A Primer for Quantitative Study Design

Chosang Tendhar, PhD
Director of Institutional Effectiveness and Assessment
Hackensack Meridian School of Medicine
February 10, 2022

Introduction

The purpose of this commentary is to briefly introduce processes involved in designing a quantitative study. There is no dearth of books on research methods. A basic google search reveals that are several primers on research design and analysis. This piece is not intended to be-all, end-all of the research design. However, the focus is on some of the basics, as it relates to the conceptualization of the study and designing it. A primer on the use of appropriate statistical analyses based on the research questions will follow next.

Like skills in any professional domains, it takes a rigorous training, extensive coursework, and many years of experiences to conduct high-quality scholarly research. This primer does not replace any of those requirements. However, this should serve as a good guide to those with no or minimal exposure to trainings in research methods. This primer includes definitions of some of the important concepts used in quantitative studies. Most importantly, this primer should enable the readers to pursue additional readings on relevant topics to successfully complete each stage of the research process.

Just as there are a plethora of research design books, various authors define research design variedly, and topics covered, and the emphases placed on each topic varied greatly as well in those books. Similarly, there are some variations in the research process, or the steps involved in conducting research that various authors explicate. I will, however, use the research process displayed in Figure 1 to guide the process of writing this primer.

Figure 1. Eight stages in Research Process

According to the model in Figure 1, there are eight steps in the research process:

i. Identifying the problem
ii. Reviewing the literature
iii. Setting research questions, objectives, and hypotheses
iv. Choosing the study design
v. Deciding on the sample design
vi. Collecting data
vii. Processing and analyzing data
viii. Writing the report

i. Identifying the Problem

Identifying a researchable problem is the most important step in the scholarly research endeavors. Many of the readers of this primer is probably familiar with the saying: "A problem well put is half solved." The role of problem formulation in scientific inquiry as Einstein and Infeld defined:

The formulation of a problem is often more essential than its solutions, which may be merely a matter of mathematical or experimental skill. To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science.

The research problem statement indicates the goal, purpose, or overall direction of the study. Such a statement implies the potential empirical investigation. However, not all problems qualify as research problems. The research problem statements are usually rather general. There are just enough information about the scope and purpose of the study to provide an initial understanding of research.

One of the most obvious aspects of problem formulation is its scope. Broad questions tend to lead to vague answers. The research problems should be workable and solvable with relationships between specific variables clearly established.

Examples of research problem statements:

- The purpose of this study is to examine the relation between teaching style and student academic achievement.

- This study investigates the effect of sex on aggression.

i.a. Sources of Problems

There is probably no single strategy for identifying research problems that works best for all researchers. Some of the common sources that have been found useful are:

Investigator's Interest and Experiences: Research ideas often emanate from the interests and practical experiences of the investigator. Some of the questions and problems teacher encounter on a regular basis may lead to researchable questions. In addition to personal interest, investigators' interests and knowledge about their profession can be the good source of research problems. As an example of investigator's interest, the course director might be

interested in examining various active learning techniques that leads to a high level of student engagement in the class.

Applying Theory: Another important source for research problems is a theory that has implications for educational practice. In scientific research, generation and testing of theories are important for establishing a body of knowledge that will generalize widely. A various definitions of theory exist, but one of the most important characteristics of a scientific theory is that testable hypotheses may be derived from it. A theory can be defined as a set of propositions that explain the relationships among observed phenomena. A theoretical framework should be included in a study, and it should guide the investigation process. For example, if a researcher is interested in examining students' commitment to a major and career in medicine, the Social Cognitive Career Theory (SCCT) model should be employed.

Replication: A previous research can be replicated with relatively minor changes. Progress in building a body of knowledge depends on a series of replications to verify and extend the initial findings. Some of the reasons for conducting replication studies are:

- To check the findings of a major or milestone study
    - Replications can confirm or disconfirm the validity of a study that produces new evidence or reports findings that challenge previous research or theory.
- To check the validity of research findings with different subjects
    - Replications often use the same procedures but change the type of subjects used to see if the original findings hold for different subjects.
- To check trends or change over time
    - Replications can be used effectively to see if initial findings hold over time.
- To check important findings using different methodologies
    - It is possible that a research finding may be unduly influenced by the way a variable is measured.
- To develop more effective and efficient interventions
    - With the recent emphasis on determining the success of educational interventions, designing research that fine-tunes an intervention is important to maximizing student achievement.

