1977

Integrated sociological theory construction: from formulation to verification

Caroline Sue Faisal

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd

Part of the Quantitative, Qualitative, Comparative, and Historical Methodologies Commons, and the Theory, Knowledge and Science Commons

Recommended Citation

Faisal, Caroline Sue, "Integrated sociological theory construction: from formulation to verification " (1977). Retrospective Theses and Dissertations. 6071.
https://lib.dr.iastate.edu/rtd/6071

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR
FAISAL, Caroline Sue, 1942-
INTEGRATED SOCIOLOGICAL THEORY CONSTRUCTION: FROM FORMULATION TO VERIFICATION.

Iowa State University,
Ph.D., 1977
Sociology, statistics-research methods

University Microfilms International, Ann Arbor, Michigan 48106

© 1978
CAROLINE SUE FAISAL
All Rights Reserved
Integrated sociological theory construction:
From formulation to verification

by

Caroline Sue Faisal

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Department: Sociology and Anthropology
Major: Sociology

Approved:

Signature was redacted for privacy.

For the Major Department
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1977

copyright © Caroline Sue Faisal, 1978. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>INTRODUCTION</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statement of Problem</td>
<td>2</td>
</tr>
<tr>
<td>Objectives</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PART I: ESTABLISHING THE NEED FOR INTEGRATED THEORY CONSTRUCTION</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER ONE: SURVEYING THE CURRENT STATE OF SOCIOLOGICAL THEORY CONSTRUCTION</td>
<td>10b</td>
</tr>
<tr>
<td>A Survey of Theory Construction Courses</td>
<td>11</td>
</tr>
<tr>
<td>A General Description of Theory Construction</td>
<td>15</td>
</tr>
<tr>
<td>Summary</td>
<td>24</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER TWO: ASSESSMENT OF CURRENT THEORY CONSTRUCTION TEXTS</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Question of Coverage</td>
<td>32</td>
</tr>
<tr>
<td>A Question of Integration</td>
<td>40</td>
</tr>
<tr>
<td>Summary</td>
<td>47</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER THREE: AN OVERVIEW OF A COMPREHENSIVE APPROACH TO SOCIOLOGICAL THEORY CONSTRUCTION</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Developing the Approach</td>
<td>52</td>
</tr>
<tr>
<td>A Comprehensive Technique for Theory Construction</td>
<td>60</td>
</tr>
<tr>
<td>Summary</td>
<td>68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PART II: A DETAILED DISCUSSION OF AN INTEGRATED COMPREHENSIVE APPROACH TO THEORY CONSTRUCTION</th>
<th>70</th>
</tr>
</thead>
<tbody>
<tr>
<td>THEORETICAL FORMULATION</td>
<td>71</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>TITLE</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>FOUR</td>
<td>WHAT IS A THEORY 74</td>
</tr>
<tr>
<td>CHANGEP</td>
<td>SUMMARY</td>
</tr>
<tr>
<td>FIVE</td>
<td>THE STATING OF ASSUMPTIONS 87</td>
</tr>
<tr>
<td></td>
<td>THEORETICAL ORIENTATION</td>
</tr>
<tr>
<td></td>
<td>LEVEL OF ABSTRACTION</td>
</tr>
<tr>
<td></td>
<td>HOW TO APPROACH THE SUBJECT MATTER</td>
</tr>
<tr>
<td></td>
<td>IMPACT OF VALUES</td>
</tr>
<tr>
<td></td>
<td>SUMMARY</td>
</tr>
<tr>
<td>SIX</td>
<td>STATEMENT OF THE PROBLEM 124</td>
</tr>
<tr>
<td></td>
<td>THEORETICAL BACKGROUND</td>
</tr>
<tr>
<td></td>
<td>SPECIFYING THE UNIT (S) OF ANALYSIS</td>
</tr>
<tr>
<td></td>
<td>A CAUSAL MODEL OF EFFECTIVENESS</td>
</tr>
<tr>
<td></td>
<td>SUMMARY</td>
</tr>
<tr>
<td>SEVEN</td>
<td>SPECIFYING CONCEPTS AND MEASURES 160</td>
</tr>
<tr>
<td></td>
<td>ADEQUACY OF CONCEPT</td>
</tr>
<tr>
<td></td>
<td>CORRESPONDENCE BETWEEN CONCEPTUAL AND OPERATIONAL DEFINITIONS</td>
</tr>
<tr>
<td></td>
<td>ADEQUACY OF INDICATOR</td>
</tr>
<tr>
<td></td>
<td>SUMMARY</td>
</tr>
<tr>
<td>EIGHT</td>
<td>LINKAGES, PROPOSITIONS AND THEIR SYSTEMIZATION 200</td>
</tr>
<tr>
<td></td>
<td>STATING THE NATURE OF THE RELATIONSHIP</td>
</tr>
<tr>
<td></td>
<td>PROPOSITIONS AND HYPOTHESES</td>
</tr>
<tr>
<td></td>
<td>SYSTEMATIZATION OF PROPOSITIONS</td>
</tr>
<tr>
<td></td>
<td>SUMMARY</td>
</tr>
<tr>
<td>NINE</td>
<td>MEASUREMENT</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Considerations Prior to Data Collection</td>
</tr>
<tr>
<td></td>
<td>Preliminary Assessment Techniques</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
</tr>
<tr>
<td></td>
<td>ANALYSIS</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER TEN: RELATING PROPOSITIONS TO STANDARD ANALYSIS PROCEDURES</strong></td>
</tr>
<tr>
<td></td>
<td>Hypotheses and Statistical Techniques</td>
</tr>
<tr>
<td></td>
<td>Correlation</td>
</tr>
<tr>
<td></td>
<td>Simple or Single Regression</td>
</tr>
<tr>
<td></td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER ELEVEN: MULTIPLE REGRESSION</strong></td>
</tr>
<tr>
<td></td>
<td>Examining Residuals</td>
</tr>
<tr>
<td></td>
<td>Comparison of Alternative Models</td>
</tr>
<tr>
<td></td>
<td>Searching for an Explanation</td>
</tr>
<tr>
<td></td>
<td>Examining the Assumptions of Multiple Regression</td>
</tr>
<tr>
<td></td>
<td>Summary</td>
</tr>
<tr>
<td></td>
<td><strong>CHAPTER TWELVE: PATH ANALYSIS</strong></td>
</tr>
<tr>
<td></td>
<td>The Role of Theory in Path Analysis</td>
</tr>
<tr>
<td></td>
<td>The Method</td>
</tr>
<tr>
<td></td>
<td>Examination of an Effectiveness Model</td>
</tr>
<tr>
<td></td>
<td>Decomposition of Effects</td>
</tr>
</tbody>
</table>
INTRODUCTION

A recently published text by William Skidmore begins with the following description of sociological theory...

The term sociological theory has a variety of meanings and usages. This variety has on occasion led to confusion among sociologists and persons studying sociology, because two or more persons may not agree about the theoretical meaning of an idea. Because of such misunderstanding, the idea itself may be lost or misinterpreted (1975:1).

If one takes this as a fairly accurate description of the current state of theory in sociology today, it should not be surprising to find that there is a similar lack of agreement on what is meant by the term theory construction. For if sociologists disagree as to what is theory and what is not, it is only reasonable that they would also disagree on the "correct" procedure to construct that theory. This is in fact what was found in a recent opinion survey to determine the existence and nature of theory construction courses at universities throughout the U.S and Canada which was undertaken in the last few years by two graduate seminars at Iowa State University.

The survey revealed a wide divergence of opinion on the meaning of theory construction as well as on the materials most relevant to teaching it. For some, theory construction represented a talent which only a few would ever possess, a talent similar to that which is required to be a great
painter or musician. For others, theory construction is not an art but an indispensable tool which more sociologists should master if the discipline is to advance. For those who thought theory construction could be taught, there still was no consensus as to how it should be taught. Thus diversity tended to be the rule rather than the exception. At the end of the survey, therefore, the question still remained "What is a good approach to theory construction?"

Statement of Problem

Over the years, as sociology has developed, its practitioners have tended to specialize in either theory or research methods. In fact, a wide gulf has developed between adherents of the two approaches with articles flying back and forth across the gulf attacking the others stand, while attempting to emphasize the strong points of their own. Researchers are said to be playing with numbers while theorists are accused of "existing in an ivory tower" playing with abstractions that have nothing to do with reality.

There seems to be a growing realization that both have been correct and wrong at the same time. What is needed is not one or the other, but both. And not even just both but both in an integrated fashion such that activities in one phase impinge upon the activities in the next. Some of the classical writers (e.g. Durkheim, Weber, and even Parsons) of sociological theory recognized this and even wrote concerning
the need for a comprehensive and integrated approach. They did not agree on what that approach should be, but at least they recognized the need for one. In the chapters to come this need will be reiterated by many of the authors of currently required texts on theory construction.

The problem remains, however, that none of those examined have published such a text, even though many cited this as their goal. It would seem therefore that a comprehensive method of theory construction is needed which will (1) incorporate theoretical formulation, measurement, analyses and inferential activities and (2) do so in an integrated context, demonstrating at each step, the consequences for those that follow. This text will attempt to do just that. A comprehensive approach to theory construction, involving activities from each of the areas in (1) above, will be presented with special emphasis on the integrated and interdependent (2) nature of those activities.

*Why is this important?*

Theory which is developed in terms of theoretical formulation only, remains little more than speculation until it has been subjected to verification in the real world. The theoretical formulation phase must, therefore become associated with some form of test which requires both data (measurement) and analysis.
On the other hand, research which has no theoretical direction can only speculate that a particular interpretation placed on analysis results is the "correct" one. It is the theoretical framework which helps us sort out and organize the most plausible alternatives.

The integration aspect is important because breakdowns in the linkage between phases and among activities, results in tenuous inferences. Thus, if a measure does not correspond to its conceptual and/or real world counterparts, what difference does it make that a statistical test proved to be significant? What does it mean? It may mean no more than the fact that a statistical relationship exists between some unknown measure and another measure. Certainly no inferences should be made to the proposition containing the ill-measured concept.

