**Lecture 4**

## Research Design

Research design is a comprehensive plan for data collection in an empirical research project. It is a “blueprint” for empirical research aimed at answering specific research questions or testing specific hypotheses, and must specify at least three processes: (1) the data collection process, (2) the instrument development process, and (3) the sampling process. The instrument development and sampling processes are described in next two chapters, and the data collection process (which is often loosely called “research design”) is introduced in this chapter and is described in further detail in Chapters 9-12.

 Broadly speaking, data collection methods can be broadly grouped into two categories: positivist and interpretive. **Positivist methods**, such as laboratory experiments and survey research, are aimed at theory (or hypotheses) testing, while interpretive methods, such as action research and ethnography, are aimed at theory building. Positivist methods employ a deductive approach to research, starting with a theory and testing theoretical postulates using empirical data. In contrast, **interpretive methods** employ an inductive approach that starts with data and tries to derive a theory about the phenomenon of interest from the observed data. Often times, these methods are incorrectly equated with quantitative and qualitative research. Quantitative and qualitative methods refers to the type of data being collected (quantitative data involve numeric scores, metrics, and so on, while qualitative data includes interviews, observations, and so forth) and analyzed (i.e., using quantitative techniques such as regression or qualitative techniques such as coding). Positivist research uses predominantly quantitative data, but can also use qualitative data. Interpretive research relies heavily on qualitative data, but can sometimes benefit from including quantitative data as well. Sometimes, joint use of qualitative and quantitative data may help generate unique insight into a complex social phenomenon that are not available from either types of data alone, and hence, **mixed-mode designs** that combine qualitative and quantitative data are often highly desirable.

#### Key Attributes of a Research Design

 The quality of research designs can be defined in terms of four key design attributes: internal validity, external validity, construct validity, and statistical conclusion validity.

 **Internal validity**, also called causality, examines whether the observed change in a dependent variable is indeed caused by a corresponding change in hypothesized independent variable, and not by variables extraneous to the research context. Causality requires three conditions: (1) covariation of cause and effect (i.e., if cause happens, then effect also happens; and if cause does not happen, effect does not happen), (2) temporal precedence: cause must

**35**

precede effect in time, (3) no plausible alternative explanation (or spurious correlation). Certain research designs, such as laboratory experiments, are strong in internal validity by virtue of their ability to manipulate the independent variable (cause) via a treatment and observe the effect (dependent variable) of that treatment after a certain point in time, while controlling for the effects of extraneous variables. Other designs, such as field surveys, are poor in internal validity because of their inability to manipulate the independent variable (cause), and because cause and effect are measured at the same point in time which defeats temporal precedence making it equally likely that the expected effect might have influenced the expected cause rather than the reverse. Although higher in internal validity compared to other methods, laboratory experiments are, by no means, immune to threats of internal validity, and are susceptible to history, testing, instrumentation, regression, and other threats that are discussed later in the chapter on experimental designs. Nonetheless, different research designs vary considerably in their respective level of internal validity.

 **External validity** or generalizability refers to whether the observed associations can be generalized from the sample to the population (population validity), or to other people, organizations, contexts, or time (ecological validity). For instance, can results drawn from a sample of financial firms in the United States be generalized to the population of financial firms (population validity) or to other firms within the United States (ecological validity)? Survey research, where data is sourced from a wide variety of individuals, firms, or other units of analysis, tends to have broader generalizability than laboratory experiments where artificially contrived treatments and strong control over extraneous variables render the findings less generalizable to real-life settings where treatments and extraneous variables cannot be controlled. The variation in internal and external validity for a wide range of research designs are shown in Figure 5.1.

*Internal validity*

*External*

*validity*

Single lab

experiment

Multiple lab

experiment

Single

case study

Multiple

case study

Cross

-

sectional

field survey

Longitudinal

field survey

Field

experiment

Math

proofs

*Validity*

*frontier*

Simulation

Ethnography

*Cone of Validity*

Figure 5.1. Internal and external validity

Some researchers claim that there is a tradeoff between internal and external validity:

higher external validity can come only at the cost of internal validity and vice-versa. But this is not always the case. Research designs such as field experiments, longitudinal field surveys, and multiple case studies have higher degrees of both internal and external validities. Personally, I prefer research designs that have reasonable degrees of both internal and external validities, i.e., those that fall within the cone of validity shown in Figure 5.1. But this should not suggest that designs outside this cone are any less useful or valuable. Researchers’ choice of designs is ultimately a matter of their personal preference and competence, and the level of internal and external validity they desire.

