

RELATIONSHIPS AND HYPOTHESES IN SOCIAL SCIENCE RESEARCH

Samithamby Senthilnathan

PhD (Bus./Fin.), MSc (Mgmt), BSc (Bus. Admin.)
Director of Studies/Quality Assurance,
International Training Institute, Papua New Guinea
Email: sennathans@yahoo.com

Abstract

*This paper highlights the variables and their relationships in a social science framework, since most of the social science studies focus on investigating relations and causal impacts between/among the variables. As a main focus, this paper illustrates identification of variables in direct causal relationship, indirect causal relationship (refers to **moderating variable**), and direct and indirect causal relationship (refers to **mediating variable**) with their relative identical linear models and hypotheses to be investigated. Further, this paper emphasizes that a social science study should consider alternative hypothesis as the direction of its investigation.*

Keywords : models, variables, moderating, mediating, hypothesis

JEL code : C12, C18, C20, C30

RELATIONSHIPS AND HYPOTHESES IN SOCIAL SCIENCE RESEARCH

Samithamby Senthilnathan

1. INTRODUCTION

A social science research can be defined in simple terms as a scientific study about the action and interaction by human that explores on the human perceptions, thoughts, understandings and behaviours with social orientation and integration. Therefore, the social science studies can identify many inter-related variables and mostly investigates such variables one over the other for their inter-actions and influences. In this context, most of the social science investigations are exemplified to explore the relationships between and among the social variables. It is identical that ***no investigation is made on two social variables, if they are not related to each other; and if a social variable is investigated independently, its relativity to another variable would be at least considered to explore its importance and significance in the social context.***

Considering the relationships of social variables, it is important to explore what type of relationships and their relative hypotheses can be formulated in a social science investigation. The main exploration of this paper is to illustrate the type of relationship of social variables and why a ***Social Science Study*** must mainly consider alternative hypothesis, rather than focusing on null hypothesis.

The rest of this paper is organised as uniqueness and relationship of social variables, hypotheses in a social science study, and a concluding remarks.

2. INVESTIGATIONS ON SOCIAL SCIENCE VARIABLES

A social science investigation mostly contributes to socially existing phenomena that adds knowledge through filling the gap in the literature. Kline (2008) demonstrates that there are three different directions of investigations that social scientists investigate. They are namely: ***Descriptive, Relational*** and ***Causal***.

2.1 Descriptive Investigations

Descriptive studies attempt to answer mostly the fundamental questions related to the social variables. These questions can be derived due to repeated occurrences of a set of observations relatively in human actions, and because of that, social scientists can pay

attention on those occurrences and attempt to explore the quality and characteristic of related individual variables on their repeated occurrences in the social context.

As the study variables, based on measurement, can be categorised as nominal, ordinal, interval and ratio, the descriptive studies can focus on these descriptive measures.¹ For a nominal variable, appropriate descriptive measure to explore the variable is the Mode; for ordinal variable, the Median of the variable becomes appropriate; and similarly, the Mean (together with related other measures like standard deviation, variance, coefficient of variance) becomes important to spell the characteristic and uniqueness of the variable.²

2.2 Relational Investigations

As most of the social variables are results of other events, social scientists mostly attempt to investigate relativity of social variables. Social science studies are mostly interested and investigate this type of investigations in general. If social scientists observe a series of occurrences and/or existing theoretical background in application with respect to two or more related variables, the studies can be directed to investigate in a different social context on the same relational variables.

As the relational investigations focuses on relationship at least between two variables, the investigation can focus on whether the study variables have relationship or association to each other. Generally, in a simple context, an associative relationship of the study variables can be explored with Chi-squared and/or Correlation Analyses.³ A relational investigation aims at establishing the relationship between study variables in a social context, irrespective independence and dependency of the variables concerned.

¹ Nominal variable can only be categorised and cannot be ranked and measured into an order. Ordinal variables can be ranked into a preferential order, and cannot be measured. Interval and Ratio variables can be both ranked and exactly measured (Bhattacharjee, 2012).