How can I contribute to the solution of the problem?

The subject matter which can be included in a text covering theoretical formulation, measurement, and analyses is vast, which is probably one of the prime reasons that others have not proposed one earlier. Obviously someone needs to organize, as well as condense, this massive amount of literature into a step-by-step procedure. The one which will be presented here will be geared to both the novice and to those who need a guide, reference, or occasional refresher. I intend to do this by accomplishing the
following objectives.

Objectives

1. Review the present state of theory construction, selecting texts which are currently being used in teaching theory construction in an attempt to establish a definition of comprehensive theory construction. Currently utilized texts will be used since this proposes to be an example of "in-use" as well as an example of "reconstructed logic".

2. Make certain that the scope of the material covered includes important activities in all four phases of the theory construction process.

3. Attempt to include material and explanations at varying depths so that the text can continue to serve as a guide as the reader gains more expertise.

4. Call attention to the interdependent nature of these activities and the subsequent necessity for their integration.

5. Demonstrating aspects of the various techniques with a single theoretical framework and data set. A theory or technique which cannot be applied is of little use. Applications will be included in order to illustrate the parts as well as the whole procedure.

How will these objectives be met?

In spite of the many differences alluded to in the survey above, some commonalities were revealed. Part One concerns itself primarily with an examination of these common threads in an attempt to establish a working definition of sociological theory construction (Chapter One). The definition, which evolves in Chapter One, stresses the need for an interdependent relationship between theoretical and
methodological activities. Having established the definition for sociological theory construction, Chapter Two is concerned with assessing the degree to which five of the most commonly cited texts for theory construction courses meet the criteria set forth in that definition. Common to practically all of the theory construction texts examined was the call for developing an approach to theory construction which could unite the disparate activities of theory formulation and methodology. However, the general conclusion of Chapter Two is that none of these texts, can be said to elucidate the kind of integrated procedure for theory construction which is contained in the definition cited above. Most of the cited texts fail to cover both theoretical and methodological activities adequately. Furthermore, none of the texts consistently demonstrated the interdependent nature of these activities. Given the definition of theory construction that is developed in Chapter One, the assumption is made that sociologists must learn to utilize both types of activities, not as disjointed techniques, but as interdependent tools leading to the construction of a well-organized, integrated and testable theory. Having shown that sociology lacks a comprehensive procedure for constructing such theories, Chapter Three attempts to suggest one. A brief overview of the general philosophy and procedure of this comprehensive strategy are outlined in Chapter Three as well as the basis
for some of the discussions that were made. This concludes part I of the dissertation.

Part II is concerned with a detailed explanation of the various activities required in the comprehensive approach to theory construction that was introduced in Chapter Three. This part of the dissertation has been divided into four sections, each section representing one of the four phases of theory construction, (theoretical formulation, measurement, analysis, and inference) that are required by the comprehensive definition. Each of these sections have then been divided into relevant chapters.

The first phase is concerned with theoretical formulation, the first chapter of which (Chapter Four) develops the definition for sociological theory that will be utilized throughout the remainder of the text. The next chapter (Chapter Five) deals with the stating of assumptions (e.g., theoretical orientation, values and etc.) that tend to set general boundaries on the subject matter that will be examined as well as how it should be examined. Chapter Six is also concerned with boundary setting topics, although on a much more specific level. This specificity is further demonstrated by introducing the theoretical framework for a substantive example that will be utilized throughout the remainder of the text as a means of illustrating the various topics that will be discussed. Selection of the concepts and
measures as well as their assessment are the major topics of Chapter Seven. The last chapter within the theoretical formulation section is primarily concerned with specifying linkages which connect concepts to form hypotheses and propositions. A number of alternative methods for systematizing those propositions in a theory or model is also reviewed. In the process, an introduction is made to causal modeling which will serve as the basis for most of the comments in the last two phases, analysis and inferences.

Although many current theory construction books are content to describe only the activities within theoretical formulation, it is the contention of this dissertation, as described above, that "theory," which has been developed strictly by these activities is little more than supposition.

To develop a solid theory, one must submit the proposed theory to the activities in the other three phases. Thus Chapter Nine covers measurement activities in terms of measurement considerations prior to data collection and preliminary assessment techniques. Analysis procedures are examined in three chapters. The subject matter in these chapters range from simple bivariate techniques (Chapter Ten) to more complicated multivariate procedures: Multiple Regression (Chapter Eleven); and Path Analysis (Chapter Twelve).
The final phase discussed in Part II is concerned with inference. Inference consists of drawing conclusions and generalizing them to the propositions that apply to the population of interest. An attempt will be made in the discussion chapter (Chapter Thirteen) to demonstrate by turning to an examination of a model using real data that inference is influenced by activities in the other phases as well as representing an activity that "puts everything together." Another aspect of this chapter which is related to accomplishing the goals of the dissertation (rather than constituting an integral part of the theory construction approach) will be to assess the degree to which the approach presented in this part of the dissertation fulfills the criteria that was originally outlined in Chapter Two. This constitutes the same criteria that was utilized to assess existing theory construction texts.

The comprehensive approach to theory construction that will be proposed here assumes a broad definition of what is meant by the term theory construction, which will in fact be developed in the next chapter. By adopting this broad notion, the text is recognizing a concern of many prominent sociologists (i.e., the need for the integration of theory and methods) which as yet, has not been satisfied. The great diversity of material that is required to teach a comprehensive course on theory construction attests to the
fact that there is no single work which emphasizes and demonstrates the necessary interdependence of these activities. Thus the procedure developed here can be said to represent a synthesis of material that has hitherto only been found scattered throughout extensive volumes of theory and methodology texts, journals and papers. The task of integrating such a large mass of material has been too time consuming for most sociologists, even for those who may be familiar with many of the separate techniques. Thus this text will attempt to be more than just a source book of different theoretical and methodological issues which is essentially what books such as Kaplan, (1964), or Kerlinger, (1964), attempt to do. What sets this work apart from these previous attempts, in addition to the range of material covered, is the way in which these theoretical and methodological activities are integrated into a whole process where each step is shown to be interdependent on all others.
PART I:
ESTABLISHING THE NEED FOR INTEGRATED
THEORY CONSTRUCTION
CHAPTER ONE:
SURVEYING THE CURRNT STATE OF
SOCIOLOGICAL THEORY CONSTRUCTION

The field, or subfield, of theory construction, is relatively new in sociology. In fact, Hans Zetterberg's *On Theory and Verification* (1966) can be said to mark the first of a growing number of recognized texts primarily concerned with describing the "proper" techniques for constructing "good" theory. However, although the field itself is new, this should not be construed as saying that only recent social thinkers have been concerned with the act of constructing or developing theory.

If one examines some of the most prominent early theoretical formulations that are still influencing social thinking today, what is revealed is an intimate interrelationship between the social writer's method and the theory he developed. Emile Durkheim's second important book was *The Rules of Sociological Method*, written in 1894, which was followed in 1897 by *Suicide*, an attempt to apply some of his methods to a concrete problem: suicide as an index of social disintegration (Fybate, 1970). His method consisted of three basic steps (1) defining the phenomenon in question in terms of easily recognizable external features, (2) refuting inadequate explanations, and finally (3) offering
his own explanation. These elements were basic to the
development of all Durkheim's theories regardless of the
particular phenomenon which was examined (Coser, 1971: 142).

Another early social thinker whose works are still
influential today was Max Weber. Unlike Durkheim, whose
methods are considered today only in relation to the works
they helped to develop, Max Weber's methods of Verstehen
(interpretive understanding) and Ideal Type (a form of
comparative analysis) are themselves the subject of study
(see Shils and Finch, 1949), as well as the works he produced
with them. But methodology for its own sake was not Weber's
concern. Methods were important tools which facilitated the
development of ideas and theories.

"Though a number of examples could be cited
where Weber did not apply his methodological
injunctions, many more instances in his work
reveal that he put his methods to brilliant use
in his substantive analysis" (Coser, 1971:226).

A more recent example of a social theorist who has had a
great deal of influence is Talcott Parsons. Like Durkheim
and Weber before him, Parsons' method is an integral part of
the theory he developed. This is perhaps one of the things
which has been repeatedly misunderstood by his critics. One
such example was, Turk (1967), who criticized Parsons because
he didn't write a deductive set of theoretical statements,
which, of course, Parsons deliberately avoided, since
according to his strategy for theory building such
inventories of propositions are written or developed only after one has constructed a system of interrelated concepts. For Parsons then, theory should be an evolutionary process whereby the building of a system of abstract interrelated concepts represents only the first step in the Parsonian methodology. Some critics, like C. Wright Mills (1959), have argued that Parsons' theory deals only in logic, definitions and highly abstract or grand theory which is too far removed from the empirical world to have any significance. He was emphasizing part of the dichotomy which has arisen in sociology between the so-called pure theorists (who either ignore empirical data and/or methods or tack them on as an afterthought) and the empiricists or methodologists (for whom the gathering and analysis of data are primary with theory seemingly ignored or tacked on). While the dichotomy is certainly quite evident in much of the sociological publications today, Mills may very well have been making a mistake, not unlike that made by Turk, (i.e., ignoring the implications of Parsonian methodology). By emphasizing only the earlier works, out of the context of the methodology, it is not difficult to agree with Mills' placement of Parsons

---

1An examination of 200 articles by Klonglan et al., (1973) related to IOR (interorganizational relations) revealed few, if any, articles which included an adequate emphasis of both theoretical and methodological concerns.
with the pure theorists. However, in his later works, particularly those on evolution (1966, 1971), Parsons did make an attempt to utilize some of his concepts to develop testable propositions. The somewhat dubious success of that attempt (Turner, 1974) does not detract from the fact that his methodology recognizes a need to combine both the abstract and the empirical, and that he is in the process of attempting to do just that.

Certainly the dichotomy between theory and methodology which seems to exist in much of the current sociological work does not appear in some of the most enduring social thought, notably the works of Durkheim and Weber or even in Parsons', as discussed above. Methodology has been shown to be an integral part of the theories which they developed. Perhaps then current sociologists who want to develop "lasting" theory should take their cue from these "masters" --- integrate methodological and theoretical concerns in the development (construction) of theory.

