Data Collection for Multivariate Statistical Analysis: The Implication of Moderators on Predictor and Outcome Variables' Correlation in Nigerian Communication Research

MURTADA, Busair Ahmad, *Ph.D.*
International Institute of Journalism, Abuja, Nigeria
(Ilorin Satellite Campus)

Abstract

To investigate communication issues in modern Nigeria, investigators have to go beyond conceptual mentioning of the key factors in such issues and do empirical search about the factors. Even though every serious research takes off with conceptual clarification of issues to be researched, it should be noted that such intellectual initiative can only provide the groundwork for data gathering. Unless issues are well-conceptualised right from the beginning and before data collection, there is always the tendency that pertinent concepts would be missed in the data. This paper discusses the significance of factoring in, in our research design, the likely mediators or moderators along with the predictors and the dependent variable to be studied. In other words, the paper discusses why the necessary concepts or variables have to be deliberately included in the research design of an empirical study and before data gathering as well as discusses reasons for elaborating the association between an independent and dependent variables. The paper also explores the literature of quantitative research methods to discuss the relevance of specific third variable as elaborators of two variables' correlation within the ambit of multivariate statistical methods and relevant theoretical or conceptual frameworks.

Keywords: Communication, Research, Research design, Data gathering, Multivariate analysis, Theoretical framework, Moderators, Predictor.

Introduction

When an issue calls for empirical inquiry, there is always the tendency on the part of the investigator, mostly at the initial stage, to study the issue from the perspective of relationship between two variables. This could be said to be normal at the starting point, especially when the two key variables are the ones noticeable at the heart of the issue being investigated. However, not most investigators, specifically the non-

empiricists, see relationship between an independent variable and a dependent variable as only a starting point. Likewise, the beginners in the field of social research would like to conceive relationship between the two variables as an end, not a means to an end. The point is that a serious and intensive investigation should elaborate on the initial relationship that has been established between the two variables. The need to elaborate the bivariate relationship is a significant one because there is always the tendency that there might be a third variable or a set of third variables facilitating such relationship. Most often the significant relationship between a predictor and a dependent variable is a moderated one because of statistical interaction of a third variable with the independent variable (Jaccard, Turrisi & Wan, 1990, p.8). Therefore, establishing relationship between two variables should be considered only as the beginning of data analysis; it should not be accepted as a final outcome upon which conclusion and recommendations should be based in a study (De Vause, 2002, p.297).

Mostly neglected in empirical inquiry into communication issues published in some emerging journals in Nigeria are not only explanations about why and how two variables correlate, but also explanations about patterns of the correlations. Put differently, the mechanisms and processes by which a variable predicts an outcome are mostly ignored in what we published in the new journals and book chapters. The import of this observation is that findings of an empirical research, which are aimed at getting to the root of significant issues and events, should go beyond figures displayed in frequency tables and cross tabulation. Such findings are expected to also go beyond testing a set of null hypotheses with the use of Chi Square and similar inferential statistics. This is because inferential statistics do not unravel the role of third factors in the whole gamut of an equation.

This conceptual discourse is necessitated by the fact that discourses on quantitative research are very scarce in the Nigerian communication research journals. Likewise, data collection and data analysis for multivariate purposes are infrequently discussed in earlier studies published in the Nigerian media/communication journals. This situation could be attributed to the fact that most of the journals and book chapters are thematically concerned about societal, cultural and political issues and phenomena that call for contributions from communication and media scholarship perspectives. Even though various statistical techniques are consistently used to analyse data in those research works, it should be noted that the works are not focused on how and why the techniques should be used for analysis nor discuss how multivariate statistical techniques work.