 **Construct validity** examines how well a given measurement scale is measuring the theoretical construct that it is expected to measure. Many constructs used in social science research such as empathy, resistance to change, and organizational learning are difficult to define, much less measure. For instance, construct validity must assure that a measure of empathy is indeed measuring empathy and not compassion, which may be difficult since these constructs are somewhat similar in meaning. Construct validity is assessed in positivist research based on correlational or factor analysis of pilot test data, as described in the next chapter.

**Statistical conclusion validity** examines the extent to which conclusions derived using a statistical procedure is valid. For example, it examines whether the right statistical method was used for hypotheses testing, whether the variables used meet the assumptions of that statistical test (such as sample size or distributional requirements), and so forth. Because interpretive research designs do not employ statistical test, statistical conclusion validity is not applicable for such analysis. The different kinds of validity and where they exist at the theoretical/empirical levels are illustrated in Figure 5.2.



Figure 5.2. Different Types of Validity in Scientific Research

#### Improving Internal and External Validity

The best research designs are those that can assure high levels of internal and external validity. Such designs would guard against spurious correlations, inspire greater faith in the hypotheses testing, and ensure that the results drawn from a small sample are generalizable to the population at large. Controls are required to assure internal validity (causality) of research designs, and can be accomplished in four ways: (1) manipulation, (2) elimination, (3) inclusion, and (4) statistical control, and (5) randomization.

In **manipulation**, the researcher manipulates the independent variables in one or more levels (called “treatments”), and compares the effects of the treatments against a control group where subjects do not receive the treatment. Treatments may include a new drug or different dosage of drug (for treating a medical condition), a, a teaching style (for students), and so forth. This type of control is achieved in experimental or quasi-experimental designs but not in nonexperimental designs such as surveys. Note that if subjects cannot distinguish adequately between different levels of treatment manipulations, their responses across treatments may not be different, and manipulation would fail.

The **elimination** technique relies on eliminating extraneous variables by holding them constant across treatments, such as by restricting the study to a single gender or a single socioeconomic status. In the **inclusion** technique, the role of extraneous variables is considered by including them in the research design and separately estimating their effects on the dependent variable, such as via factorial designs where one factor is gender (male versus female). Such technique allows for greater generalizability but also requires substantially larger samples. In **statistical control**, extraneous variables are measured and used as covariates during the statistical testing process.

Finally, the **randomization** technique is aimed at canceling out the effects of extraneous variables through a process of random sampling, if it can be assured that these effects are of a random (non-systematic) nature. Two types of randomization are: (1) **random selection**, where a sample is selected randomly from a population, and (2) **random assignment**, where subjects selected in a non-random manner are randomly assigned to treatment groups.

Randomization also assures external validity, allowing inferences drawn from the sample to be generalized to the population from which the sample is drawn. Note that random assignment is mandatory when random selection is not possible because of resource or access constraints. However, generalizability across populations is harder to ascertain since populations may differ on multiple dimensions and you can only control for few of those dimensions.

#### Popular Research Designs

As noted earlier, research designs can be classified into two categories – positivist and interpretive – depending how their goal in scientific research. Positivist designs are meant for theory testing, while interpretive designs are meant for theory building. Positivist designs seek generalized patterns based on an objective view of reality, while interpretive designs seek subjective interpretations of social phenomena from the perspectives of the subjects involved. Some popular examples of positivist designs include laboratory experiments, field experiments, field surveys, secondary data analysis, and case research while examples of interpretive designs include case research, phenomenology, and ethnography. Note that case research can be used for theory building or theory testing, though not at the same time. Not all techniques are suited for all kinds of scientific research. Some techniques such as focus groups are best suited for exploratory research, others such as ethnography are best for descriptive research, and still others such as laboratory experiments are ideal for explanatory research. Following are brief descriptions of some of these designs. Additional details are provided in Chapters 9-12.