² Mean, Median and Mode are descriptive statistical measures that explore the individuality, characteristic and uniqueness of a variable. The mode represents highest number of occurrences (useful to represent Nominal Variable) in a set of observations; median represents the middle observation (useful to represent Ordinal Variable) of the orderly arranged set of observations; and mean represents average (useful to represent the Interval/Ratio Variable) of the observed measurements (Bhattacharjee, 2012).

³ Chi-squared Test is mostly carried out mostly for the nominal (categorical) variables to justify the extent to which the associations of the study variables exist (Moore, Notz & Flinger, 2013). Similarly, Correlation Analysis is called bivariate analysis, which demonstrates how two measurable variables have association to each other. The correlation analysis always implies what is the direction and size of association between the variables, and cannot mean for how such relationship is built in (Sharma, 2005).

2.3 Causal Investigations

Causal investigations are mostly in place that consider how a dependent variable can be explained and possibly predicted/estimated relatively to its contributing variable(s). These contributing variables to have the occurrences of dependent variable are known as its constituents. These causal investigations should be based on the relationship variables identified in the above relational studies. If the relational study fails to signify a relationship between two variables, it is hardly possible to devise the causal investigation on those variables. Hence, to a most certain extent, the causal investigation between independent and dependent variables only becomes possible after establishing the relational state of the variables as indicated above.

The impacts of the independent variable(s) on the dependent variable are measured and explored to establish a fact in relation to the study (independent and dependent) variables. The most common way of exploring the impacts and influences of independent variable(s) on a dependent variable is through carrying out regression (and multivariate) analysis.⁴ The results of the analysis provides evidence for whether the independent variables significantly important to explain and/predict the dependent variable.

All in all, irrespective of the above type of investigations, a study should consider descriptive measures of the study variables as an important analysis to explore the uniqueness and characteristics of the study variables in a theoretical background.⁵ Depends on the requirement of the investigation, the study can be carried on appropriate analyses for establishing or demonstrating a social fact whether the study variables have significant relationships or causal relativity, one over the other. It is important to understand that ***all of the analyses should be directed to answer the main research question of the study***. In this context, the hypothesis is important to provide directions of conceptualization and analyses of the study variables.

⁴ Regression Analysis explores how a set of independent variables can have impact on a dependent variable, i.e., how these independent variables transformed into the dependent variable (Sharma, 2005).

⁵ The term "***Theory***" simply means "***empirically established and logically interrelated facts/proposition***". According to Kerlinger (1986, p.9), a "***Theory***" is: "***a set of interrelated constructs, definitions, and propositions that present a systematic view of phenomena by specifying relations among variables, with the purpose of explaining and predicting phenomena***".

3. HYPOTHESIS OF A SOCIAL SCIENCE STUDY

3.1 Simple Meaning of a Hypothesis in a Social Science Study

As a hypothesis in a social science investigation is meant to be a tentative answer to the research question of the study, the hypothesis becomes the empirical direction and benchmark to justify the results and findings of the study (Bhattacharjee, 2012). The social scientists require extensive knowledge of study area to highlight a research problem in a social context. However, they seek evidences from other previous investigations to formulate the research problem and then the relative hypothesis as the tentative answer to the research question. All along, it is important to understand that a research question cannot be just investigated without single evidence at least. In accordance with the research question, social scientists often formulate a social phenomenon hypothesis to design, direct and conclude their investigations, where the hypothesis provides a base for setting up other requirements of the investigation (Labaw, 1980).⁶

3.2 Types of research in Social Sciences

Basically, social science studies can be identifiable into two: (a) Inductive investigation and (b) Deductive investigation. Inductive studies are to direct the investigations to establish a theory from the observations and patterns of social occurrences; and the deductive studies are to investigate whether the established theory (a result of an inductive research) has existence in the present context of the society. In a process, inductive investigation begins with **observations**, identifying the **pattern** of observed outcomes, positioning the **ideal outcome** (preposition/hypothesis) and establishing the **fact** or **theory**; and deductive investigation as a result of an inductive study begins from the established **theory**, emphasizing related **hypotheses**, **observing** the occurrences in social science context, and **confirming** the theory established.⁷

