Approaches to Theory and Method in Criminal Justice

THEORY

Theory in criminal justice represents an attempt to develop plausible explanations of reality, which in this case is crime and criminal justice system. Theory attempts to classify and organize events, to explain the causes of events, to predict the direction of future events, and to understand why and how these events occur (Turner, 1997, p.2). It represents a reasonable and informed guess as to why things are as they appear and to explain their underlying nature and meaning. The generation of theoretical explanations is what distinguishes a theory from merely a collection of war stories and carefully documented encyclopedic accounts. Theory asks: What is the point of all of this? What does it mean? Why are things this way?

Methodology, (methods) on the other hand, involves the collection of accurate facts and/or data regarding the nature of crime and criminal justice policy. *In short, while theory addresses the issue "why", methodology concerns itself with "what is"*. In any field, usually exists a certain division between those who are primarily interested in generating theory and who view their efforts as classical scholarship akin to philosophy and those (methodologists) who are viewed as technical and scientific in their approach. Good criminal justice requires both. Theory devoid of method - explanation without accurate supportive data - is just as mush ritualistic dead end as method devoid of theory. Both theory and method should be viewed as means to an end, the end being sound criminal justice knowledge.

PURE VERSUS APPLIED RESEARCH

Pure (basic) research is concerned with the acquisition of new knowledge for the sake of science or the development of the field, whereas **applied research** is practical research concerned with solving immediate policy problems. Criminal justice has experienced conflict between two camps, *the applied practitioner* and *the nonapplied academic*. Being on the front lines of the criminal justice system, practitioners are more concerned with *applied research*, studies, and findings that speak directly to policy issues. Academics, on the other hand, are more interested in *pure research*, which may have no immediate applicability but contributes to the knowledge base and scientific development of the discipline.

QUALITATIVE AND QUANTITATIVE RESEARCH

In **quantitative research** concepts are assigned numerical value, whereas in **qualitative research** concepts are viewed as sensitizing ideas or terms that enhance our understanding.

PRIMARY AND SECONDARY DATA ANALYSIS

Primary data analysis is the term used to describe a researcher's analysis of data that he or she has collected. So, if you create your own questionnaire or interview schedule, draw a sample of respondents, have them respond to your questions, code and enter the data into a computer, and then use a statistical package such as SPSS to complete you're doing primary data analysis. If you are working on a senior research seminar or master's or doctoral thesis that requires you to do original data collection, you will end up doing primary data analysis. This text will help you figure out how to plan the development of a questionnaire, draw samples, create codes, enter data, and do statistical analysis.

If you bring together data that other people or organizations originally collected and then use SPSS to complete the analysis, you're going **secondary data analysis.** This is a particularly important part of criminal justice research because of the availability of so many important secondary data sets, such as the official criminal justice system statistics that describe the operation of police courts, correctional facilities, juvenile justice organization, and so on. In addition to this official statistic, criminal justice researchers and other social scientists have conducted to a **data archive** maintained for use by others. The largest social science data archive is the Inter-university Consortium for political and Social Research (**ICPSR**), located at the University of Michigan. It has a special collection of criminal justice studies (the National Archives of Criminal Justice Data) maintained at the request of the National Institute of Justice. With the skills you'll learn in this book and a tool like SPSS , you could use this archive of data sets. Not only are these secondary data sets available for the United State, but four different international surveys have now been conducted under the sponsorship of the United Nations. One of the most useful sources of secondary data of American research is the **General Social Survey (GSS),** conducted with the federal support by the National Opinion Research Center (NORC). You'll be using the GSS data throughout this book, so it's worth an introduction (see Section 2.2)

There's even a third kind of data analysis, called **meta-analysis,** in which the "data" are the results of other completed analyses, so that each "case" is

a completed study. For example, several researchers (Ennett, Tobler, Ringwalt, & Flewelling, 1994) gathered all the published and unpublished studies of Project DARE, the largest drug prevention program for American children. Each "case" was a statistical summary of how large an effect (if any) the DARE program had on its participants. The researchers report that when they summarize all that is known about the program, there was little evidence that it had any preventive effect.

---------- page break ----------

Researchese: The Language of Research

When a person is suddenly confronted with an unfamiliar style of presentation, he/she experiences a sense of disorientation. In social sciences, the language of research is almost like being exposed to a foreign language, and this sense of disorientation can be easily described as **research shock** (Hagan, 2003, p. 20).