Turning attention back to current efforts at theory construction, it was mentioned that theory construction as a separate field or subfield in sociology is a surprisingly new area of concentration. Kuhn (1962) has pointed out that in any new endeavor, there is a period of time, in the beginning, when everyone goes his own way. With time, however, a synthesis of ideas begins to occur and eventually
the field gains systematization. One might well ask to what extent systematization has occurred in the new area of sociological theory construction. In an effort to answer this question, graduate students at Iowa State University conducted a survey during the winter quarters of 1973 and 1974. A questionnaire was sent to all sociology departments in the U.S. and a number in Canada. The results of that survey will be examined in the remainder of this chapter in an attempt to

1) discover what is currently meant by sociologists when they use the term theory construction,

2) outline some of the elements which are considered as most basic to the understanding and teaching of sociological theory construction,

3) establish a working definition of sociological theory construction by building on the results of both the survey and the previously noted examples set by earlier theorists, i.e., integrating methodological and theoretical concerns.

A Survey of Theory Construction Courses

In the winter quarter, 1973, a questionnaire was sent out to sociology departments throughout the U.S. and Canada as part of a class project in a theory construction seminar at Iowa State University. The following year when the seminar was again held, a new group of students sent follow-up letters and questionnaires to those departments who had failed to answer the first time. A total of 200
schools were contacted. Of these, responses were received from sixty-four. The purpose of the questionnaire was to ascertain the number and nature of theory construction courses as they are being taught throughout American universities.

As Table 1.1 on the next page demonstrates, a majority of the responding schools have begun to deal with the area of theory construction. Only seventeen, or roughly one-fourth, of the schools did not have such an emphasis, even though many of these answered in the affirmative to the question, "Do you teach theory construction either as a separate course or as part of other courses?" These schools were placed in the "none" category because it was clear from the material covered in the courses that the subject matter dealt with critiques of social thought and/or theoretical perspectives rather than the construction of theory. Comments of the others included in the "none" category ranged from "at the present time we only have a masters program but if and when we get our Ph.D. program going we intend to add such a course." to "consider it only a passing fad." Another frequent comment was to the effect that the construction of theories must be considered in the same way as the construction of a fine piece of music—it represents an art for which one either has a talent or one does not. Those who exhibit such a talent can be encouraged and directed but it
is futile to attempt to instruct the majority of sociologists in an art which can only be nurtured not instilled.

Table 1.1: Nature of current theory construction courses

<table>
<thead>
<tr>
<th>Nature of Courses</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>17</td>
</tr>
<tr>
<td>As part of theory course(s)</td>
<td>18</td>
</tr>
<tr>
<td>As part of methodology course(s)</td>
<td>10</td>
</tr>
<tr>
<td>Separate course</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
</tr>
<tr>
<td>Total schools responding</td>
<td>64</td>
</tr>
<tr>
<td>Total schools contacted</td>
<td>200</td>
</tr>
</tbody>
</table>

^Sixty-four schools responded, however, three had two courses and one had three courses which included an emphasis on theory construction, making a total of 69 courses.

Of the remaining schools with at least some interest in theory construction, twenty-four had separate distinct courses. In addition, there were eighteen courses in sociological theory (covering such varied topics as theoretical perspectives, various substantive areas, or philosophy) which incorporated one or more units on aspects or techniques of theory construction. In ten other courses, theory construction techniques were included within basically methodological courses.

The results of this survey would seem to be far from definitive in establishing the exact nature of theory construction as utilized throughout the sociology departments in the country, especially since the response rate was not
particularly good. Certainly of the schools that responded, a majority of them (three-fourths) do consider theory construction to be a relevant activity of skill, particularly for Ph.D. candidates. The age-old division of theory and empiricism or methods seems to be reflected in the syllabi of the courses which place theory construction with theory courses or with methodology courses. For the former, the required texts generally include books such as Stinchcombe (1968) or Dubin (1969) which approach theory construction from a generally theoretical side. On the other hand, those placing their theory construction within methodology courses generally have texts such as Blalock's (1969) which is much more mathematically and empirically oriented. Even the courses which are completely separate often take an emphasis which is either one or the other. So how does one proceed to get a grasp of the content of sociological theory construction? One approach would be to attempt to find some types of commonality in these divergent approaches. While it is true that no two instructors will teach even the same course exactly alike, there were a number of texts which tended to be required with some frequency. It was decided then that an examination of the most frequently required texts might give more of an insight into the nature of sociological theory construction than the summary of courses did.
While there certainly was a wide variety of material required in the various courses, seven books seemed to be more consistently required than the others. Table 1.2 lists these seven along with the number of citations each received. The most frequently cited of any of the other texts not included in the table received fewer than four citations and thus were not included in the examination.

Table 1.2: Most frequently cited texts

<table>
<thead>
<tr>
<th>Number of Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zetterberg (1966)</td>
</tr>
<tr>
<td>Stinchcombe (1968)</td>
</tr>
<tr>
<td>Blalock (1969)</td>
</tr>
<tr>
<td>Dubin (1969)</td>
</tr>
<tr>
<td>Kuhn (1960)</td>
</tr>
<tr>
<td>Kaplan (1964)</td>
</tr>
<tr>
<td>Reynolds (1971)</td>
</tr>
</tbody>
</table>

Kuhn's selection reflects the emphasis of some on the nature of explanation, outlining the changing perspectives (paradigms) and the impact each new perspective has had on the prevailing notions of that time. As such it is considerably different from the other books in Table 1.2. It tends to be more concerned with philosophical problems than with construction, per se, although, this is not to say that the nature of explanation is unimportant to the construction of theory. Rather, it tends to deal with only a small
portion of the overall, process of constructing theory.

In many ways, Kaplan is also somewhat different from the others included in the list since he deals discursively with separate issues rather than presenting a method or methods for constructing theory. The issues he deals with include theoretical formulation, measurement, analysis and inference issues and problems but on a somewhat general, philosophical level. He is more interested in the "essence" of an activity rather than in the what or how.

The remaining five books are concerned more with describing necessary steps in a procedure(s) for constructing sociological theory. Zetterberg is by far the most frequently cited text. It also happens to be the first such formal attempt and has been around for over ten years. Certainly, many more instructors are likely to be familiar with it than with some of the more recent works in the field. Stinchcombe's book has also been around for some time. In addition it tends to approach theory construction from a number of different perspectives, demonstrating that the perspective one chooses has important implications for the type of theory one develops as well as the way one constructs that theory. In this sense then, it considers a dimension not discussed by Zetterberg. Of course, the depth of the discussion for any one perspective is far more limited than that found in Zetterberg or most of the others. Another
frequently cited book was Blalock's (1969) text on theory construction. It tends to begin where Zetterberg left off, attempting to demonstrate how axiomatic systems and verbal formulations can be converted into more formal, causal systems which are more mathematically manipulatable. Dubin, on the other hand covers the same area which is stressed by Zetterberg, but has the advantage of time. Thus Dubin's discussion offers a more concise method of theory formulation than that presented in Zetterberg. The last work cited in Table 1.2 is that of Reynolds. Reynolds represents a short, simplified, readable introduction to many of the issues which most theory construction texts discuss on a more complex level.

Although the above comments about the seven books listed in Table 1.2 have been brief, it should be obvious that most of them approach the subject of theory construction from different positions. The basic question, then, "What does (or should) theory construction mean?" remains unanswered. Obviously, one can infer from the previous discussion of the survey and the texts that theory construction can mean different things to different people.

This is not necessarily true, however, if one examines the intent of many of the works just discussed. A common thread can be found among most of these works, particularly in what the various writers feel has been lacking in much of
present-day sociological theory, i.e., the integration of theory and research. This is true regardless of the fact that they seem to disagree on the best way to bring about this integration.

The debate between theory and research (methods or empiricism) has been long and heavy. More and more sociologists, however, seem to be recognizing that it has also been both futile and detrimental. The following comments, taken from introductory statements to some of the texts in Table 1.2, illustrate this desire to bring an end to the wasteful debate, emphasizing instead, the complementarity of the features within each approach.

"As a science, sociology has already bridged the gulf between theory and research; this is true both in principle and in the work of several gifted scholars. The question now is to teach students to run back and forth across this bridge. Our compartmentalized instruction in theory and research might obscure the connection between the two for the students..." Zetterberg (1966:viii).

"...existing sociological theories are rarely clearly enough specified to lend themselves to model building ...I am convinced that (this is partly) a result of a rather unfortunate and false dichotomy between 'theory' and 'research' that has developed within our discipline. I have therefore decided to deal with theory building from the peculiar slant or perspective of applied multivariate analysis..." (Blalock, 1969:vi).

"...it is the intent (of this book) to bridge the gap between theoretical models and empirical research.... The bridge here proposed is viewed as a functioning rather than a static bridge -- it is concerned with the interaction between theory and research. Emphasis is put upon the traffic between theory and research and the existential links that
maintain their constant interaction" (Dubin, 1969:2). Reynolds' book is meant to be short and introductory, of necessity then it was not intended to cover everything. However, in his introduction Reynolds makes the suggestion that his book be used in conjunction with an introduction to research methods:

"In combination with an introduction to research methods...the reader should be provided with a broad and integrated introduction to theory construction and theory testing in empirically based social science" (Reynolds, 1971:vii).

Thus Reynolds recognized the necessity of integrating theoretical formulation not only with testing but also with the real world.

Stinchcombe seems to be somewhat less dedicated to an integrated method of theory construction than the other writers mentioned above; however, even he recognized that the two can not be totally separated. Accordingly in his introduction he states that the sociologists' job...

"(at least at the current stage of development of sociology) is primarily to invent theories, and only secondarily to test them. But theories ought not to be invented in the abstract by conceptual specialists; they should be adequate to the tasks of explanation posed by the data" (Stinchcombe, 1968:3).