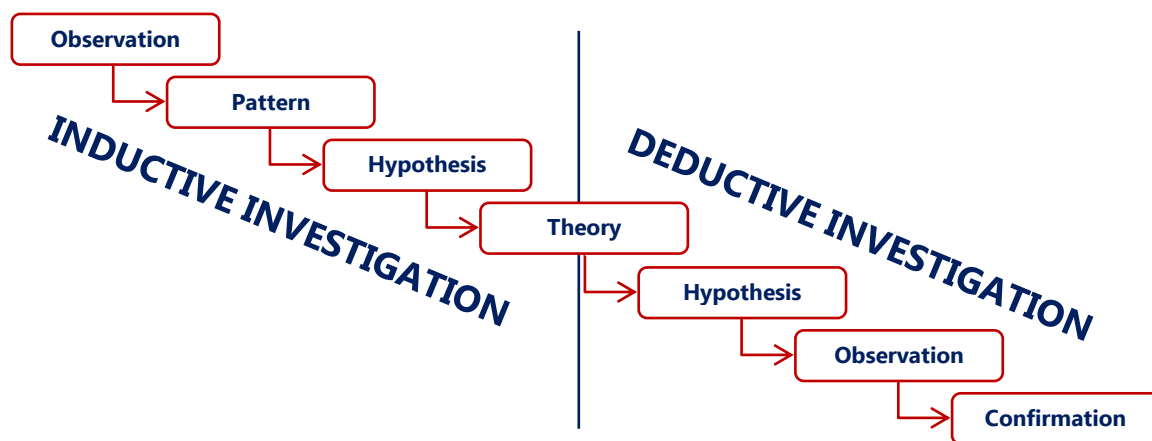
⁶ Labaw (1980, p. 35) indicates “Developing the hypothesis provides the key ingredient to structure all subsequent parts of the project: the questionnaire, the sample, the coding, the tabulation forms, and the final report itself”.

⁷ Though some studies alternatively use the terms prepositions and hypotheses, Bhattacharjee (2012, p.13) identifies the difference between a preposition and a hypothesis as indicating “Because propositions are associations between abstract constructs, they cannot be tested directly. Instead, they are tested indirectly by examining the relationship between corresponding measures (variables) of those constructs. The empirical formulation of propositions, stated as relationships between variables, is called hypotheses.”

A social science theory as a finding of inductive investigation can be established through observing social/variable phenomena; and therefore, the inductive investigations provide the bases for deductive investigations. As the deductive studies depend on inductive studies, Figure 1 shows the relationship and distinction between the inductive and deductive investigations.

Nowadays, most of the social studies are deductive studies. They do attempt to investigate and confirm the established social facts/theories in different social, geographical, economic, political, legal and environmental contexts.

Figure 1: Relationship and Distinction between Inductive and Deductive Studies



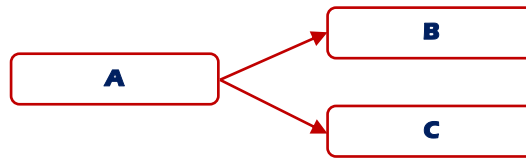
3.3 Relationship Variables

It is general and obvious that social scientists are merely interested in establishing relationships between variables. There is lack of evidence that non-related variables have been investigated in a social science study. As most of the studies are interested in establishing or confirming the relationship between social variables, such relationships can be categorized mainly into three.

3.3.1 Direct Causal Relationship

This relationship is meant for a variable that has a nature/capacity to **directly** influence and modify other variable(s). For example, the variable “A” as an independent variable has a direct impact on variables “B” and “C”, individually (see Figure 2).