Existing research works on quantitative and qualitative data analysis by Nigerian scholars such as Amadi (2011) and Amadi (2014) only discuss the bias of Nigerian communication researchers for quantitative methods at the expense of

qualitative methods in their research works though, they did not mention that those scholars made quantitative methods and statistical techniques the central themes of their discourses published in the local journals. This paper fills the gap, exploring data analysis literature in quantitative research with emphasis on how relevant data on communication related issues should be collected as well as how to maximize such data for elaboration for the benefit of mass communication research community and policy makers. The general objective of this paper is to discuss not only the reason for inclusion of all pertinent concepts in the research design of our investigation because such design would determine the information to look out for during data collection, but also the reason for elaborating the initial correlation of an independent variable and a dependent variable. The specific objectives are to:

1. Explain why it is necessary to ensure that pertinent and significant variables are captured in the research design of the communication issue to be studied before data collection
2. Identify statistical techniques for evaluating indicators
3. Explain the purpose of elaborating relationship between two variables with the aid of a third factor or a set of third factors
4. Identify the benefits of multivariate analysis of communication research data

Clarifying Concepts at Research Design Stage
Before embarking on data collection, it is necessary to break the issue to be studied into specific concepts that capture its different aspects. Definition of a concept would depend largely on what we mean by each of the concepts within the context of our investigation. Bearing in mind that concepts do not have set meanings, we would have to be very clear about what we mean by a concept and how to define it. We would better play safe, if we only use the words that are commonly understood and contextually make sense. Clarification of concepts takes different approaches notwithstanding, it is easily simplified by going through three basic steps: obtaining a range of definitions for each of the concepts; deciding on a definition; and delineating the dimensions of each concept (De Vaus, 2002, p.44-46).

It would be somewhat unreasonable to pick a definition for a concept without giving consideration to how the concept has been defined by others. Relevant sources have to be read in search for applicable definitions out of which a choice could be made. Dictionaries, encyclopaedias, web pages, textbooks and journals are distinctive sources of definitions. From a range of definitions, an investigator can make selection of a more appropriate definition. At times, there might be the need to identify the common elements of two or more definitions, which could be extracted for creating a newer and apposite definition. Most importantly, the decision to opt for a precise definition or for creating one for the purpose of our study should be logically

justified and aligned with the theory that drives discussion in the investigation. By clarifying the key concepts through assigning definitions to them, we would be making them nominal and functional. They can thereby provide focus for our research as well as guide us on the pertinent information to gather.

It is also important to know that most concepts have different anchor points or dimensions; therefore clarification of concepts goes beyond their definition. Conceptual clarification includes determining their dimensions so that further focus could be attained for our research and so that we can precisely measure what we actually intend to measure. Put differently, identifying and isolating the actual anchor point or dimension from a range of anchor points of a concept provides the ground for valid measure. A valid measure is the "one which measures what it is intended to measure" (De Vause, 2002, p.53). An investigator is expected to identify a specific dimension of each of the concepts for easy generation of indicators or items for every concept (Austin & Sutton, 2015). Most often, not all dimensions of a concept should be studied.

Locating dimensions of a concept requires exploring relevant literature in which a specific concept is extensively discussed from different viewpoints. If we are dealing with two or more concepts, it is germane to do literature review on each concept with focus on its diverse dimensions. Reviewing literature for Mass Communication research criss-crosses different disciplines such as sociology, psychology, economics, political science, and of course, literary studies. The hybrid nature of communication research and theories demands reading of existing research works in the sister disciplines in order to discover sundry dimensions for each of the pertinent concepts in focus.

Some concepts might be too broad and abstractive so much that the researcher might not be able to come up with a unique definition. In such a situation the needful would be to locate two initial dimensions of a concept sub-dimensions and, of course further steps of sub-dimensions which can provide concrete meaning of the concept or a set of indicators of the concepts. Until such process is sufficiently accomplished, the concept in question would remain latent. Concepts or latent variables cannot be directly measured because they cannot be directly observed (Bryne, 2010, p.4). A typical example can be seen in how we could conceptually define 'TV audience perception of Nigerian Shiite Muslims', a supposedly dependent variable in a proposed research work. The question raised here is on how we define the concept **viewers' perception**, considering different dimensions or angles to the concept. In clarifying the concept, we have to first and foremost descent from the vague meaning to a concrete meaning. We have to do this in order to come by indicators or items to tap the concepts.