You have been introduced to **researchese** in CRJ 715 and may be familiar with it. When you use this jargon around others not in the field, you may be surprised that they are unaware of these terms. This, incidentally, explains why many occupational groups tend to stick together socially.

CONCEPTS

Concepts are abstract tags we put on reality and are the beginning point in all scientific endeavors. They are symbolic human creations or constructs that attempt to capture the essence of reality. In deciding on a name for a phenomenon, we are attempting to describe, understand, classify, or become more sensitized to some element of reality. Examples of concepts in criminal justice studies include crime, recidivism, police patrol, etc. Age, sex, race, religion, and social class are other concepts with which we are quite familiar. Concepts may be viewed as qualitative, sensitizing/global notions or they can be converted into *variables* through *operationalization*.

OPERATIONALIZATION

Operationalization defines concepts by describing how they will be measured. It can be defined in response to the statement: "I measured it by _____." Completion of this sentence constitutes the operationalization of the concept.

VARIABLES

Variables are concepts that have been operationalized. A variable, then, is *any entity that can take on different values*. OK, so what does that mean? Anything that can vary can be considered a variable. For instance, *age* can be considered a variable because age can take different values for different people or for the same person at different times. Similarly, *country* can be considered a variable because a person's country can be assigned a value. Theoretically, variables can be of a qualitative nature. For example, qualitative distinctions could be made regarding a person's age (*old* or *young*). The variable gender consists of two text values: *male* and *female*.

But, we can, if it is useful, assign quantitative values instead of the text values, but we don't have to assign numbers in order for something to be a variable. It's also important to realize that variables aren't only things that we measure in the traditional sense. For instance, in much social research and in program evaluation, we consider the treatment or program to be made up of one or more variables (i.e., the 'cause' can be considered a variable). So even the program can be considered a variable.

An *attribute (or category)* is a specific value on a variable. For instance, the variable *sex* or *gender* has two attributes: *male* and *female*. Or, the variable *agreement* might be defined as having five attributes:

- 1 = strongly disagree
- 2 = disagree
- 3 = neutral
- 4 = agree
- 5 = strongly agree

---------- page break ----------

LEVELS OF MEASUREMENT: TYPES OF VARIABLES

Variables may be measured on four levels:

- Nominal
- Ordinal
- Interval
- Ratio

Nominal level variables represent the simplest level of measurement. Objects are usually placed into mutually exclusive categories or types, and there is often no necessary quantitative or statistical meaning to numbers assigned to these categories, except as a convenience in distinguishing groups. Thus, any numbers assigned are merely qualitative descriptions, or labels, that enable us to keep track of differences. Demographic variables such as race, sex, religion, and county are examples of nominal variables. Values might be assigned, such as 1 to Protestant, 2 to Catholics, 3 to Jewish, 4 to Muslim, and 5 to other. Three Protestants, however, do not equal one Jewish. The numbers merely assist in categorizing qualitative distinctions. We refer to these measurements as a nominal because they "name." **Nominal variables** simply name the different categories constituting them.

Another example of this is writing down five telephone numbers of friends and calculating the "average" telephone number from these data, then calling it and waiting for the "average" friend to answer. In criminal justice, numbers could be arbitrarily assigned to different types of crimes in order to categorize them, for instance:

Categorization 1 (Most serious crime)	**Categorization 2 (Alphabetically)**
1.Homicide	1.Assault
2.Assault	2.Burglary
3.Robbery	3.Homicide
4.Burglary	4.Robbery

Obviously, numbers assigned to above have no mathematical meaning. That is, homicide (1) plus robbery (3) does not equal burglary (4), nor do 4 homicides equal one burglary.

Ordinal level variables. Many social scientific variables go a step beyond simply naming the different categories in some order: from low to high, from more to less, and so on. Whereas the nominal variables "religious affiliation" classifies people into different religious groups, "religiosity" might order them in groups, such very religious, somewhat religious, and not at all religious. And whereas the nominal variable "political party identification" simply distinguishes different groups (e.g., Democrats and Republicans), an ordinal measure of "political philosophy" might rank the very liberal, the somewhat liberals, the middle-of-road, the somewhat conservative, and the very conservative. Similarly, one can imagine a variable that ranges from no support to total support of death penalty.