**Experimental studies** are those that are intended to test cause-effect relationships (hypotheses) in a tightly controlled setting by separating the cause from the effect in time, administering the cause to one group of subjects (the “treatment group”) but not to another group (“control group”), and observing how the mean effects vary between subjects in these two groups. For instance, if we design a laboratory experiment to test the efficacy of a new drug in treating a certain ailment, we can get a random sample of people afflicted with that ailment, randomly assign them to one of two groups (treatment and control groups), administer the drug to subjects in the treatment group, but only give a placebo (e.g., a sugar pill with no medicinal value). More complex designs may include multiple treatment groups, such as low versus high dosage of the drug, multiple treatments, such as combining drug administration with dietary interventions. In a **true experimental design**, subjects must be randomly assigned between each group. If random assignment is not followed, then the design becomes **quasi-experimental**. Experiments can be conducted in an artificial or laboratory setting such as at a university (laboratory experiments) or in field settings such as in an organization where the phenomenon of interest is actually occurring (field experiments). Laboratory experiments allow the researcher to isolate the variables of interest and control for extraneous variables, which may not be possible in field experiments. Hence, inferences drawn from laboratory experiments tend to be stronger in internal validity, but those from field experiments tend to be stronger in external validity. Experimental data is analyzed using quantitative statistical techniques. The primary strength of the experimental design is its strong internal validity due to its ability to isolate, control, and intensively examine a small number of variables, while its primary weakness is limited external generalizability since real life is often more complex (i.e., involve more extraneous variables) than contrived lab settings. Furthermore, if the research does not identify ex ante relevant extraneous variables and control for such variables, such lack of controls may hurt internal validity and may lead to spurious correlations.

**Field surveys** are non-experimental designs that do not control for or manipulate independent variables or treatments, but measure these variables and test their effects using statistical methods. Field surveys capture snapshots of practices, beliefs, or situations from a random sample of subjects in field settings through a survey questionnaire or less frequently, through a structured interview. In **cross-sectional field surveys**, independent and dependent variables are measured at the same point in time (e.g., using a single questionnaire), while in **longitudinal field surveys**, dependent variables are measured at a later point in time than the independent variables. The strengths of field surveys are their external validity (since data is collected in field settings), their ability to capture and control for a large number of variables, and their ability to study a problem from multiple perspectives or using multiple theories. However, because of their non-temporal nature, internal validity (cause-effect relationships) are difficult to infer, and surveys may be subject to respondent biases (e.g., subjects may provide a “socially desirable” response rather than their true response) which further hurts internal validity.

**Secondary data analysis** is an analysis of data that has previously been collected and tabulated by other sources. Such data may include data from government agencies such as employment statistics from the U.S. Bureau of Labor Services or development statistics by country from the United Nations Development Program, data collected by other researchers (often used in meta-analytic studies), or publicly available third-party data, such as financial data from stock markets or real-time auction data from eBay. This is in contrast to most other research designs where collecting primary data for research is part of the researcher’s job. Secondary data analysis may be an effective means of research where primary data collection is too costly or infeasible, and secondary data is available at a level of analysis suitable for answering the researcher’s questions. The limitations of this design are that the data might not have been collected in a systematic or scientific manner and hence unsuitable for scientific research, since the data was collected for a presumably different purpose, they may not adequately address the research questions of interest to the researcher, and interval validity is problematic if the temporal precedence between cause and effect is unclear.

**Case research** is an in-depth investigation of a problem in one or more real-life settings (case sites) over an extended period of time. Data may be collected using a combination of interviews, personal observations, and internal or external documents. Case studies can be positivist in nature (for hypotheses testing) or interpretive (for theory building). The strength of this research method is its ability to discover a wide variety of social, cultural, and political factors potentially related to the phenomenon of interest that may not be known in advance. Analysis tends to be qualitative in nature, but heavily contextualized and nuanced. However, interpretation of findings may depend on the observational and integrative ability of the researcher, lack of control may make it difficult to establish causality, and findings from a single case site may not be readily generalized to other case sites. Generalizability can be improved by replicating and comparing the analysis in other case sites in a **multiple case design**.