In an attempt to delineate 'viewers' perception', it could be traced to

perceptions of social reality which were essentially described as divergent because of perceivers' dissimilar communication experiences of events (McLeod & Chafee, cited in Eveland & Glynn, 2008). Therefore, viewers' perception can be explored at two levels of abstraction, one leading to the other. At the individual level of abstraction, it refers to the view of the world held cognitively by individuals. At the societal level, it implies a perception of the world commonly shared in the society (McLeod & Chafee, cited in Eveland & Glynn, 2008). Because viewers' perception in our hypothetical research has to do with Nigerian Shiite Muslims, it is a contextual one. Thus, it can be linked to social reality perceptions which are synonymous with the construct stereotypes in that stereotypes could be the result of external stimuli dialectic and the cognitive processes that occur at intra-individual level (Leyens, Yzerbyt & Schadron, 1994, pp.10-13). Viewers' perceptions of social reality has been conceptualised and measured in different cultivation-related studies from various dimensions (Gross, Morgan, Nancy & Signorelli, 1986; Gerbner & Gross, 2002, p.229; Griffin, 2009, p.353; Morgan and Shanahan, 2010).

Viewers' perception may be differentiated in terms of scope, anchor point, and intensity. The scope of *viewers*' perceptions implies the pervasiveness and generality of the perceptions held. For example, viewers may negatively perceive an entire region, an entire ethnic group, or certain subgroups or collectives within a social order. Viewers' perceptions may also be differentiated based on the anchor point of perception. Viewers may anchor their perceptions in the observable behaviour of a target group like violent or non-violent protest conduct of Nigerian Shiite Muslims or in rather less noticeable mannerisms like tolerance, empathy, insensitivity, and sociability. A third dimension may be characterised by the strength or intensity with which the perceptions may be held. Intensity implies the orientation of the audience members towards the target object of the perceptions. For instance, some may like to share interaction with the target; others may like to share the same schools, same workplaces, or same neighbourhood with the target of the perception. Therefore perceptions play a twofold filter role: one of commonality of outlook and one of actions (Leyens, Yzerbyt & Schadron, 1994, p.10).

Emanating from the myriad of definitions of perception sourced from the literature reviewed which has helped us to delineate the concept and considering the purpose of our assumed investigation, we can go on to say that *viewers*' perception of Nigerian Shiite Muslims is conceptually defined as shared beliefs of Nigerian television viewers about the Nigerian Shiite Muslims as a group of people. Perception is treated here as essentially an attribute of the audience rather than an attribute of the target entity.

Developing Indicators for a Concept
Concepts are only measurable and, of course, researchable only when they are made

to descend from the ladder of abstraction. A clearer picture of the concept 'viewers' perception' has emerged through the movement of the concept from its generic meanings to the specific ones. Until the concept's dimension was made to descend from upper level of abstraction to the possible lower and observable level, its meaning stood vague and lofty.

The question that should be asked here is: why is it necessary to develop indicators for every concept? The answer is that without a set of indicators, it would be impossible to measure the concepts. No matter how clearer and meaningful a concept might be defined, it cannot be measured directly as it would remain, to some extent, a latent variable. It is therefore the duty of the researcher to operationally define the latent variable in observable terms. Put differently, the concept has to be interpreted in terms of behaviour that signifies it to enable direct measurement. Direct measurement of the signifying behaviour includes items or statements, attitudinal scales, and scores on a specific measuring instrument such as a self-administered questionnaire (Bryne, 2010, p.4). The statements in the instrument should be functional so they can tap physical activities or normal tasks that characterised the variable. Most importantly, the indicators should be made to concentrate on and carefully capture the identified dimensions of our concepts or variables of interests so that respondents' answers or responses could be relied upon.

This raises the issue of validity which suggests that we have to ensure that the indicators measure what they are constructed to measure. It also raises the issue of reliability which looks at the consistency in the way respondents respond to items in the instrument at a time and the manner they respond to the same items in another day. A reliable instrument is that which generate the same results on repeated occasions. Whenever indicators are devoid of validity and reliability, there would be measurement error (De Vaus, 2002, p.52). In addition, indicators should be consistent with the theory and research questions of concern.

In some places wherein there are well-established indicators for specific concepts or variables that have been measured consistently by experts, there would not be any need to reinvent new ones. In fact, researchers can take advantage of such existing scales. They can easily be obtained online for use. It is advantageous to use existing scales because they do not only provide the ground for easy tracking of change in the way people respond to same items or question overtime as new findings are compared with the old ones, but they also facilitate uniformity of concepts' definition and coding classification (De Vaus, 2002, p.50). However, modifications can be made in the wording of the existing indicators for local relevance.