Ordinal variables share the nominal-variable quality of distinguishing differences between people and add the quality of ranking those differences. At the same time, it is not meaningful to talk about the distance separating the categories that make up on an ordinal variable. For example, we have no basis for talking about the "amount of liberalism" separating the very liberal from the somewhat or the somewhat liberal from the middle-of-the-road. We can say that the first group in each comparison is more liberal than the second, but we can't say by how much.

Ratio level variables not only assume the interval quality of data, but they also have a fixed meaningful zero point. Such data enable us to show how many time greater one value is from another. Some variables allow us to speak more precisely about the distances between the categories constituting a variable. Consider age for a moment. The distance between 10 years old and 20 years old is exactly the same as that between 60 years old and 70 years old. Similarly, the time served in a prison is a ratio variable: 30 months and 20 months are separated by the same amount of time 40 months and 50 months. Moreover, **ratio variables** such as age have the additional quality of containing a genuine zero point. This is what allows us to examine ratios between the categories constituting such variables. Thus, we can say that a 20-years-old is twice as old as a 10-years-old. A prisoner who has served 40 months has been in prison exactly twice as long as one serving 20 months. By comparison, notice that we would have no grounds for saving one person is twice as religious as another. Ratio variables, then, share all the qualities associated with nominal and ordinal variables but have additional qualities not applicable to the lower-level measures. Other examples of ratio measures include income, age, years of schooling, and number of delinquent acts.

Rarer in social research are variables that have the quality standard intervals of measurement but that lack a genuine zero point: **interval variables.** One example is the intelligence quotient (IQ). Although is it calculated in such a

way as to allow for a score of zero, that would not indicate a complete lack of intelligence because the person would have at least been able to take the test. Moving outside the social sciences, consider temperature. The Celsius and Fahrenheit measures of temperature both have zero-degree marks, but neither represents a total lack of heat, given that it is possible to have temperatures below zero, the Kelvin scale, by contrast, is based on an absolute zero, which does represent a total lack of heat (measures in terms of molecular motion); it is therefore a ratio variable.

As a way of summarizing the levels of measurement, we suggest that you inspect the following table, which lays out very clearly how each succeeding level of measurement adds even more information to what you know about a variable.

	Nominal Variables	*Ordinal Variables*	*Interval Variables*	*Ratio Variables*
Mutually Exclusive	*Yes*	*Yes*	*Yes*	*Yes*
Exhaustive	*Yes*	*Yes*	*Yes*	*Yes*
Ranking		*Yes*	*Yes*	*Yes*
Intervals			*Yes*	*Yes*
Zero point				*Yes*

DEPENDENT AND INDEPENDENT VARIABLES

Another important distinction having to do with the term *variable* is the distinction between an ***independent*** and ***dependent*** variable. This distinction is particularly relevant when you are investigating cause-effect relationships. We must learn this distinction.

(In all fairness, it's as "easy" as the signs for *arrivals* and *departures* at airports?

Do I go to arrivals because I'm arriving at the airport? or -- Does the person I'm picking up go to arrivals because they're arriving on the plane?!

The **dependent variable** (*outcome*) is the variable one is attempting to predict. By convention is represented by the letter Y. Common dependent variables in criminal justice are concepts such as crime and recidivism. The **independent variable** (*predictor*) is the variable that causes, determines, or precedes in time the dependant variable and is usually denoted by the letter X. An independent variable in one study could become a dependent variable in another. For example, a study of the impact of poverty (X) upon crime (Y) [poverty crime] finds poverty as the independent variable, whereas a study that looks at race (X) as a predictor of poverty (Y) [race poverty] finds poverty as a dependent variable. As a rule of thumb, the treatment variable is always an independent variable, as are demographic variables, such as age, sex, and race. The dependent variable usually is the behavior/attitudes.

Finally, there are two traits of variables that should always be achieved.