Kaplan makes no such introductory statements about the need for integrating methods and theoretical formulation, however a quick glance at the table of contents demonstrates that he
feels the important issues facing social scientists include both. Furthermore, the progression of subject matter moves logically from a general discussion of the "Scientific Method" to concepts, laws, experiment, measurement, statistics, models, theories, explanation and finally values. Certainly there is a feeling of flow in the direction of an integrated procedure, even if that integration is not specifically dealt with.

From these statements one can readily see that the need for an integrated approach to theory construction emphasizing both theoretical and methodological activities is a concern to these writers. In the next section, an attempt will be made to build upon these statements and the comments made earlier by classic theorists to develop a somewhat broad, but concise, statement concerning the general nature of comprehensive theory construction.

A General Description of Theory Construction

Much of the discussion in this chapter has evolved around the issue of theory vs. methods and its relationship to the subject of theory construction. It was shown in the discussion of the survey that this debate has had implications on the manner in which theory construction has been approached in course work. Nevertheless, those who have written sociological theory (i.e., Durkheim, Weber, and Parsons) as well as many who have attempted to instruct
others on how to write sociological theory, seem to agree that the only issue should be how to integrate, not which is more necessary to the development of sociology. The following statement then is an attempt to incorporate these concerns in a concise statement about the nature of sociological theory construction.

Theory construction should be a process whereby both theoretical and methodological activities interact to produce a generalized conception of reality, commonly referred to as a theory or model. As such the end product of this process should be that of explanation as opposed to simply prediction.

This process can be summarized by the diagram in Figure 1.1. Each phase is impinged upon by both theoretical and methodological activities of earlier phases.

---

Figure 1.1: The four phases of a comprehensive approach to theory construction
The four phases depicted in Figure 1.1, (formulation, measurement, analysis, and reformulation), encompass the concerns expressed earlier by Zetterberg, Blalock, Dubin, and the others in their quest for a more integrated conception of theory and methods. This conceptualization is also somewhat similar to that proposed by Reynolds (1971:155) in his description of the explanatory stage of scientific activity which consisted of a continuous cycle of (1) theory construction (used in a more narrow sense than is being used in the present discussion); (2) theory testing, attempts to falsify with empirical research; and (3) theory reformulation, back to (1). The major differences of course arise in the eclectic division of theory and research inherent in Reynolds' discussion as opposed to the integration of these activities emphasized in the present discussion. Another book, not previously mentioned, The Logic of Science in Sociology by Walter Wallace (1971) contains a much more detailed diagram which, essentially is not all that different from the ideas expressed here or in Reynolds. However, as with Reynolds, the Wallace book is insufficient in both coverage and integration of the necessary steps depicted in their diagrams.

One final point should be emphasized. The statement was made earlier that theory construction is a process which
should attempt to produce a generalized conception of reality. Therefore the relationship between the elements or phases of this process and the real world are important to consider. In Figure 1.1 it was noted that each successive phase is impinged on by all previous phases. Figure 1.2 illustrates that each phase must also reflect the real world thing (object, process or relationship) that it is supposed to represent. Thus no one stage is separate from either the earlier phases or the real world.

---

**THEORETICAL CONCEPTION**

**REAL WORLD THING**

**OPERATIONAL VARIABLE**

**STATISTICAL UNIT**

---

Figure 1.2: Relationship of elements in each phase to the real world

The importance of reality throughout the entire process of theory construction is unfortunately something that has often been overlooked or disregarded, especially in the second and third phases. It is easy to lose sight of the real world when confronted with the practical problems of
measurement and analysis, however, this does not negate the fact that precision, no matter how good it may look on paper, can still be irrelevant and insignificant if it bears no relationship to the real world. Thus a study may be a success mathematically (e.g., it may yield statistically significant results) but be a flop substantively (e.g., if the measures were faulty), or any of a wide variety of ways that one loses sight of the real world being studied.

Summary

Theory construction as an area of interest is relatively new, especially compared to more substantive concerns. In a survey of theory construction courses in universities throughout the country, this newness was reflected by a real lack of consensus on the very nature of the term, "theory construction". No two schools agreed on the exact nature of the content of such a course. Even so, a number of works were found to be fairly consistently cited as required texts. Since the number of possible sources is so large and varied as to make a comprehensive analysis unfeasible, it was decided that these frequently cited texts might serve as candidates for an analysis of current theory construction concerns. Utilizing introductory comments of the intent of some of these writers, a brief generalized description of theory construction was presented, thus setting the stage for reviewing these works from the standpoint of the authors' own
aspirations. This then is the main intent of the next chapter: To examine current texts on theory construction (consisting primarily of those listed in Table 1.2) to see whether any or all have been successful in outlining a method of integrated theory construction as derived from statements made by the authors themselves.
CHAPTER TWO:

ASSESSMENT OF CURRENT THEORY CONSTRUCTION TEXTS

A recent survey of the content of theory construction courses throughout the country revealed a great diversity of opinion on what is generally meant by the term, "theory construction", as well as how it should be taught. Nevertheless a number of texts tended to appear rather consistently on the required book lists of the courses surveyed. These were described briefly in the previous chapter and included Zetterberg (1966), Stinchcombe (1968), Blalock (1969), Dubin (1969), Kaplan (1964), Kuhn (1962), and Reynolds (1971).

An examination of these texts showed that most of the authors agreed on the need for the integration of theory and methods. Each affirms this as an explicit or implicit goal of their text. In an attempt to synthesize this generally held goal, then, a broad definition of theory construction was proposed as follows:
a process whereby both theoretical and methodological activities interact to produce a generalized conception of reality, commonly referred to as a theory or model. As such the end product (theory) is a resultant of constant interaction between theory formulation, measurement, analysis and reformulation activities.

Theory construction for the purpose of the present discussion will be viewed in the broad sense of the term, as stated above, rather than in the more limited confines of abstract thought. While it is true that not all theory has been constructed in this broad sense, (Parsons' social system theory has never been completely put to empirical test), the comments of the earlier cited authors of theory construction texts as well as those of the "master theorists", (Durkheim, Weber, and Parsons) point to their concern for the need to utilize such a broad approach.

If one accepts this broad interpretation of theory construction, then two important questions arise in relation to the degree to which present theory construction texts satisfy this definition. The first concerns coverage, or the degree to which the text includes material from the various phases of an integrated approach to theory construction. The second question concerns the degree to which these texts demonstrate the interdependent nature of the activities in
one phase with the activities in other phases. The remainder of this chapter will consist then, of assessing the extent to which frequently cited theory construction texts answer the questions of coverage and integration of theoretical and methodological material.

A Question of Coverage

As a means of approaching the question of coverage, an attempt was made to list some of the most important activities involved in the four phases (theoretical formulation, measurement, analysis, and inference) of a comprehensive approach to theory construction. Table 2.1 was constructed to give a visual perception of the degree to which these texts attempted to deal with various activities of theory construction from the standpoint of material covered in each of the texts. Even from such a simplistic approach as is shown in the table, one can readily see that none of the eight texts listed in the table can claim to describe a comprehensive approach which adequately covers both the theoretical and methodological activities that would be required by the present definition of theory construction.

The list of activities included in Table 2.1 reflects concerns raised by a variety of sources. For instance, theoretical orientation was included here, although it has not been explicitly included in most traditional discussions, because I agree with Sjoberg and Nett (1968) that it has
Table 2.1: Material covered by current theory construction texts

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Z</th>
<th>S</th>
<th>B</th>
<th>D</th>
<th>K</th>
<th>R</th>
<th>A</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical Orientation</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical Background</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Unit of Analysis</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population and Sample</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concepts</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Linkages</td>
<td>***</td>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Propositions</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Systems or Models (arrangement of) Propositions</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>
Table 2.1 (continued)

<table>
<thead>
<tr>
<th>THEORY CONSTRUCTION</th>
<th>Z</th>
<th>S</th>
<th>B</th>
<th>D</th>
<th>K</th>
<th>R</th>
<th>A</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Empirical Indicators</td>
<td>*</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td></td>
</tr>
<tr>
<td>Research Design</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Gathering</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technique</td>
<td>***</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preliminary Analysis</td>
<td>**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis Procedures</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
</tbody>
</table>
important implications for the type of theory that is
developed. Unit of analysis and population are two more
topics which are not often dealt with explicitly. But these
are also part of what Sjoberg and Nett were referring to when
they said assumptions should be made explicit. Several items
in the list have been practically universally recognized as
necessary topics. These include theoretical background;
concepts; linkages (usually included only in so far as they
are a necessary part of a proposition); propositions; and
some approach(es) for arranging them into a theory. There is
a growing demand in the literature (which is also reflected
by Hage, 1972) for theorists to at least indicate the nature
of the indicators or measures they have in mind for their
concepts. What usually occurs in the absence of such
statements is confusion rather than understanding since each
researcher often substitutes a different measure. This
naturally makes comparison of results difficult at best. In
addition to the above activities, which can be considered to
be theoretical formulation activities, the table also
includes a few measurement and analysis topics, in keeping
with the definition of comprehensive theory construction.
[Note: all of the above selections will be discussed in
greater depth in Chapter Three. This brief discussion should
suffice for the present, however, in terms of establishing
the logic behind their inclusion in the table.] Having
discussed the criteria, I shall now turn to the examination of the texts in terms of the topics that each has covered.

Of the six (excluding Kuhn)original texts discussed earlier, Zetterberg and Kaplan seem to be the most inclusive. However, Kaplan is written on more of a philosophical level than on a practical level, and was not really intended to present an approach to theory construction, per se, rather, it was intended to deal with issues that generally confront the social scientist. In the case of Zetterberg, the emphasis is clearly on theoretical concerns since he takes six chapters to cover the six primarily theoretical points but only one to examine most of the methodological points that he includes. Furthermore, the nature of the methodological points he includes is largely limited to research design and preliminary analysis issues such as reliability, validity, and etc.

The Blalock book makes the assumption that the reader has a certain background to build on, such as a familiarity with Zetterberg's axiomatic format, as well as a degree of understanding of mathematical and statistical procedures and terminology. Thus even though he calls his book, Theory Construction, and even though he calls for a more integrated

---

1The Kuhn book was not included since the subject matter in the table was not relevant to the content of the book.
procedure of methods and theory, he chose to limit the context of his book to the narrow concern of converting verbal theories into mathematical equations.