Clarification of Contradictory Findings: The contradictory findings on many topics in the literature present good sources for research problems. For example, research on the effect of charter schools is mixed. Some studies indicate that student attending charter schools outperform students in regular schools whereas other studies conclude that it makes no difference in student achievement.

ii. Reviewing the Literature

The past research on the phenomenon under investigation should play a key role in the process of problem formulation and the design of a study. Therefore, it is important to conduct a thorough review of the relevant literature. It is imperative for the researchers to have exposure to the larger body of knowledge. This will serve as a safeguard against attempts to reinvent the wheel. However, as indicated in the previous section, the sources of research problems extend beyond the existing literature. An important benefit, however, of an extensive review of the literature is that it will enable researchers to demonstrate the significance of their studies sharing the piece(s) of the puzzle their studies contribute to the existing knowledge base.

The review of literature should not be restricted to one's own discipline. It is important to break

out of the almost inevitable parochialism that training and specialization engender. Failure to do so will foster tunnel vision. Restricting the review of literature to one's own discipline and subdiscipline will be inimical to cumulative knowledge.

Interdisciplinary studies are more impactful and significantly expands the body of knowledge in one's own discipline. The study of medical education significantly overlaps with educational research in general, and particularly with other professional education, such as engineering education. For example, the study of professional identity formation is an emerging theme within the medical education community. However, there is a rich literature on the professional identity formation as a construct in engineering education. Therefore, the field of medical education stands to gain if we integrate literature from other fields in addressing medical education research questions.

Some of the core purposes of reviewing literature include:

- Refining the research problem
- Establishing the conceptual or theoretical orientation
- Identifying contradictory findings
- Developing research hypotheses
- Learning about new information

iii. Setting Research Questions, Objectives, and Hypotheses

The research problem statement, as stated earlier, provides an overall direction of the research and that it is general in nature. In other words, the problem statements typically lack specificity. However, the researcher presents specific and narrow questions to obtain measurable and observable data on variables to demonstrate that it is researchable. A researchable question is one that can be investigated empirically. An empirical study is one that gathers evidence (data) that is objective, evidence that is based on observation, measurement, or experience that can be replicated by others.

Using one of the examples of general research problem statements listed in section #i, let's pose a researchable question. *What is the effect of sex on aggression?* A question with an appropriate level of specificity accompanied by an adequate level of knowledge of the current state of affairs of the field will enable a researcher to develop hypotheses and identify dependent variables and independent variables.

In this context, let me define hypothesis and variables.

Hypothesis: A hypothesis is a conjectural statement about a relation between two or more variables, or about how the variables in the study are related. It is an educated "guess" or tentative expectation about a correct solution to a problem, possible relationships, or differences. In research, a hypothesis is typically the investigator's prediction or expectation of what the results will show, and such a prediction is made prior to data collection.

For the research question stated above, based on the researcher's personal experiences and knowledge of past research, a researcher might hypothesize that *there is a significant difference between males and females in aggression and that males tend to be more aggressive*.

Variable: A variable is a characteristics (personal aspects such as grade level, age, or income level) or attributes (represents how an individual or individuals in an organization feel, behave,

or think) of an individual or organization that researchers can measure or observe and varies among the individuals or organizations. Variables are the key ideas or concepts about which researchers collect information to address the purpose of their quantitative study. For example, intelligence, achievement, and cognitive style each involves a range of values, which is usually expressed numerically. However, some variables are better described as containing two or more categories, for example, male and female, and cooperative vs. individualized instructions.

There are serval types of variables in educational research. Let me enumerate five of them:

Categorical Variable: It is also referred to as nominal level data. A nominal scale is one in which there are mutually exclusive categories without any ordered implied. A categorical variable is used to assign an object or person to a group (level) defined by having specified characteristics. Some of the examples of categorical variables include sex, race, socioeconomic status (SES), voting preference, and religious preference.

Continuous Variable: A continuous variable reflects an infinite number of values within a given range of scores. That is, they can assume any value along a scale of values such as height (inches), weight (pounds), time (in seconds), or income.

Independent Variable: An independent variable is the presumed cause. In other words, an independent variable is an attribute or characteristic that influences or affects an outcome or dependent variable. It is useful to classify independent variables into manipulated or non-manipulated ones.

Dependent Variable: A dependent variable is the presumed effect. Stated differently, a dependent variable is an attribute or characteristic that is dependent on, or influenced by, the independent variables. This is generally the variable that the researcher wants to measure or predict—the one of most interest.

Confounding Variable: Confounding variables are attributes or characteristics that the researcher does not directly measure, but that may influence the relationship between the independent and dependent variables.

It is essential that researchers are clear in their minds and research questions concerning the relationship between their variables – which ones are doing the influencing (independent) and which one are being affected (dependent). The independent variable is the antecedent (intervention), the dependent variable is the consequence (outcome). Predictions are made from independent variables to dependent variables.

iv. Choosing the Study Design

A research design is a set of procedures that researchers use to collect, analyze, and report their data in a research study. Quantitative research designs are a set of procedures for collecting, analyzing, and reporting numeric data to assess specified variables to answer research questions that call for explanation.