Figure 2: Direct Causal Relationship



The significant contribution of the independent variable (A) in predicting the dependent variables (B and C) can be explored with simple regression models, individually, whereas the dependent variables (B and C) have no relationship to each other, but impacted by the independent variable (A).

These relationships in regression analysis can be modeled as:

$$V_B = \beta_0 + \beta_A \cdot V_A + u_i \quad (1)$$

$$V_C = \beta_0 + \beta_A \cdot V_A + u_i \quad (2)$$

Where, V_A = independent variable "A", V_B = dependent variable "B"
 V_C = dependent variable "C" β_0 = intercepting value
 β_A = coefficient estimate of variable "A" in predicting dependent variables "B" and "C", individually.
 u_i = error term

The significant contribution of variable "A" in predicting the variables "B" and "C" can be confirmed with individual respective coefficient estimate of variable "A", significantly.

3.3.2 Indirect Causal Relationship

In this relationship, the contribution of an independent variable (A) towards the dependent variable (C) is moderated and impacted through another variable (B), where the moderating variable (B) has played its role in transforming the contents of independent variable (A) into dependent variable (C). In this context, the contents of independent variable (A) cannot be directly converted into dependent variable (C), if the moderating variable is in place. Rather, the independent variable (A) has first impact on the moderating variable (B), which then impact/affect the dependent variable

(C). Notably, all along, each of the variables has direct significant relationship to each other (see Figure 3).

Figure 3: Indirect Causal Relationship



The significant contribution of the independent variable (A) in predicting the dependent variable (C) can be explored with regression models. In this context, setting four different models to examine such indirect causal relationship is important, as given below.

$$V_B = \beta_0 + \beta_A \cdot V_A + u_i \quad (3)$$

$$V_C = \beta_0 + \beta_A \cdot V_A + u_i \quad (4)$$

$$V_C = \beta_0 + \beta_B \cdot V_B + u_i \quad (5)$$

$$V_C = \beta_0 + \beta_A \cdot V_A + \beta_B \cdot V_B + u_i \quad (6)$$

Where, V_A = independent variable "A", V_B = moderating variable "B"

V_C = dependent variable "C" β_0 = intercepting value

β_A = coefficient estimate of variable "A"

β_B = coefficient estimate of variable "B"

u_i = error term

The significant contribution of variable "A" in predicting the variables "B" and "C" (see equations (3) and (4)) can be confirmed with individual respective coefficient estimate of variable "A", significantly. This confirms that variable "A" has impact on both variables "B" and "C". Also, it is important to confirm the significant relationship of variable "B" to the dependent variable "C" and therefore, devising and confirming a model for it (see equation (5)) is important. Equation (6) is to confirm the moderating role of variable "B", where the both explaining variables (A and B) of dependent variable (C) are accommodated.

The regression results of the coefficient estimates of independent variables (A and B) in equation (6) need to be considered to confirm the moderating role of variable “B”.

If variable “B” act as a moderating variable to transform the information of variable “A” to predict variable “C”, the equation (6) should result in that the coefficient estimate of variable “B” should be significant, while the coefficient estimate of variable “A” is insignificant at the required significant level (normally $p < 0.05$). However, the both variables “A” and “B” should have significant correlative relationship with dependent variable “C”.

3.3.3 Direct and Indirect Causal Relationship

In this relationship, the contribution of an independent variable (A) towards the dependent variable (C) is mediated and impacted through another variable (B), where the mediating variable (B) has played its role in transforming the contents/information of independent variable (A) into dependent variable (C). In this context, the contents of independent variable (A) are converted into dependent variable (C), both directly and through mediating variable “B”. Indicatively, independent variable (A) has impact on both the mediating variable (B) and dependent variable (C), and then the mediating variable (B) again impacts/affects the dependent variable (C). Notably, all along, each of the variables has direct significant relationship to each other (see Figure 4).