A text which is quite thorough on activities involved in the theoretical formulation phase is Dubin's. He uses a narrow definition of theory construction. Thus for Dubin, the act of theory construction ends with the formulation of hypotheses for testing. All other activities in the table are considered research activities and thus outside the realm of "theory construction" and the content of his book. Even so, he spends the last chapter of his book discussing the relationship between theory and research and research and theory building. It would seem that although Dubin (as well as many other sociological writers) recognizes the need for a more intimate blending of theoretical and methodological activities, he is either unable or not inclined to do it himself, even though he continues to draw attention to the problem. Thus he writes that his book is intended to serve as a bridge for the gap between theoretical models and empirical research (1969:2). However, the bridge never seems to get past the foundation (on the theoretical model side) or at best extends shakily out over the water (his one chapter on the relationship of research to theory) because construction on the other side of the bridge (the methodology side) is never attempted.
Likewise Reynolds seems to have been in pretty much the same frame of mind as Dubin when he wrote his book. He comments on the need for integration but concludes that this integration should be achieved by the use of a methodological text in conjunction with his own, which emphasizes the theoretical issues.

Stinchcombe takes one chapter to discuss what he calls, "the logic of scientific inference". Included in that discussion is a general introduction to the tools (variables, definitions, observation, concepts, and theory) and procedures (multiple tests, alternative theories, crucial experiments, and etc.) which he more or less utilizes to develop example substantive theories within various theoretical perspectives or orientations. The discussion of these tools and procedures is thus necessarily limited in both depth and breadth.

Two additional theory construction texts have been included in Table 2.1. It was felt that their relative newness (Abell, 1971, and Hage, 1972) may have prevented their being included in the survey as required texts. This statement applies to Abell despite the fact that the Reynolds text was published the same year. Abell's text is a much more thorough and complex treatment of theory construction than is the Reynolds' book. Thus, although much of the Abell book seems well worth the extra effort needed for a thorough
understanding of it, it would be less likely to be as readily adopted as the more simplistic Reynolds' text. Like most of the books listed in the table, Abell is very thorough on theoretical formulation issues but fails to include an equal emphasis on methodological ones. Unlike the others, however, Abell does devote one complete chapter to a discussion of analytical techniques (correlation, regression and etc.) and their relationship to the theoretical aspects discussed in the earlier chapters of his book. Certainly Abell's book must be considered at least a start toward developing a more integrated approach to theory construction, even though it too falls short of comprehensive coverage of methodological issues.

The last book included in the table, is the most recent of those being considered here. It was written by Jerald Hage (1972) and deals almost exclusively with theoretical formulation issues. This is in keeping with his notions of what is a theory: "Practically, a theory can be considered as fairly complete if it contains concepts, definitions, statements and linkages" (p. 173). Hage alludes to the distinction between theory and empiricism that has characterized sociology but his reference remains on an essentially formulation level.

"The danger of this concentration (on operational definitions and linkages) is present in the proponents of path analysis. Whole path diagrams give us measurement and testability;
they do not provide meaning and plausibility" (p. 174).

Thus it is not surprising that his book contains nothing on measurement or analysis, even though he recognizes that these do have a bearing on truly "good" theories when he states: "Likewise, theories that cannot be measured and tested are not very desirable" (p. 174).

Of the books listed in Table 2.1 none have incorporated both the theoretical and methodological issues necessary to meet the broad and comprehensive definition of theory construction which was developed earlier. In fact most of the texts considered only the theoretical formulation aspects, even though all apparently recognized the need for integrating both methodological and theoretical activities (see quotes from the various authors in the previous chapter) in the process of constructing truly lasting theory.

A Question of Integration

The question of integration is not so easily discussed as was the previous question of coverage. There is no such criteria which can be listed and checked off as was done with coverage. Instead, the content of each text had to be scanned for possible statement which could be said to advocate or demonstrate the interrelationship of one activity with another. It was found that integration was clearly not a primary goal of any of the texts examined. Rather
statements of an integrating nature were few and far between. A sample of these statements for each text will be included and discussed here.

When one examines the question of integration with respect to Zetterberg (1966), an interesting debate becomes evident, not unlike that alluded to earlier in reference to Dubin. Thus, while the following statements are concerned with the integration of methods and formulation, there is evident an increasingly disenchantment with methodological concerns. It should also be pointed out that these statements on integration do not begin to occur until the sixth and seventh chapters (where methodological issues are first raised). In the first two sample statements below, Zetterberg is emphasizing the need for more integration,

"...definitions (concepts) and indicators should embrace each other in the most intimate way. When we ask how valid the indicators are, we are asking about the intimacy of this embrace" (p. 113).

"The question of validity thus goes to the core of the relation between theory and data...the progress toward validity lies in a continuous adjustment of theorizing to the techniques of research and in a continuous adjustment of techniques of research to theorizing. Unfortunately contemporary sociologists sometimes seem to lack good understanding of this principle. New methods are often developed in a theoretical vacuum sometimes in response to practical needs. And whole conceptual systems are published without the slightest hint as to how their concepts should be translated into research operations" (p. 115).

In the next two statements the disenchantment spoken of
earlier becomes evident.

"The role of measurement in the development of theoretical sociology should not be exaggerated. Comparatively few theoretical advances in other sciences seem to have been inspired by refined measurement techniques and I cannot think of any existing sociological proposition that owes its existence and plausibility primarily to a careful control of the errors of measurement" (p. 126).

"Quantification and statistical analysis are helpful in making the decision (to accept or reject a proposition). However, we should make clear that the use of statistics is no substitute for theorizing" (p. 139).

In the final analysis, one gets the feeling that for Zetterberg the activities of theoretical formulation and methodology remain separate concerns, even though they should remain consonant with one another.

"The task of sociology is to continue with great dedication to sum up its knowledge in the form of theory and to use this theory to gain control over its research efforts" (p. 177).

The major thesis that Stinchcombe (1968), seems to be trying to get across is the necessity for an intimate relation between theory, theoretical (substantive) perspectives, observation (research data), and the real world. For Stinchcombe the theoretical perspective from which one views the world will to a certain extent influence what one sees or cares to
investigate. Likewise it will have an impact on the method of observation as well as on the abstract theory that is developed. Similarly, there should be as close a relationship between what is (real world), what is seen (research data), and the explanation (theory) that is derived. Thus the three integrating statements that were found for Stinchcombe revolved around these four elements. Only two of the statements will be included here.

"It is quite useless to discuss concepts without reference to substantive theory about what goes on in the world, about what causes what. And such substantive theory is merely wind without observation (research) to find out whether it is true or not" (p. 40).

"...measurement is scientific theory in action for a specific purpose...measurement is not only a devise for testing theory, it is a part of the theory...improvement of measurements is usually due to the advance of theory" (p. 43).

The scope of Blalock's book (1969), in terms of coverage of material was shown in the previous section of this chapter to be quite limited. Indeed he makes no pretense of it being otherwise.

"It is not within the scope of this short book to discuss questions of design and analysis..." (p. 25).

Nevertheless, although it is true that he does not pursue these topics directly, he often indicates the interdependence of verbal theory, mathematical models and analytical techniques and problems. It would be difficult to list in its entirety all the statements that Blalock makes of an
integrating nature, since nearly every page or so contains some such reference. Some of these remain on a general level

"It would be highly misleading to suggest that theories are first arrived at by a deductive process and then (Blalock's emphasis) tested. The actual process is much more fluid than this and undoubtedly always involves an inductive effort. One formulates the best theory he can in the light of existing evidence. He then should formalize this theory in order to spell out its implications. These implications are then checked against new data and the theory modified" (p. 8).

or they may deal with the more specific implications of one type of activity on another

"Though it is not within the scope of the present work to consider problems of measurement error, the tendency to introduce dichotomies on the theoretical level would seem to encourage inadequate measurement" (p. 34).

after spending some 153 pages describing the need for greater interdependence of formulation and testing activities, it seems unfortunate that Blalock would end his book on a somewhat equivocal note. He seems to suggest that separation of the two processes is essential even though he still seems to recognize that theory construction (formulation) can not or should not be done without regard for methodological implications.

"This will require that anyone wishing to test (Blalock's emphasis) a general theory must construct an auxiliary theory appropriate for the particular population, measuring instruments, and research design with which he is dealing. Those who are more concerned with the process of theory construction (Blalock's emphasis) should at least suggest the kinds of
operational procedures and possible disturbing influences that should be considered in developing these auxiliary theories. Such a division of labor seems absolutely essential, given the magnitude of the task that lies ahead" (p. 154).

It was stated above, that Dubin (1969), describes his book as a bridge between theory and research. This was questioned on the basis of material covered, even though Dubin offers the explanation that most students are well versed in methods but lacking in theory formulation, (a somewhat questionable assertion in itself). When it comes to the integration of the two, examples can be found, but in the course of his 249 pages, they seem to be the exception rather than the rule. Rather, he seems more committed to bridging the gap between the social scientist and the philosopher. As such, most of the integrating statements that can be found remain on a fairly general level as demonstrated by the following two examples

"Theorizing is an integral part of empirical investigation just as empirical analysis has meaning only by reference to a theory from which it is generated" (p. 7).

"The burden of this section of the chapter (on Research) is that theory building must be empirically rooted" (p.239).

In a similar fashion, Reynolds (1971) apparently depends on readers making the integration themselves by reading a methodology text in conjunction with his own on theory formulation. Thus the few integrating statements that he
makes are on a very general level

"The most important criterion for evaluating the usefulness of any statement for the purpose of science is the degree of correspondence between the statement and the results of empirical research" (p. 115).

There would seem to be some initial attempt to demonstrate the interdependence of formulation and testing in Reynolds' description of his composite approach for developing a scientific body of knowledge. He lists the following steps in such a procedure

1. Exploratory (observation type of research)
2. Descriptive (development of patterns)
3. Explanatory
   (a) Theory construction
   (b) Theory testing, attempts to falsify with empirical research
   (c) Theory reformulation, back to step 3a (pp. 154-155).