If no existing measurement scale is relevant to your study, it becomes necessary to create your own indicators. How do we develop indicators? First and foremost, it should be known that it would be unwise to write statements or items

from our own thoughts. That is what some of us do when developing data gathering instruments. It is rather useful to obtain information from people who are part of the study population through unstructured interviews or through focus group discussion (Austin & Sutton, 2015). Such approach would enable us to have insight and firsthand knowledge about the thinking of the people regarding the subject matter we are studying. It would provide us with the guide to the right worded questions, statements or items. Focus Group Discussion, a qualitative method, is mostly useful for understanding an issue at a deeper level as well as for adding meaning and understanding to existing knowledge of why and how of the issue (Academy Editors' Picks Resources, 2017).

In the case of our hypothetical study, we can decide to develop indicators for 'TV audience perception of Nigerian Shiite Muslims', the supposedly dependent variable, by gathering data from a focus group discussion. The participants or discussants should comprise TV viewers who live in the northern part of Nigeria or who do not live in the north but have concern for religious events, issues and, of course, conflicts in north of Nigeria. We open the floor for a brainstorming session and exploratory discussion about activities of the Nigerian Shiite Muslims, their behaviour, and how people view them in the society. We would ensure that all concepts for our proposed study, one after the other, form the basis for brainstorming. Definitely, we would have background knowledge about each and every concept discussed, a part of which is our dependent variable. We would also know about how an average TV viewer is likely to perceive the Islamic sect through the lens of a group of people who are part of our study population. Information obtained would be useful for developing indicators to measure or help to operationalise the variable. Similar method applies to measurement of other variables in the study. See Table 1 for a typical set of indicators or indexes that were derived from a focus group discussion to purposefully measure TV audience perception of Nigerian Shiite Muslims.

Table 1 Indicators used to measure TV audience perception of Nigerian Shiite Muslims

1.	Shiite Muslims behave violently.
2.	Shiite Muslims are impatient.
3.	Shiite Muslims are the ones who initiate trouble.
4.	Shiite Muslim religious scholars are a threat to the society
5.	Shiite Muslims are opposed to modern values.
6.	Shiite Muslims want political power at all cost.
7.	Shiite Muslims behave in uncivilised ways.
8.	Shiite Muslims do not socialise with others.
9.	Shiite Muslims are less educated.
10.	Shiite Muslims look down upon other Muslims.
11.	Shiite Muslims are less accommodating.
12.	Shiite Muslims practice mischief.
13.	Shiite Muslims are unforgiving of others.
14.	Shiite Muslims are fatalistic.
15.	Shiite Muslims are fundamentalists.
16.	Shiite Muslims are not tolerant.
17.	Shiite Muslims are reactionaries.
18.	Shiite Muslims do not intermingle with other Muslims and non-Muslims.
19.	Shiite Muslims do not sympathise with other Muslims and non-Muslims in misfortunes.

Table 1 is provided here to showcase the indicators or statements used as a part of the whole process of the composite measure which include the Likert-scale or a Likert-scale type. Respondents would typically be asked, in a self-report fashion, to index their perceptions of Nigerian Shiite Muslims by indicating their degree of agreement or disagreement on a ten-point Likert-type scale ranging from 1 very strongly disagree to 10 very strongly agree.

Statistical Techniques for Evaluating Indicators
It is safe to develop as many indicators to measure a variable from differing dimensions. Definitely, an unstructured interview and/or focus group discussion would, at the end of the day, provide us as many as thirty or more indicators for one concept. However, most often all the indicators for measuring the concept or variable might not necessarily be used. It is up to us to decide which dimension of the concept to take. Such dimension would drive our choice of a set of indicators to use for measuring the concept (Austin & Sutton, 2015).