1. Each variable should be ***exhaustive***, it should include all possible answerable responses. For instance, if the variable is "religion" and the only options are "Protestant", "Jewish", and "Muslim", there are quite a few religions I can think of that haven't been included. The list does not exhaust all possibilities. On the other hand, if you exhaust all the possibilities with some variables -- religion being one of them -- you would simply have too many responses. The way to deal with this is to explicitly list the most common attributes and then use a general category like "Other" to account for all remaining ones.
2. In addition to being exhaustive, the attributes of a variable should be *mutually* ***exclusive***, no respondent should be able to have two attributes simultaneously. While this might seem obvious, it is often rather tricky in practice. For instance, you might be tempted to represent the variable "Employment Status" with the two attributes *"employed"* and *"unemployed"*. But these attributes are not necessarily mutually exclusive -- a person who is looking for a second job while <u>self-employed</u> would be able to check both attributes! But don't we often use questions on surveys that ask the respondent to "check all that apply" and then list a series of categories? Yes, we do, but technically speaking, each of the categories in a question like that is its own variable and is treated dichotomously as either "checked" or "unchecked", attributes that *are* mutually exclusive.

---------- page break ----------

HYPOTHESES

Hypotheses are specific statements regarding the relationship between (usually two) variables and are derived from more general theories.

A *research hypothesis* states an expected relationship between variables in positive terms, For example: poverty causes crime.

The *alternative hypothesis* is the exact opposite statement of the research hypothesis. For example: poverty doesn't cause crime.

The *null hypothesis* is a hypothesis of no difference and is the one actually tested statistically. For example: poverty is not related to crime.

One approach to research involves the formulation of hypotheses, the operationalization or measurement of the variables, and the testing or bringing the evidence. The figure below outlines a model of the research process.

DEDUCTION AND INDUCTION

Deduction involves moving from a level of theory to a specific hypothesis, whereas **induction** entails inferring about a whole group on the basis of knowing about a case or a few cases. Thus, theory to fact is *deduction* and fact to theory is *induction*. Sherlock Holmes' famous compliment to Dr. Watson, "Brilliant deduction, my dear Watson" should probably have read "induction" because Watson, in helping Holmes solve a case, was proceeding from specific facts or evidence to a conclusion of a theory.

---------- page break ----------

UNIT OF ANALYSIS

One of the most important ideas in a research project is the *unit of analysis*. The unit of analysis is the major entity that you are analyzing in your study. For instance, any of the following could be a unit of analysis in a study:

- individuals
- groups
- artifacts (books, photos, newspapers)
- geographical units (town, census tract, state)
- social interactions (dyadic relations, divorces, arrests)

Why is it called the 'unit of analysis' and not something else (like, the unit of sampling)? Because *it is the analysis you do in your study that determines what the unit is*. For instance, if you are comparing the children in two classrooms on achievement test scores, the unit is the individual child because you have a score for each child. On the other hand, if you are comparing the two classes on classroom climate, your unit of analysis is the group, in this case the classroom, because you only have a classroom climate score for the class as a whole and not for each individual student. For different analyses in the same study you may have different units of analysis. If you decide to base an analysis on student scores, the individual is the unit. But you might decide to compare average classroom performance. In this case, since the data that goes into the analysis is the average itself (and not the individuals' scores) the unit of analysis is actually the group. Even though you had data at the student level, you use aggregates in the analysis. In many areas of social research these hierarchies of analysis units have become particularly important and have spawned a whole area of statistical analysis sometimes referred to as *hierarchical modeling*. This is true in education, for instance, where we often compare classroom performance but collected achievement data at the individual student level.

---------- page break ----------

TYPES OF QUESTIONS

There are three basic types of questions that research projects can address:

1. **Descriptive**. When a study is designed primarily to describe what is going on or what exists. Public opinion polls that seek only to describe the proportion of people who hold various opinions are primarily descriptive in nature. For instance, if we want to know what percent of the population would vote for a Democratic or a Republican in the next presidential election, we are simply interested in describing something.
2. **Relational**. When a study is designed to look at the relationships between two or more variables. A public opinion poll that compares what proportion of males and females say they would vote for a Democratic or a Republican candidate in the next presidential election is essentially studying the relationship between gender and voting preference.
3. **Causal**. When a study is designed to determine whether one or more variables (e.g., a program or treatment variable) causes or affects one or more outcome variables. If we did a public opinion poll to try to determine whether a recent political advertising campaign changed voter preferences, we would essentially be studying whether the campaign (cause) changed the proportion of voters who would vote Democratic or Republican (effect).

The three question types can be viewed as cumulative. That is, a relational study assumes that you can first describe (by measuring or observing) each of the variables you are trying to relate. And, a causal study assumes that you can describe both the cause and effect variables and that you can show that they are related to each other. Causal studies are probably the most demanding of the three.