However, the entire discussion of this promising strategy is accomplished in the course of only 2 1/2 pages and thus not much is made of it, either in terms of coverage or integration, even though he does make the statement that this strategy demonstrates that theory construction is both difficult and time-consuming and "should not be attempted in a vacuum" (p. 155).

The last book to be examined for integration is that of Kaplan (1964). It should be noted that unlike the others discussed in this section on integration, Kaplan was not intending to write a book on how to construct theories. Rather, his text represents a philosophical examination of
issues involved in scientific activity as it relates to the behavioral scientist. Thus the question of integration is not really of concern to Kaplan here, since he is more interested in the essence of these issues than in discussing a systematic approach.

Summary

In this chapter an attempt was made to assess the degree to which present theory construction texts present a broad and comprehensive approach to theory construction. This was done on the basis of two criteria: (1) coverage of material and (2) integration of that material to demonstrate the interdependence of the various activities. On the question of coverage, none of the texts examined, including two more recent texts, could be said to be comprehensive, at least in terms of methodology. In similar fashion only one (Blalock's) of the texts examined could be said to have paid much attention to demonstrating how one activity depends on or influences other activities.

The general feeling that was derived from this examination was rather ambivalent. On the one hand most of the authors expressed a desire to see theoretical formulation and methodology more integrated, or as Blalock was quoted as saying, described in a more fluid fashion. On the other hand, of the theory construction texts (excluding both Kuhn and Kaplan which were not written for this purpose), only
Zetterberg and Abell made any attempt (incomplete as they were) to incorporate both formulation and methodological concerns. For the most part there was a "let someone else do it notion" or a feeling that the student would somehow do it himself. This is in fact the way most teaching of these two activities has been handled. We teach the student theoretical formulation, methodology, and statistics in compartmentalized courses and then expect him or her to bridge the gap. Perhaps the reason that we are still discussing a gap in theory and methods in sociology is that few have actually been able to achieve this goal on their own.

The second part of this text represents another attempt to fill this void. It will be a step by step presentation of the activities which one usually must consider in a comprehensive approach to theory construction. As such it will include both theoretical formulation and the methodological activities of measurement and analysis. It will be different from the theory construction texts examined here in terms of both coverage of material and a conscious attempt to demonstrate, throughout the procedure, the interdependence of the various activities as well as the way they relate to the real world. A further difference relates to examples. The current text will attempt to work through a theoretical framework and a set of data, in order to
demonstrate in an ongoing process, the various topics as they are discussed here. Few, if any authors have attempted to demonstrate their approach in such detail. It should therefore, represent more of a *practical* guide than most previous attempts. Before undertaking the detailed discussion, however, the last chapter of Part I will attempt to give a brief overview of the proposed approach.

It is hoped that the overview will give the reader a feel for the method as a whole, before examining the detailed descriptions of its parts.
CHAPTER THREE:
AN OVERVIEW OF A COMPREHENSIVE APPROACH
TO SOCIOLOGICAL THEORY CONSTRUCTION

In the first section of this dissertation, attention was
drawn to the importance of integration of theory and method
by several of the "master theorists" in the development of
their theory. It was pointed out that present-day authors of
theory construction texts have also held this, more or less,
as an ideal, albeit for some (such as Blalock, 1969) perhaps
an unattainable one.

The purpose of this chapter will be to elucidate one
such approach which emphasizes the blending or flowing of one
activity into the next. It describes the theory construction
process in terms of four phases: theoretical formulation,
measurement, analysis, and inference or reformulation. Also
unlike much of current thought, the present technique
emphasizes the interdependence of the four phases and the
activities within those phases. Thus the decision to choose
a certain theoretical perspective does have and should have
implications throughout the remainder of the construction
process. Likewise, the concepts one chooses has implications
for measurement decisions, for analysis decisions, and even
the kinds of inferences that can be drawn.
A comprehensive approach to theory construction which describes formulation, measurement, analysis, and inferential activities will require the bringing together of a vast amount of material which is usually discussed separately in theory, methodology and statistics books. The enormity of the task may well have been the major obstacle in preventing others from doing it earlier, particularly those who write wistfully in introductions and prefaces of the need for such an approach but fail to carry through with the task themselves.

The need for such an approach is obviously present, not only for the beginning student but for others as well. The development of a specialization in methodology has led some sociologists to believe that they do not need to know anything about measurement statistics, computer programing and etc. When problems arise they just contact the staff methodologist and let him/her take care of it. The problem with that attitude is that most of the choices made in the theoretical formulation phase have implications for activities within the measurement, analysis, and inferential phases. One should be aware of those implications before the choices are made. Otherwise the theorist may find that the things he/she wants to prove can not be done with the information at hand because prior choices did not allow for collection of appropriate data. Few, if any, methodologists
have missed the experience of having an individual walk into his/her office with a mass of raw data to ask "What do I do now?" Most will say their first reaction was a desire to deposit the mess in the trash. The time to decide what to do is at the beginning when the options are still open.

Admonitions to the contrary (such as the concluding statement by Blalock quoted in the last chapter), the purpose of this chapter will be to develop such a comprehensive approach to theory construction. Although some readers may want a more complex treatment of particular issues, the average reader should be able to gain an adequate understanding of the issues covered as well as a feeling for the way each activity relates to the other activities so that he or she will be able to put the approach to practical use.

Developing the Approach

Having decided to undertake the project of developing an integrated and comprehensive approach to theory construction, the first step was to accumulate the necessary resources. Since extensive volumes have been written on each of the first three phases (i.e. theoretical formulation, measurement and analysis) it was obvious that inclusion of topics must of necessity be selective rather than all embracing. This meant that choices had to be made in terms of
1. what to include
2. the order of inclusion
3. the terminology that should be used

In so far as the choices that were made eliminated other, perhaps equally good choices, the reasons for selecting as I did should be made clear. This will enable the reader to be more fully aware of the basis for this approach and to decide whether those choices meet his/her own priorities and inclinations.

**What to include**

Defining a theory was included as a part of the approach because the term theory has come to mean so many different things to different people. Therefore the theorist should specify what he/she means by the term as a part of establishing what needs to be done.

The stating of assumptions form a part of the definition of theory that will be developed in Chapter Four. Assumptions help to set the boundaries for what and how a phenomena should be studied. They therefore constitute an important, though usually unstated part of that theory. Hage (1972) as well as Sjoberg and Nett (1968) suggest the inclusion of assumptions in the theory construction process.
Theoretical orientation, values, and level of abstraction are perhaps more obviously called assumptions than "how to approach the subject matter." Some might even consider the latter as a measurement question. It was placed with the assumptions because of its boundary setting aspect. The choice of an exploratory, descriptive, or analytical approach is made at the beginning of a study before the concepts, linkages and etc., are selected.

Assumptions set rather general boundaries, the domain setting activities in Chapter Five that were selected, have the effect of setting the specific boundaries of interest. Statement of the problem defines the goals and substantive boundaries. Unit of analysis specifies who or what is being examined and the population/sample spells out the specific group(s) of units. The boundary setting activities were included because they relate to the kinds of inferences that can be made. They are not always included in theory construction texts but Dubin (1969) and Abell (1971) argue for their inclusion.

Concepts and propositions are almost universally included in theoretical formulation. Linkages on the other hand are often dealt with as a part of the proposition. Zetterberg (1966), Dubin (1969), Abell (1971), and Hage (1972) are exceptions. They make more of an effort to discuss linkages separate from propositions. A great deal of
emphasis will be placed in this text on linkages, primarily because tradition rather than meaning has often guided the selection of the linkage terms in propositions.

At one time measures and empirical (operational) hypotheses were considered to be methodological concerns, not theoretical. This no longer seems to be the case, however. More and more theorists have come to agree that they must be an integral part of the theory (Hage, 1972) if others are to fully understand the theorist's intent.

Considerably more selection was involved with regard to the measurement phase than was true of theoretical formulation. Pragmatics demanded that something be cut. It was not possible to do so with the theoretical formulation activities which meant that whatever cuts were made had to come in either the measurement or analysis phases. The problem that was confronted, then, was what could be shortened or deleted without seriously affecting the goals of this text, namely developing and illustrating an integrated and comprehensive approach to theory construction. For it to be comprehensive no phase could be omitted and enough had to be included in each phase so that the feeling of continuity and integration could be maintained.

In terms of the measurement phase it was decided that much of the "how to do it" (e.g., how to construct and administer a questionnaire) could be omitted without too much
damage since previously collected data is now often used and is likely to be utilized increasingly due to ever increasing costs and decreasing funding sources. Even if that assumption is made, however, certain topics must be included, because they have important implications for the types of analyses and inferences that can be made. Thus issues such as study design (important in terms of both analysis and inference and directly related to the selection of a theoretical orientation) sampling, data collection technique and scaling were included in the first part of the measurement chapter.

The second part of the measurement chapter will concern preliminary measurement assessment. This is a fairly recent (within the last 10 years or so) emphasis, although the need has been recognized for a much longer time. It will involve such issues as reliability, validity and measurement error, all of which have a great deal of influence on analyses and inference.

There are, of course, a multitude of analysis procedures that could have been selected. An attempt was made to select some of the more popular parametric procedures that are currently being used. The parametric approaches were selected rather than the nonparametric for several reasons.
1. According to Borgatta and Bohrnstedt (1972) most of the parametric statistics are robust even with ordinal data so that violating the interval assumption does not present too much problem in terms of inference.

2. Parametric statistics offer more information on which to base ones inferences.

3. Nominal and ordinal data can be handled within regression by using dummy variables.

The specific choices of correlation, simple and multiple regression, analysis of variance, and path analyses were made because they seem to be the ones with the most relevance for current sociological research and theory building.