There are two mostly used mechanisms to assess reliability of indicators in a scale. They are the test-retest reliability and internal consistency. The test-retest reliability of a scale is a statistics that is calculated whenever we administer a questionnaire to the same people on two different instances. We thereafter calculate the correlation between the two scores obtained and see whether there is high test-retest correlation which normally indicates a more reliable scale. With the invention of the SPSS, there has been less use of test-retest. Researchers therefore make use of

the SPSS by which the scale reliability can be assessed in terms of its internal consistency (Pallant, 2007, p.6). Internal consistency of a scale refers to the extent to which the items, indicators or statements that are included in the scale hang together to measure the underlying attribute. Internal consistency is mostly measured by the use of a statistics called Cronbach's coefficient alpha which is available within the *reliability analysis* unit in the *scale* component on the SPSS. Cronbach's coefficient alpha tells us of the average correlation among all of the items that are put in the scale (Pallant, 2007, p.6). Its values ranges from 0 to 1, the higher the value the greater the reliability of the scale. Nunally recommends a Cronbach's alpha value of 0.7 as the minimum level a scale should attain for consistency and reliability (Nunally cited in Pallant, 2002, p.6). It should be noted that a scale with small number of indicators most often generate a small Cronbach's alpha value. That is why it is advised that researchers generate as many indicators for each variable during focus group discussion or during interview with a group of people who are part of a research population.

Use of Multivariate Data Analysis Techniques
Multivariate statistics are used for elaborating relationship between two variables. Elaboration analysis normally commences with focus on causal relationship between an independent and dependent variables. An inquisitive researcher would not be satisfied with a correlational outcome of two variables no matter how logical such outcome might look. The researcher would want to go on investigating in order to detect the reason behind the correlation. Most often, the quantitative empiricists dig for *other factors* that intervene between the predictor and the dependent variable. They would also like to know the kind of intervention the other factors make. They would like to specifically know whether the variables are providing a moderating type of intervention or a mediating type. At times, they might just like to know the most powerful independent variable among fewer others in terms of predicting an outcome in a given situation.

In our hypothetical study, it could be found that the dependent variable (TV *viewers*' perception of Nigerian Shiite Muslims) was predicted by two independent variables at two different levels of significance. Specifically, the dependent variable could be found to have been predicted by the first independent variable (viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims) at a level of significance lower than the level of significance at which it was predicted by the second independent variable (viewers' real life experience with Nigerian Shiite Muslims' conflict with others)**.**

That is why it is significant to include relevant concepts at the research design stage and before data collection. Otherwise, it would be impossible to create a non-

existing variable at the data analysis stage because there would be lack of components to create it in the data that have been collected. Data that are full of relevant conceptual ingredients would always provide the ground for multivariate or multiple explanations of why and how an event takes place or why and how an issue takes effect (Pallant, 2007, pp.4-5). The why and how explanations are statistically given by interaction effects which are normally created by formation of an interaction variable from two existing continuous independent variables. In essence, one of the independent variables is deemed to moderate the relationship between the other independent variable and the dependent variable (Jaccard, Turrisi & Wan, 1990, pp.8-9). This designation can be typically explained with our   hypothetical study. Let say 'monthly income level of TV viewers' is the moderator variable (i.e. the variable that moderates the relationship between viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims [the independent variable] and TV viewers' perception of Nigerian Shiite Muslims [the dependent variable]. This interaction effect implies that the relationship between the independent variable and the dependent variable depends on the income levels of individual viewers which are differing.

There are different techniques for multivariate analysis. The choice of a multivariate technique for data analysis in a study would depend on the type of research questions or objectives that guide the study, the nature of data gathered for the study (Pallant, 2007, p.100), and ultimately the theoretical framework and reasoning in which the study is situated. Even though the standardised and sophisticated techniques help evaluate and clarify our models, they cannot be substituted for theoretical sophistication (de Vause, 2002, p.319).

Basic knowledge and understanding of diverse statistical techniques, what they stand for, the type of research questions they handle and their underlying assumptions and principles are very essential for empirical investigators. Therefore, use of standardised and sophisticated data analysis techniques is not about mere decision to impress readers of our research works.