**Focus group research** is a type of research that involves bringing in a small group of subjects (typically 6 to 10 people) at one location, and having them discuss a phenomenon of interest for a period of 1.5 to 2 hours. The discussion is moderated and led by a trained facilitator, who sets the agenda and poses an initial set of questions for participants, makes sure that ideas and experiences of all participants are represented, and attempts to build a holistic understanding of the problem situation based on participants’ comments and experiences. Internal validity cannot be established due to lack of controls and the findings may not be generalized to other settings because of small sample size. Hence, focus groups are not generally used for explanatory or descriptive research, but are more suited for exploratory research.

**Action research** assumes that complex social phenomena are best understood by introducing interventions or “actions” into those phenomena and observing the effects of those actions. In this method, the researcher is usually a consultant or an organizational member embedded within a social context such as an organization, who initiates an action such as new organizational procedures or new technologies, in response to a real problem such as declining profitability or operational bottlenecks. The researcher’s choice of actions must be based on theory, which should explain why and how such actions may cause the desired change. The researcher then observes the results of that action, modifying it as necessary, while simultaneously learning from the action and generating theoretical insights about the target problem and interventions. The initial theory is validated by the extent to which the chosen action successfully solves the target problem. Simultaneous problem solving and insight generation is the central feature that distinguishes action research from all other research methods, and hence, action research is an excellent method for bridging research and practice. This method is also suited for studying unique social problems that cannot be replicated outside that context, but it is also subject to researcher bias and subjectivity, and the generalizability of findings is often restricted to the context where the study was conducted.

**Ethnography** is an interpretive research design inspired by anthropology that emphasizes that research phenomenon must be studied within the context of its culture. The researcher is deeply immersed in a certain culture over an extended period of time (8 months to 2 years), and during that period, engages, observes, and records the daily life of the studied culture, and theorizes about the evolution and behaviors in that culture. Data is collected primarily via observational techniques, formal and informal interaction with participants in that culture, and personal field notes, while data analysis involves “sense-making”. The researcher must narrate her experience in great detail so that readers may experience that same culture without necessarily being there. The advantages of this approach are its sensitiveness to the context, the rich and nuanced understanding it generates, and minimal respondent bias. However, this is also an extremely time and resource-intensive approach, and findings are specific to a given culture and less generalizable to other cultures.

#### Selecting Research Designs

Given the above multitude of research designs, which design should researchers choose for their research? Generally speaking, researchers tend to select those research designs that they are most comfortable with and feel most competent to handle, but ideally, the choice should depend on the nature of the research phenomenon being studied. In the preliminary phases of research, when the research problem is unclear and the researcher wants to scope out the nature and extent of a certain research problem, a focus group (for individual unit of analysis) or a case study (for organizational unit of analysis) is an ideal strategy for exploratory research. As one delves further into the research domain, but finds that there are no good theories to explain the phenomenon of interest and wants to build a theory to fill in the unmet gap in that area, interpretive designs such as case research or ethnography may be useful designs. If competing theories exist and the researcher wishes to test these different theories or integrate them into a larger theory, positivist designs such as experimental design, survey research, or secondary data analysis are more appropriate.

Regardless of the specific research design chosen, the researcher should strive to collect quantitative and qualitative data using a combination of techniques such as questionnaires, interviews, observations, documents, or secondary data. For instance, even in a highly structured survey questionnaire, intended to collect quantitative data, the researcher may leave some room for a few open-ended questions to collect qualitative data that may generate unexpected insights not otherwise available from structured quantitative data alone. Likewise, while case research employ mostly face-to-face interviews to collect most qualitative data, the potential and value of collecting quantitative data should not be ignored. As an example, in a study of organizational decision making processes, the case interviewer can record numeric quantities such as how many months it took to make certain organizational decisions, how many people were involved in that decision process, and how many decision alternatives were considered, which can provide valuable insights not otherwise available from interviewees’ narrative responses. Irrespective of the specific research design employed, the goal of the researcher should be to collect as much and as diverse data as possible that can help generate the best possible insights about the phenomenon of interest.