Figure 4: Indirect Causal Relationship



The significant contribution of the independent variable (A) in predicting the dependent variable (C), individually and through mediating variable (C), can be explored with regression models. In this context, as indicated above,

setting four different models to examine such direct and indirect causal relationships is important, as given below.

$$V_B = \beta_0 + \beta_A \cdot V_A + u_i \quad (7)$$

$$V_C = \beta_0 + \beta_A \cdot V_A + u_i \quad (8)$$

$$V_C = \beta_0 + \beta_B \cdot V_B + u_i \quad (9)$$

$$V_C = \beta_0 + \beta_A \cdot V_A + \beta_B \cdot V_B + u_i \quad (10)$$

Where, V_A = independent variable "A", V_B = mediating variable "B"

V_C = dependent variable "C" β_0 = intercepting value

β_A = coefficient estimate of variable "A"

β_B = coefficient estimate of variable "B"

u_i = error term

The significant contribution of variable "A" in predicting the variables "B" and "C" (see equations (7) and (8)) can be confirmed with individual respective coefficient estimate of variable "A", significantly. This confirms that variable "A" has impact on both variables "B" and "C". Also, it is important to confirm the significant relationship of variable "B" to the dependent variable "C" and therefore, devising and confirming a model for it (see equation (9)) is important. Equation (10) is to confirm the mediating role of variable "B", where the both explaining variables (A and B) of dependent variable (C) are accommodated.

The regression results of the coefficient estimates of independent variables (A and B) in equation (10) need to be considered to confirm the mediating role of variable "B".

If variable "B" act as a mediating variable to transform some of the information of variable "A" to predict variable "C", equation (10) should result in that the coefficient estimates of both variable "A" and "B" should be significant at a required level (normally $p < 0.05$).

It is important to highlight the importance of mediating variable (B) in a study in two ways.

- a) The extent to which the existing literature support the importance of the mediating variable; and
- b) The extent to which the results of analysis support the above (a) literature context. In this context, the quantum of mediating variable (B) can reveal how much it transforms information considerably from independent variable (A) to predict dependent variable (C).

3.4 Hypothesis and Why Alternative Hypothesis in a Social Science Study?

In general, two types of hypotheses need to be highlighted in a study: (a) ***Null Hypothesis***, and (b) ***Alternative hypothesis***. Though there is no a concrete and identical way of devising hypotheses for an investigation, it is reasonably agreeable that a social science study needs to emphasise more on "***Alternative Hypothesis***". A simple and logical interpretation of ***this proposal*** is that ***a social science study DO NOT attempt to investigate: (1) a factual prove established by a natural science study, and (2) NON-RELATED variables***. Because of this nature, the alternative hypothesis is always referred to as the ***Working or Functional Hypothesis***, which has the direction of investigation in establishing a social science phenomenon. Most of the social science studies investigate on relationship variables. It is identical that if two variables considered have no theoretical background of established relationship in a social science study, no any investigation takes place on those variables. In this context, it is notable that ***a base for a social science investigation is the expected relationship between the variables***. The study apparently attempts to establish the relationship of such variables, consistently with the existing literature. Hence, the study should highlight its alternative hypothesis(es) as the base(s) for its investigation, rather than pointing its relative null hypothesis, which has no a theoretical and/or literature base for the investigation by the study.

To exemplify the ways of formulating the hypotheses in a social science study, now consider, for examples, Figures (3) and (4) for indirect causal relationship, and direct and indirect causal relationship. Referring to the models in the figures, the related hypotheses can be formulated as indicated below.

Hypotheses for Figure (3): Indirect Causal Relationship

The models (regression equations (3) – (6)) of Figure (3) consider three relationships among the variables. Accordingly, with respect to the correlative relationship of variables, the following hypotheses can be formulated.⁸

H3: There is significant relationship between the variables “A” and “B”.

H4: There is significant relationship between the variables “A” and “C”.

H5: There is significant relationship between the variables “B” and “C”.