Correlation and Correlation Coefficient

What is correlation? Correlation means relationship between two variables and it is calculated by correlation coefficients. **Correlation coefficient** is a numerical summary that tells us about the type and strength of a relationship between variables. A correlation coefficient is numerically indicated thus: $r_{ab} = +/- x$

While *r* stands for the coefficient, a and b stand for the variables that are set for correlation. The plus (+) and minus (-) respectively indicate the positive and negative directions of the relationship between two variables. Definitely a relationship goes in one direction only. Unlike the minus sign that always appears for negative

relationship, the plus sign does not appear. Therefore a coefficient with no sign indicates positive correlation. There are different correlation coefficients that are calculated for detecting the type and strength of relationship between variables.

The choice of the coefficient to use in a given situation largely depends on how the two variables are measured. For ratio and interval-scale generated variable, Pearson product moment correlation is calculated. Pearson product moment correlation is also called product-moment correlation or Pearson's *r*. To calculate coefficient for relationship between ordinal variables, spearman correlation coefficient (rho) is used. The procedure for examining correlation between two nominal variables is based on chi-square. For effective understanding of coefficients and what they stand for, interested researchers should read books that teach basic statistics.

While correlation tells us whether two variables are related, it does not indicate how changes in one variable produce changes in another variable. Correlation is not causation. Correlation means two variables co-vary. Such relationship does not mean one variable causes the other, but because a third variable instigates the two variables to relate. Put in different words, for every correlation between two variables, there is always a third explanation or a set of multiple explanations. Frey, Botan and Kreps (2000) illustrated a good example in a Chinese fable about the extent to which people in ancient times recognised the danger of confusing correlation with causation. The fable reads:

> While hunting for prey, a tiger caught a fox. The fox quickly thought and said, "You can't eat me! The Emperor of Heaven appointed me the king of the beast. If you eat me, you'll be disobeying his orders. If you do don't believe me, follow me. You 'll soon see whether the other animals run away at the sight of me or not." Agreeing to this, the tiger accompanied him. Every beast who saw them coming dashed away. Not realizing they were afraid of him, the tiger believed they were afraid of the fox (Frey, Botan & Kreps, 2000, p.365).

That is it. The tiger saw the correlation between the beasts' sighting the fox and the beasts running away. The tiger did not know 'he' was the third variable to which the relationship between the two variables should be attributed! Therefore, it would be investigative suicidal for a researcher to infer causation from correlation. In the case of our hypothetical investigation, we should be cautious about attributing whatever type of perception TV *viewers*' might share about Nigerian Shiite Muslims to the independent variable (viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims) at a zero order level. There are other variables that should be 'partialled out' or controlled for. Otherwise, the factors behind such initial correlation would remain unknown.

Multivariate Data Analysis is Correlation-Focused

Mass communication researchers cannot do without multiple data analysis and multiple correlations because of the nature of communication process virtually in every society. Thus, they investigate the complexity of a communication issue or event by examining relationship among more than two variables in order to discover, at least, one or two of the reasons one thing predicts another. Because communication processes are anchored on multiple concepts for investigation, it behoves an inquisitive communication researcher to build in the research design the relevant concepts and subsequently go for a set of data that embraces all the concepts in the design. Such must be ensured in preparation for correlational analysis.

Multiple correlation means relationship between two or more variables and it is calculated by a multiple correlation coefficient. A multiple correlation coefficient indicates a set of independent variables interrelate to the dependent variable in focus. A multiple correlation coefficient shows the direction and the strength of the relationship between a set of independent variables and the outcome variable. A multiple correlation coefficient is numerically indicated thus:

$R_{abc} = +/-x$

Where R stands for the coefficient, a, b and c stand for the variables that are computed for correlation. Further, clearer explanation of multiple correlations is indicated by Coefficient of Multiple Determination. A coefficient of multiple determination depicts the amount of variance explained in the dependent measure by a combination of the independent variables. A multiple determination coefficient is calculated by squaring the multiple correlation coefficients.

Going back to our hypothetical investigation, let us assume that we found the dependent variable (TV *viewers*' perception of Nigerian Shiite Muslims) to have correlated with a combination of the two predictors (viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims and viewers' real life experience with Nigerian Shiite Muslims' conflict with others), $R_{abc} = .65^2 = .42$. The finding would suggest that 42% of the variance in the dependent variable was explained by the two independent variables as an entity. Therefore, 58% of the variance in the dependent variable was left unexplained by the two independent measures and we would say that .58 is the coefficient of non-determination. The coefficient of non-determination is that part of the variance in the criterion variable that is not explained by the two predictors.