There are three broad classes of designs: experimental, quasi-experimental, and non-experimental. What serves to distinguish among the three classes of designs is the presence or absence of (a) manipulation of independent variables, and (b) randomization.

Experimental or True Experimental Design: An experiment is a study in which at least one variable is manipulated, and units are randomly assigned to the different levels or categories of the manipulated variable(s). In other words, randomization assign individual participants to different conditions of treatment variable.

Because subjects are randomly assigned to treatment and control, it may be assumed that the groups are equivalent, in the probabilistic sense, on all other variables affecting the dependent variable, except for the treatment. True experiments are also called randomized experiments, intervention trials, and randomized control trials. Researchers use experiments when they want to establish probable cause and effect between their independent and dependent variables. In other words, they seek to understand causal relationships among variables of interests.

Quasi Experimental Design: This refers to investigations in which treatments are administered but randomization is absent. In other words, it is an investigation that has all the elements of an experiment, except that subjects are not randomly assigned to groups.

Nonexperimental Design: In nonexperiment, both manipulation and randomization are absent. Nonexperimental research essentially describes participants, traits, and other characteristics without direct or active intervention. Some of the common types of nonexperimental designs are descriptive studies, comparative studies, correlational studies, and survey research.

v. Deciding on the Sample Design

A sample design refers to the methods to be followed in selecting a sample from the population. The purpose of sampling in quantitative studies is to obtain a group of participants who will be representative of a larger group of individuals. The degree of representativeness and the quality of the information obtained are based on the sampling techniques employed. The best procedure for selecting such a sample is to use probability sampling.

Two major types of sampling procedures are probability sampling and non-probably sampling. Qualitative studies use nonprobability sampling exclusively. Participants are selected purposefully in qualitative studies. Types of sampling procedures for qualitative studies include criterion sampling, typical case sampling, extreme case sampling, and snowball sampling.

The most rigorous sampling strategy for quantitative research is probability sampling. Using this technique, researchers select individuals from the population through a random process so that each individual has a known chance (or probability) of being selected. Some of the common examples of this technique are simple random sampling, systematic sampling, and stratified sampling.

vi. Collecting Data

Across all quantitative research designs, the collection of data has common characteristics. Researchers use instruments to measure the variables in the study. An instrument is a tool for recording quantitative data. It contains specific questions and response possibilities that the researcher establishes in advance of the study. Examples of instruments are survey questionnaire, checklists that can be used to observe an individual's behaviors, standardized tests, secondary data analysis, and structured interviews.

vii. Processing and Analyzing Data

In quantitative research, researchers analyze the data using mathematics procedures, called statistics. Many might know how to calculate correlation coefficient or perform a *t*-test, but not all of them might be able to pick an approach that is appropriate to address a particular research question. A statistical technique to be used will be determined by the research question, whether exploring relationships or differences between groups, and the nature of dependent and independent variables, specifically whether they are categorical or continuous variables.

Depending on the sample size, it is important to understand when to use parametric tests or non-parametric tests. Furthermore, it is helpful to have a familiarity with the parametric techniques and their non-parametric techniques counterparts. For example, the counterpart of an independent sample *t*-test is Mann-Whitney U Test. Similarly, counterparts of Paired-Sample *t*-test and One-way Between Groups ANOVA are Wilcoxon Signed Rank Test and Kruskal-Wallis Test, respectively. There will be a separate primer on statistics soon.

vii. Writing Reports

The entire research process culminates in a report that is disseminated to audiences who may benefit from what the researchers learned during the study. Different audiences also evaluate the research process as described in the research report. The overall format for a report of quantitative study follows a predictable pattern: introduction, review of the literature, methods, results, and conclusions; essentially all of the steps involved in conducting research, discussed above.

Concluding Thoughts

A point that I cannot overemphasize is that, if anyone needing research assistance, including data analyses, they should see experienced researchers with quantitative analytic skills during the early stage of their planning process. It is important to ensure that the research questions and hypotheses are testable. Similarly, it is important to ensure that a researcher is using a reliable and validated instrument or develop an instrument that adheres to fundamental principles of scale constructions. Most importantly, keep in mind that fancy statistics and expensive statisticians will not enhance the quality of your paper. But a researchable question, an adequate literature review, sound theoretical framework, reliable and validated instruments, appropriate data for your research questions, and application of appropriate statistical techniques will strengthen the quality of your paper. The findings of such a study will make significant contributions to the current body of knowledge.