---------- page break ----------

CAUSALITY

The ultimate purpose of all scientific investigation is to isolate, define, and explain the relationship between key variables in order to predict and understand the underlying nature of reality. The problem of **causality** has been a subject of continuing philosophical discussion, but scientific investigation is based on the a priori assumption that the fundamental nature of reality can be known - that causation lies at the basis of reality.

Resolution of the Causality Problem

To approach this matter, scientific investigation entails basically three *essential* **Steps for Resolving the Causality Problem.**

1. *The first step* involves the demonstration of a **relationship or covariance between variables**. That is, one variable is related, increases or decreases in value, in some predictable manner along with increases or decreases in value of another variable.
2. *The second step* consists of specifying or indicating **the time sequence of the relationship**. Which variable id the independent variable X, and which is the outcome or dependent variable Y? Generally, logic or knowledge of which variable comes first gives one the direction of causation. For instance, it would make more sense that the criminality of parents (X) would precede in time and possibly predict criminality in offspring (Y), rather than vice versa.
3. *The third step* is the stage where many studies bog down and where research findings are subject to prolonged debate. It involves the **exclusion of rival causal factors**, or the elimination of other variables that could conceivably explain the original relationships the researcher had claimed.

---------- page break ----------

TYPES OF RELATIONSHIPS

A relationship refers to the *correspondence* between two variables. When we talk about types of relationships, we can mean that in at least two ways: the *nature* of the relationship or the *pattern* of it.

THE NATURE OF A RELATIONSHIP

While all relationships tell about the correspondence between two variables, there is a special type of relationship that holds that the two variables are not only in correspondence, but that one *causes* the other. This is the key distinction between a simple *correlational relationship* and a *causal relationship*. A correlational relationship simply says that two things perform in a synchronized manner. For instance, we often talk of a correlation between inflation and unemployment. When inflation is high, unemployment also tends to be high. When inflation is low, unemployment also tends to be low. The two variables are correlated. But knowing that two variables are correlated does not tell us whether one *causes* the other. We know, for instance, that there is a correlation between the number of roads built in Europe and the number of children born in the United States. Does that mean that is we want fewer children in the U.S., we should stop building so many roads in Europe? Or, does it mean that if we don't have enough roads in Europe, we should encourage U.S. citizens to have more babies? Of course not. (At least, I hope not). While there is a relationship between the number of roads built and the number of babies, we don't believe that the relationship is a *causal* one. This leads to consideration of what is often

termed the *third variable problem*. In this example, it may be that there is a third variable that is causing both the building of roads and the birthrate, that is causing the correlation we observe. For instance, perhaps the general world economy is responsible for both. When the economy is good more roads are built in Europe and more children are born in the U.S. The key lesson here is that you have to be careful when you interpret correlations.

If you observe a correlation between the number of hours students use the computer to study and their grade point averages (with high computer users getting higher grades), you *cannot* assume that the relationship is *causal*: that computer use improves grades. In this case, the third variable might be socioeconomic status -- richer students who have greater resources at their disposal tend to both use computers and do better in their grades. It's the resources that drive both use and grades, not computer use that causes the change in the grade point average.

PATTERNS OF RELATIONSHIPS

We have several terms to describe the major different types of patterns one might find in a relationship. First, there is the case of ***no relationship*** at all. If you know the values on one variable, you don't know anything about the values on the other. For instance, I suspect that there is no relationship between the length of the lifeline on

your hand and your grade point average. If I know your GPA, I don't have any idea how long your lifeline is.

Then, we have the ***positive relationship***. In a positive relationship, high values on one variable are associated with high values on the other and low values on one are associated with low values on the other. In this example, we assume an idealized positive relationship between years of education and the salary one might expect to be making.

On the other hand a ***negative*** relationship implies that high values on one variable are associated with low values on the other. This is also sometimes termed an ***inverse*** relationship. Here, we show an idealized negative relationship between a measure of self esteem and a measure of paranoia in psychiatric patients.

These are the simplest types of relationships we might typically estimate in research. But the pattern of a relationship can be more complex than this. For instance, the figure on the left shows a relationship that changes over the range of both variables, a curvilinear relationship. In this example, the horizontal axis represents dosage of a drug for an illness and the vertical axis represents a severity of illness measure. As dosage rises, severity of illness goes down. But at some point, the patient begins to experience negative side effects associated with too high a dosage, and the severity of illness begins to increase again.