Order of inclusion

The arrangement of topics is for the most part a logical one, activities are arranged into respective phases in the order the theorist would take in his/her construction of theory. Thus theoretical formulation activities take place before measurement activities which also precede analysis and inference. Some topics required consideration at more than one point in the process. Thus sampling is discussed in connection with both domain setting (specifying a theoretical sample as part of specifying the boundaries) in the problem statement chapter, and measurement (with respect to the

---

1A statistical technique is said to be robust if it has the ability to remain stable even though an assumption may have been violated.
extraction of information from the actual sample of respondents).

What terminology should be used?

Terminology is an area which definitely needs some systemization in sociology. Every new text seems to include one or more terms for which the author has substituted his/her own designation. For instance Dubin (1969) rejects the words theory and concepts and substitutes model and unit (respectively). Propositions and hypotheses are not always differentiated in a clear manner. What is empirical to some is considered to be a low level of abstraction to others, (thus extracted data is not empirical to Blumer (1969) who reserves that term for the real world exclusively). Likewise the use of concept, variable, empirical measure, operational measure, and indicator can be confusing.

An effort has been made to explicitly recognize these differences in terminology when they occur. The term(s) which is used in the text will usually represent the one which is most often utilized in the sociological literature, unless otherwise stated. At times, a second term will be included within parentheses when it is felt that it would add to the clarity of the statement, for example: empirical measure (indicator).

These represent some of the problems that were encountered in the development of the comprehensive approach
which will be outlined below and described in detail in Part II. Before I describe this method, however, I would like to comment on the examples that will be utilized in Part II. Since this is supposed to be a comprehensive method, it was thought that the reader would get more of a feel for the method in terms of Kaplan's "logic-in-use" if a single set of actual data with an appropriate theoretical framework was utilized throughout the text to illustrate the various activities. Characteristically in most texts, examples, many of which are often hypothetical are drawn from a variety of sources. It was felt however that this approach would give neither the feeling of continuity and integration that was desired nor an adequate picture of many of the problems that are often encountered with real as opposed to hypothetical data. Since these represent important goals of the text and the method, the single actual data set and framework was selected in preference to the more common procedure.

This represents the framework within which the following approach was developed. I shall give a short overview of the method first and then list the procedures in a step-by-step outline in preparation for the more detailed discussion in Part II.
A Comprehensive Technique for Theory Construction

Because of the great division between theory and methods, the theorist has generally confined himself to the realm of the abstract and the methodologist or empiricist has generally remained on a more "empirical" realm. In many cases neither of these two realms have had much to do with reality or the real world. Herbert Blumer (1969) makes a similar, though stronger, observation. (Note: his use of the term, empiricism, means something different from that used by the usual empiricist. Empiricism means the real world to Blumer, not the data that someone has extracted from that real world.) In describing the activities of most researchers or methodologists, Blumer says

"More serious is their almost universal failure to face the task of outlining the principles of how schemes, problems, data, connections, concepts, and interpretations are to be constructed in the light of the nature of the empirical world under study" (Blumer's emphasis) (p. 27).

Likewise he indicts the majority of social theorists,

"...social theory in general shows grave shortcomings. Its divorcement from the empirical world is glaring. To a preponderant extent it is compartmentalized into a world of its own, inside of which it feeds on itself" (p. 141).

In contrast to the situation which he saw as existing in current sociology, Blumer suggested a relationship more in keeping with the one that is being advanced here, as is clearly seen by this statement
"Theory, inquiry, and empirical fact (real world or reality) are interwoven in a texture of operation, with theory guiding inquiry, inquiry seeking and isolating facts and facts affecting theory. The fruitfulness of their interplay is the means by which an empirical science develops" (p. 141).

What is being said by both Blumer and this text then, is that there must be an intimate relationship between theory, methods and the real world that each of these (theory and methods) is supposed to reflect.

The notion of flow from one type of operation to the next is present in Blumer's admonition, as it is in the statement quoted by Blalock earlier where he states "the actual process is much more fluid than this" (p. 8). This same notion of movement or flow from one activity to another is depicted in the diagram in Figure 3.1.

Before describing the substance of the diagram, however, a few preliminary comments might be in order to emphasize important differences between this approach and others that are currently in use.

Ordinarily measurement and analysis activities, as well as the collected data itself, are usually placed in the empirical realm. However, this is not the case with the present text. Measurement activities before data collection are seen to be an attempt to move from the abstract to the real world whereas data collection (extraction) actually moves back to the abstract realm. Thus unless the data is
the phenomenon, it represents an abstraction, albeit a very low level of abstraction. Another important point about the diagram is that only the fluid aspect of the procedure is emphasized. Just emphasizing this aspect creates a rather complex diagram. It did not seem practical, therefore, to draw lines which would demonstrate interdependence, (e.g., the nature of concepts, indicators, and type of analysis are all dependent on each other), even though this is an important part of the approach. The diagram in Figure 3.1 therefore, should be thought of in terms of the general flow of activities it depicts rather than as a comprehensive illustration of the approach as a whole.

Looking at the diagram more closely, the starting position is the real world. The social scientist uses the process of abstraction to describe aspects about certain phenomena that he or she sees in the real world. The abstract properties or aspects that are defined are called concepts. Relations between properties are observed and are also abstracted out to become propositions (statements relating two or more concepts), which can then be systematized into sets of propositions, such as axiomatic statements, causal models and etc. (various approaches to organizing a theory). At this point many theorists turn to describing empirical indicators of the more abstract concepts. However, as shown in the diagram, one should first
Figure 3.1: Flow diagram of a comprehensive approach to theory construction as it relates to the real world
make a comparison back to the real world. How well does the concept, as defined, describe the real world property it supposedly represents? When one is assured of the correspondence between concept and real property, as well as the correspondence between proposition and real relation, the process of operationalization can begin.

Operationalization is a process of turning concepts into variables "by mapping them into a set of values" (Abell, 1971:38). In other words, it is an attempt to develop a way to measure that concept or property in the real world. As the diagram demonstrates, this is a movement back toward the real world but the measure or indicator can only approximate the real world property (which is represented by the concept), it can never actually be the property. It is important that the relationship between all three of these elements (real world property, concept, and indicator) be as close as possible. Note the three sets of double-arrow slashed lines calling for comparison and forming a sort of triangle. Concepts and indicators should both be compared to the real world property they represent as well as being compared for fit to each other.

Operationalization can also be thought of in terms of turning abstract propositions (or general hypotheses) into empirical (or operational) hypotheses. Empirical hypotheses are on a lower level of abstraction than propositions or
general hypotheses. As is indicated in the diagram, analysis of the data leads directly to a confirmation or rejection of empirical hypotheses, not propositions. Propositions are only "tested" indirectly in the form of inferential conclusions following the testing of the empirical-statistical hypothesis.

The confirmation or rejection of a single proposition may or may not affect the original theory, depending upon how vital it was to the system. Quite often however, results of the testing lead to some reformulation of the propositions and eventually to the theory. The process of reformulation sometimes includes the retesting of revised hypotheses with the same set of data before stating the revised propositions or theory. Caution must be exercised, however, in revising a theory on the basis of test results from one set of sample data. There may be a problem with the sample, the measures, the specific relationship tested in statistical hypotheses, or a number of other things which may "mask" the true relationship in the real world. The final theory should be relevant to the population of phenomena in the real world having similar characteristics.

The model in the diagram in Figure 3.1 is essentially the method which will be proposed here as an alternative approach to a more comprehensive construction of theory. This particular method strives to avoid the criticism leveled
by Blumer against both social theorists and methodologists in its constant reference back to the real world. It also differs from the other theory construction texts discussed in chapter two in terms of both content (since it will include measurement and analysis, in addition to theoretical formulation and inference) and in its explicit integration of that content. Because of the complexity of the project at hand (to describe a comprehensive approach to theory construction) an attempt was made to keep the diagram as brief and simple as possible so that the reader might more easily get a grasp of the approach as a whole. It would also seem worthwhile, however, to have a more detailed description of the approach. The following outline has been included for that purpose.

I. Theoretical Formulation
   A. Definition of theory
   B. Stating assumptions
      1. Theoretical orientation
      2. Level of abstraction
      3. How to approach the subject matter
      4. Impact of values
   C. Statement of the problem
      1. Theoretical background
      2. Importance of the problem
      3. Unit of analysis
4. Population at risk
5. Objectives
6. How to meet objectives

D. Steps in formulating a theory
   1. Specifying concepts and measures
   2. Linkages
   3. Propositions and their systematization

II. Measurement
   A. Considerations prior to data collection
      1. Research design
      2. Study design
      3. Sampling
      4. Data collection technique
      5. Scaling
   B. Measurement assessment
      1. Functional unity
      2. Reliability and validity
      3. Measurement error
      4. Examination of descriptive statistics

III. Analysis
   by stating specific statistical hypotheses.
   A. Relating propositions to analysis techniques
   B. Multiple regression
   C. Path analysis

IV. Inference
Summary

In response to a much observed need for an integrated, comprehensive approach to theory construction, this chapter has attempted to demonstrate how one such approach was developed. A brief overview and outline of the approach was also given.

A comprehensive approach which would satisfy the working definition of theory construction that was developed in Chapter One, would require a synthesis of theoretical formulation, measurement and analysis activities. In attempting to achieve that goal it was found that selective inclusion, rather than an all encompassing one, would be a more appropriate standard for a dissertation. The approach that was outlined above is limited, to the extent, then, that some features of measurement and analysis were not examined. However, an attempt was made to include those aspects of measurement that are necessary for bridging the theoretical formulation activities with the selected analysis activities when the theorist is using data that has already been collected. This does not mean that the approach can not be used when data is collected first hand. Rather, the method should still prove useful, although the beginning theorist would probably need to supplement the measurement chapter with further reading.
This chapter represents only a brief introduction to the activities which are encompassed by the proposed approach. Part II of this text will be concerned with a detailed examination of these activities and will be presented in four sections, each representing one of the four phases of the comprehensive approach to theory construction that was developed above.
PART II:
A DETAILED DISCUSSION OF AN INTEGRATED COMPREHENSIVE APPROACH TO THEORY CONSTRUCTION
THEORETICAL FORMULATION

It was demonstrated in Part I that the term *theory* construction has a variety of meanings and interpretations. In some cases, it has been more or less equated to what will be discussed in this section under the heading, Theoretical Formulation. However, this approach has frequently led to the construction of "theory" which according to some, is more suitable to the realm of philosophy or fiction than it is in providing sociological understanding of real phenomenon (Blumer, 1969). In the present text, theory construction has taken on a much broader meaning, so that theoretical formulation represents only one of several interrelated phases that form a comprehensive approach to theory construction.