The limitation of multiple correlation lies in the fact that it can only tell us about relationship among variables, it cannot tell us about relationship between two variables, when the effects of other variables are controlled for or removed. **Partial**

correlation explains the relationship between two variables, when influence of another variable or effects of a set of other variables are statistically disconnected. Partial order correlation is indicated numerically thus: $r_{abc} = +/-x$

The *r* stands for the coefficient, a and b stand for the two variables whose relationship is being examined, and c stands for the control variable. When we control for one variable, the partial correlation is called first-order partial correlation. When we control for two or more variables, the partial correlation is called higher-order partial correlation (distinctively, when we control for two variables and three variables, we can say we are calculation second-order partial correlation and third-order partial correlation etc). When we control for no variable while examining relationship between two variables, we would be having a zero-order correlation. It is a rule of thumb that we do not compute for a first-order partial correlation when we cannot established that two variable significantly correlate at the zero-order level. Partial correlation, regardless of type, is all about the effects of influence of a third or Z variable in a bivariate relationship.

In our hypothetical investigation, we tried to know whether there was relationship between an independent variable (viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims) and a dependent construct (viewers' perception of Nigerian Shiite Muslims). Our initial interest was about relationship between the two variables. So, we went on to calculate a coefficient of determination which indicated a positive correlation of .40 for two variables (at the significant level of .02). What this suggests is that 40% of the variance in the dependent variable was explained by the independent variable at zero order level. We went further in our investigation to know what might have accounted for the bivariate relationship. The quest for other plausible factors was instigated by the fact that those factors might help explain something of the zero order relationship. We therefore recomputed the two variables along with another variable (viewers' real life experience with Nigerian Shiite Muslims' conflict with others) to be controlled for, using the literature and theory that guided our path of reasoning.

In a nutshell, we applied first order partial correlation for elaborating the relationship and we found that the two variables still correlated although at a lower level of significance (.05). This result suggests that the relationship between viewers' exposure to TV news reports on violent activities of Nigerian Shiite Muslims and viewers' perception of Nigerian Shiite Muslims was connected to viewers' real life experience with Nigerian Shiite Muslims' conflict with others. Otherwise the level of significance at which the two variables correlated would have remained unchanged.

Both zero order correlation and partial order correlation are prerequisites to multiple correlation analysis. What goes for multiple regression, ANOVA, path

analysis and similar others are expected to have been examined at both zero order correlation and partial order correlation analysis level.

Regression and Similar Multivariate Statistical Techniques

Quantitative researchers most often go beyond establishing relationship and patterns of relationship among variables. Most of the time, quantitative investigators would like to predict and explain respondents' score on the criterion or dependent variable by knowing their score on the predictor or independent variable. The investigators therefore go for regression analysis. The lowest of the regression family is called linear regression. Linear regression 'is used to predict or explain scores on a criterion variable on the basis of the obtained scores on a predictor variable and knowledge of relationship between the two variables' (Frey, Botan & Kreps, 2000, p.370). The highest cadre of regression family is multiple linear regressions. A multiple linear regression allows researchers to predict or explain scores on a criterion variable on the basis of the obtained scores on two or more predictor variables and knowledge of relationship among all the variables (Frey, Botan & Kreps, 2000, p.371). Within the fold of multiple linear regressions are hierarchical multiple regression and stepwise regression. A researcher choice of a regression technique to use would depend on what he intends to do on the basis of the concerned theory and conceptual arguments in the relevant literature.

Some of the merits of regression lie in the fact that it allows researchers to control for variables, identify the strength of predictors relative to their relationship with the outcome measures through the 'beta' or standardised coefficient, and create interaction variables.

Closely related to regression is ANOVA for multivariate analysis. ANOVA shares some characteristics with regression, but it lacks the ability to control for a third or a set of third variables. Its application to data analysis hinges on theoretical explanations.