The hypotheses for the regression analysis should be formulated as follow.⁹

H_{R3}: Independent variable (A) significantly contribute to moderating variable (B)

H_{R4}: Independent variable (A) significantly contribute to dependent variable (C)

H_{R5}: Moderating variable (B) significantly contribute to dependent variable (C)

H_{R6}: Variable (B) significantly acts as a moderator between independent variable (A) and dependent variable (C).

Meaningful contribution of every independent variable in the models (equations (3) – (6)) can be highlighted with the significant coefficient estimate of explaining variable. However, to explore the significance of moderating variable (B’s) role, the coefficient estimate of independent variable (A) in equation (6) should be insignificant. *In case*, the results reveal that the coefficient estimates of both explaining variables (A and B) are significant in equation (6), then the variable “B” act as “Mediator”, not as a moderator.

Hypotheses for Figure (4): Direct and Indirect Causal Relationship

In the similar way, the models (regression equations (7) – (10)) of Figure (4) consider three relationships among the variables. Based on the correlative relationship of variables, the following hypotheses can be formulated.

H7: There is significant relationship between the variables “A” and “B”.

H8: There is significant relationship between the variables “A” and “C”.

H9: There is significant relationship between the variables “B” and “C”.

⁸ Hypotheses can be set with Positive or Negative relationship of the variables considered in the study; and related such hypotheses should have one tail test appropriately.

⁹ In H_{R3}, R indicated hypothesis for a regression model. It is considered for understanding purpose. However, it is not necessary to consider accordingly.

The hypotheses for the regression analysis should be formulated as follow.

HR7: Independent variable (A) significantly contribute to mediating variable (B)

HR8: Independent variable (A) significantly contribute to dependent variable (C)

HR9: Mediating variable (B) significantly contribute to dependent variable (C)

HR10: Variable (B) significantly acts as a mediator between independent variable (A) and dependent variable (C).

To explore the significant mediating variable (B's) role, the coefficient estimates of both explaining variables (A and B) in equation (10) should be significant. *In case*, the results reveal that the coefficient estimate of variable "A" is insignificant in equation (10), then the variable "B" act as "Moderator", not as a mediator.

All along, the opposite of each of the hypothesis is the NULL hypothesis, which is indirectly understood in every testing of the hypothesis.

4. CONCLUDING REMARKS

This paper highlights the variables and their relationships in a social science framework, thus exploring three different directions (*Descriptive, Relational* and *Causal*) of investigations to be conducted in it. However, it is identical that most of the social science studies focus on investigating relations and causal impacts between/among the variables.

With a simple meaning of hypothesis, the types of studies in social sciences (*inductive and deductive investigations*) are explained with their distinction and relationship. Next, as a main focus, this paper illustrates identification of variables in direct causal relationship, indirect causal relationship (refers to *moderating variable*), and direct and indirect causal relationship (refers to *mediating variable*) with their relative identical models and hypotheses to be investigated. Further, this paper argues and emphasizes that a social science study should consider alternative hypothesis as the direction of its investigation, rather than pointing on its relative null hypothesis.

For young and future researchers, this paper gives a base for conceptualising and formulating hypotheses in social science investigation, and identifying the independent, moderating and mediating variables throughout the analyses.

References

Bhattacharjee, A. (2012). *Social Science Research: Principles, Methods, and Practices*, Textbook Collection, University of South Florida, USA.

Kerlinger, F.N. (1986). *Foundations of Behavioral Research*, (3rd ed.). Fort Worth, TX: Holt, Rinehart and Winston

Kline, R. B. (2008). *Becoming a Behavioral Science Researcher: A guide to producing research that matters*. New York, NY: Guilford

Labaw, P. J. (1980). *Advanced Questionnaire Design*. Cambridge, MA: Abt Books.

Moore, D. S., Notz, W. I., & Flinger, M. A. (2013). *The Basic Practice of Statistics*, 6th ed., New York, NY: W. H. Freeman and Company.

Sharma, A. K. (2005). *Text Book of Correlation and Regression*, Discovery, New Delhi.