tally counting, one must **check** the total number of
Subjects/responses in each question to make sure that there are no omissions or double counts.

(c) Computer compilation
Before you decide to use a computer, you have to be sure that it will save time or that the quality of the analysis will benefit from it. Note that feeding data into a computer costs time and money. The computer should not be used if your sample is small and the data is mainly generated by open questions (qualitative data), unless there is a resource person who is competent in using a program for qualitative data analysis. The larger the sample, the more beneficial in general the use of a computer will be.
Computer compilation consists of the following steps:
1. Choosing an appropriate computer program
2. Data entry
3. Verification or validation of the data
4. Programming (if necessary)
5. Computer outputs/prints

Choosing an appropriate computer package
The identification of an appropriate statistical package is the first step in using a computer for data analysis.
Examples of computer packages commonly used are: SPSS, STATA, MINITAB, TORA, EVIEW, R-statistics and SQL etc.

Data entry
To enter data into the computer you have to develop a data entry format, depending on the program you are using. After deciding on a data entry format, the information on the data collection instrument will have to be coded (e.g., Male: M or 1, Female: F or 2). During data entry, the information relating to each subject in the study is keyed into the computer in the form of the relevant code.

Verification validation of the data

During data entry, mistakes will definitely creep in. The computer can print out the data exactly as it has been entered, so the printout can be checked visually for obvious errors, (e.g., exceptionally long or short lines, blanks that should not be there, alphabetic codes where numbers are expected and wrong codes)

Programming

A certain amount of basic knowledge of computer programming is needed to give the appropriate commands during the analysis of data.

Computer outputs

The computer can do most of the analysis and the results can be printed. It is important to decide whether each of the tables, graphs, charts and statistical tests that can be produced makes sense and should be used in your report. That is why one needs to plan or design the method of data analysis at initial stage of research work.

IV. Conclusion

No meaningful development strategy for combating the millennium challenges can be effective without development of adequate techniques in science and technology. Based on the above statement, there is no doubt that in writing a paper or journal that capable of providing solutions to any of the millennium challenges, the concepts of the methodology (key aspect of paper or journal) should not be overlooked.

References

[1]. Creswell, J.W (2003). Research Design: Sage publications, London
[2]. Miles, M.B et al (1974). Quantitative Data Analysis 2nd edition: Sage publications, London
[3]. Altman, D.G (1991). Practical Statistics for Medical Research: Chapman and Hall, London.
[4]. Colton, T. (1974). Statistics in Medicine: Boston Little, Brown and Company 9INC.
[5]. Corlien, M.V; Indra Pathmanathan and Ann Brownlee (2003). Designing and conducting health systems research projects: Volume 1 Proposal development and fieldwork: KIT/IDRC
[6]. Degu, G and Tessema, F. Biostatistics for Health Science Students: lecture note series: Department of Community Health, Jimma Institute of Health Sciences.
[7]. Development in primary health care (1998). Ethical considerations in research focus ESTC-EPHA/CDC PROJECT (2004), Training modules on health research.
[8]. Fletcher, M (1992), Principles and Practice of Epidemiology. Department of Community Health, Faculty of medicine, Addis Ababa University.
[9]. Manktelow, B; Hewitt, M.J and Spiers, N (2002). An introduction to Practical Statistics Using SPSS: Trent Focus, Gondar (1995). Manual for field training Jimma (1996). Manual for student research project
[10]. Mathers, Nigel; Howe, Amanda and Hunn, Amanda. Trent focus for research and The Carter Center 9EPHTI), Addis Ababa.