More complex than regression are path analysis and factor analysis. Path analysis investigates hypothesised relationship among variables for the purpose of establishing their causal connections by displaying the paths the causal influences take (Frey, Botan & Kreps, 2000, p.374). Path analysis requires that the researcher formulate a model for his analysis, using a pictorial flow graph which serves as a means of result presentation (De Vause, 2002, p.330).

Conclusion

Data gathering and data analysis are serious phenomena in communication research and should always be preceded by a research. The essence of research design is to guide on what to search for during data collection. It looks somehow outlandish and

reductive to use quantitative data for only descriptive analysis or for frequency tabulation. Data collection should be based on a well-determined research design that speaks of relevant concepts that are formulated on the basis of the issue or event to be studied, the research problem and, of course, the theoretical framework. The concepts must be well-defined and interconnections ought to be structured of how each concept taps a specific angle of the issue or event to be investigated. Definitely, every concept in our research has a mission to deliver in the research and in the overall data analysis. Once concepts or constructs or variables are made readily available in the data collection instrument, they guide the process of data gathering and provide a requisite for data analysis.

Data analysis tends to be informative and lucid whenever it depicts how variables interconnect to explain the situation of things in the social order that has been investigated. The correct picture of how things are and how they work in the locale of the research of concern cannot be given by bivariate analysis of variables or by mere descriptive portrait of variables and their sub-categories in a cross tabulation. An in-depth analysis would always be required with the use of multivariate statistical methods that have the capability not only to inform, elaborate and elucidate about what predict what and at the instance of what, but also tells the role played by each mechanism or factor among several mechanisms or factors in the equation.

Even though numerical values ought to have been given in this conceptual paper, they are avoided in order to make the explanation simple and friendly. The paper does not say it all about data gathering and its rudiments. Yet, it elaborates the basic things to be done to have relevant data and meaningful analysis. It is recommended that empirical research, specifically in the field of communication, in this part of the world should be enriched with multivariate explanation of the issues investigated. Editors of our journals should promote this tradition of quantitative investigation and writing.

References

Academy Editors' Picks Resources. (2017). *How to conduct a sussceful focus group discussion*. Retrieved 24th of April, 2018 from https://blog, socialcops.com> resources

Amadi, F. (2011). Broadening mass communication research for enhanced media practice. *Global Media Practice*, 5(1), 81-96. Retrieved 24th of April, 2018 from https://globalmedia.journal.ac.za

Amadi, F. (2014). Trends in mass communication research in Nigeria: a Guerilla warfare dissertation. *Asia Pacific Journal of Multidisciplinary Research*, 2(5), 78-86. Retrieved 24th of April, 2018 from www.apjmr.com

Austin, Z. & Sutton, J. (2015). Qualitative research: Data collection, analysis and management. *The Canadian Journal of Hospital Pharmacy, 68*(3), 226-231.

Bryne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, application, and programming*. 2nd Edition, New York: Routledge.

De Vause, D. (2002). *Survey in social research*. 5th edition, Australia: Rutledge.

Gerbner, G., Gross, L., Morgan, M. & Signorielli, N. (1986). *Television's mean world: Violence Profile No. 14-15.* Pennsylvania: The Annenberg School of Communication, University of Pennsylvania.

Gerbner, G., & Gross, L. (2002). Living with television: The Violence Profile. In M. Morgan (ed.), Against *the Mainstream*: *The Selected Works of George Gerbner* (pp.227-253). Hillsdale, NJ: Lawrence Erlbaum Associates. (Reprinted from *Journal of Communication*).

Grifin, E. (2009). *A first look at communication theory*. New York: McGraw-Hill Education.

Frey, L.R., Botan, C.H, & Kreps, G.L.(2000). *Investigating communication: An introduction to research methods*, 2nd Edition, Boston: Allyn and Bacon.

Jaccard, J., Turrisi, R., & Wan, C. K. (1990). *Interaction effects in multiple regression*. London: Sage Publications.

Morgan, M., & Shanahan, J. (2010). A state of cultivation. *Journal of Broadcasting and Electronic Media, 54*(2), 337-355.

Pallant, J. (2007). *SPSS survival manual*. 3rd edition, Australia: Allen and Unwin.