The major goal of this section is to describe the various activities involved in the theoretical formulation phase. These activities will be discussed in terms of two categories: boundary-setting activities and the steps involved in formulating a theory. The latter is what has commonly been equated by other writers with theory construction. The former is primarily concerned with the thinking processes that one usually goes through in deciding upon what to examine (in terms of general assumptions as well as what the specific boundaries will be for that examination,
e.g., stating the problem, selecting a unit of analysis and designating a population).

But a theory which ends at the formulation stage is still largely supposition, since it has neither been supported nor refuted in the real world. Thus before any substantial conclusions can be drawn about the validity of the propositions in a theory, empirical hypotheses, derived from those propositions must be put to the test in the real world. This requires that the theoretical formulation phase be followed by both measurement and analysis phases. Thus, although the major intent of this section is to discuss the activities involved in theoretical formulation, an attempt will also be made to demonstrate (1) how formulation decisions often guide measurement and analysis decisions, and in turn, (2) how the final theory (as arrived at in the inferential phase) is ultimately dependent upon the activities and outcomes of each of the previous phases. In other words an attempt will be made throughout this section, as well as the others, to demonstrate the kind of interdependence of activities from the various phases that will lead to the development of an integrated, comprehensive theory.

But what is a "comprehensive" theory? as mentioned in an earlier chapter, it will probably be difficult for sociologists to agree on a method of constructing theory as
long as they continue to disagree as to what is a theory. Homans has further suggested that sociologists stop using the term theory with students until we have taught them what it means. Therefore the first chapter in this section (Chapter Four) will be concerned with establishing a comprehensive definition of sociological theory which will then be assumed and referred back to, throughout the remainder of the text.
CHAPTER FOUR:
WHAT IS A THEORY

The term "theory" has been applied to a wide variety of phenomena. Denzin has in fact referred to theory as "the most voiced concept in modern sociology" (1970: 34). As one might expect, this has led to confusion, not clarification or understanding, and has further led some to suggest that theory is also the most "overworked" concept. Apparently everyone wants to be a "theorist" so they apply the label, "theory," to whatever they happen to be writing at that moment. Both the Zetterberg and Merton discussions below illustrate this rather haphazard usage with which the term has been applied.

According to Zetterberg (1966), the variety of works to which the term has commonly been applied can be roughly classified under four headings. These are as follows:

A. Sociological Writings of Older Vintage
more properly described as sociological classics. Parsons' Theories of Society (1961), an example of this type, is actually an anthology of classical passages of sociological literature rather than explicit "theories" about society.

B. Criticisms or commentaries on sociological writings Becker and Boskoff (eds.), Modern Sociological Theory: In Continuity and Change (1957) represents an example of this type of anthology. These are usually written from a
historical perspective, tracing continuities from one era to the next in the development of sociological thought.

C. Taxonomies—providing an orderly schema for classification and description of something social. Parsons and Shils' (eds.), Toward a General Theory of Action (1959) was cited by Zetterberg as an example of this type. It should be noted that Turner (1974) has said that although Parsons' "theory" is currently only at this level, it is supposed to be followed later by something more in keeping with Zetterberg's fourth category below.

D. A set of systematically related propositions leading to explanation and verificational studies. George Romans' The Human Group (1950) was cited as an example of this type of activity.

All of these various types of works have been commonly labeled as sociological theory, however, Zetterberg maintains that the term, "theory" should be reserved for the last type of endeavor.

"Only at this stage does it make sense to speak of 'testing a theory,' 'derivation,' and—most important of all—'explanation'" (1966:28).

He notes further that distress over the "corruption of the concept of social theory" has led some writers to substitute the term "model" for theory (p. 29) leading in some cases to even greater confusion.

Merton (1967:139-153) has also commented on the wide variety of activities to which the term sociological theory has been applied. He lists six general categories of
activities which often bear the label "theory".

a. Methodology
b. General sociological orientations
c. Analysis of sociological concepts (taxonomies)
d. Post factum sociological interpretations
e. Empirical generalizations in sociology
f. Sociological theory

His definition of sociological theory refers to "logically interconnected sets of propositions from which empirical uniformities can be derived" (1967:39). Obviously, the first three categories of activities above do not fit his definition of theory, since they are not sets of substantive propositions. Methodology is really the logic of scientific procedure, and is therefore not content bound. General sociological orientations chiefly function to provide a general context for inquiry and the development of theory. Taxonomies, which deal with the organization and clarification of definitions of concepts, are helpful for the development of clearly stated theory but are not in themselves theory.

In contrast to the activities discussed above, all of the last three activities included in Merton's list above, are generally stated in the form of hypotheses. Nevertheless, only the last type is considered by Merton to be theory.
Post factum interpretations arise after the fact, (i.e. after the observations have been made). As such this type of interpretation is tailor-made to fit the data, so that what one has is an illustration of a plausible explanation and not confirmation (compelling evidence) in support of a previously developed theory. In the case of empirical generalizations, they constitute "isolated propositions summarizing observed uniformities of relationships between two or more variables" (1968:149). Therefore, while disjointed empirical generalizations can be said to provide the raw material for the development of sociological theory, they can not, in that form, be considered to be sociological theory.

According to Merton then, one must have a set of interrelated propositions which have evolved from a stated rationale of why they are so interrelated (see the term logically interconnected in the definition above). A further stipulation is that the propositions should not be so abstract or remote as to make their linkages with reality difficult to comprehend (as is often the case with general theory).

A third noted sociologist, George Homans, (1967) comments that before we begin to discuss substantive sociological theory with students, we have to teach them what a theory is (and is not). A theory, according to Homans, is not a set of concepts or definitions of concepts, nor is it a
proposition stating the existence of a relationship. It is rather a set of propositions taking the form of a deductive system having the following characteristics:

a. contains both descriptive and operative concepts

b. states relationships between those concepts in the form of propositions

c. such that they form a deductive system where lower order (empirical generalizations) are derivable from higher order (general or abstract) propositions.

A further requirement states that good theory involves explanation and thus provides grounds for prediction. Thus at least some of the concepts of a good theory must be contingent (measurable) so that the truth or falsity of the theory, or its derived propositions, are testable.

Building on Homans' conception of theory, as outlined above, Denzin (1970) offered a definition which emphasized the explanatory and predictive power of a theory.

"...a set of propositions that are interrelated in an ordered fashion such that some may be deducible from others thus permitting an explanation to be developed for the phenomenon under consideration. A theory (then) is a set of propositions that furnish an explanation by means of a deductive system. Theory is explanation...when a deductive system provides explanation, it also permits prediction" (pp. 34-35).

Summarizing these various definitions, then, a theory could be said to be a set of logically interrelated
propositions which lead to the explanation and prediction of social phenomena. However, according to Sjoberg and Nett (1968), this definition is incomplete. Theory is more than explanation (and/or prediction); more than propositions; a theory must include an explicit statement of the underlying assumptions on which it rests. It will be demonstrated in the next chapter that there is no single overarching theoretical orientation or perspective in sociology. Instead, there are many, each with its own set of assumptions about man, society, and data which influence what, as well as what is not, examined by its adherents. As Ritzer (1975) notes below, this means that the selection of an orientation (paradigm) has important implications for the way reality is perceived and theorized about.

"...although adherents of each paradigm (orientation) claim all of sociology, their approach seems best suited to a particular facet of social reality. The social behaviorists appear to be best able to deal with behavior and contingencies of reinforcement. The social definitionists (social phychologists) apparently cope best with the social construction of reality and the ensuing action. Finally, the social factists (social organizational approaches) seem to work best with social structures and social institutions. I have emphasized the words appear, apparently and seem in the preceding sentences because it is my view that none of the paradigms is able to cope well with its facet of reality. In fact, no aspect of social reality can be adequately explained without drawing on insights from all of the paradigms" (Ritzer, 1975:211).
Therefore, since there is no one set of assumptions, and since the assumptions that are made tend to emphasize only portions of reality (conflict, cohesion, exchange relationships, and etc.), Sjoberg and Nett stress that to be complete, a theory must clearly specify the orientation (and/or the assumptions) from whence it evolved.

If this requirement is added to the summary on the previous page, a theory would then be said to consist of:

A set of logically interrelated propositions based on an explicit set of assumptions which lead to explanation and prediction about social phenomena.

Hage (1972) has suggested that even this is not really complete. In addition to the set of theoretical statements, the theorist must include definitions and linkages at both the theoretical and operational levels as a way of clearly delineating the theorist's conception of the social phenomenon about which he is theorizing. This would not be necessary if only one definition and measure existed for each concept. However, this is not the case, there are often a variety of definitions and empirical measures available, each with a slightly different meaning than the others. This is particularly important in terms of testing a theory. As Hage observes below, what is empirically tested is often quite different from what is described theoretically, which can also be quite different from what the theorist originally intended—all because of the confusion that results when the
theorist fails to specifically state his intended meanings, measures and linkages.

"We sometimes find it necessary to change both theoretical concepts and statements when we have operationalized one of the concepts. Indeed one is continually struck by articles in the journals that appear to be testing hypotheses very different from those they say they are testing because their definitions are so different" (1972:175).

If the definition of theory being developed here is expanded to include definitions and linkages, as suggested by Hage, a rather comprehensive definition of sociological theory results.

AN INTEGRATED, COMPREHENSIVE SOCIOLOGICAL THEORY CONSISTS OF: A set of logically interrelated propositions; which have been evolved from clearly stated definitions and linkages, at both theoretical and operational (empirical) levels; as well as an explicit statement of underlying assumptions; all of which lead to explanation, and some degree of prediction, about social phenomena.

This then will be the definition of sociological theory that will be assumed and utilized throughout the remainder of the text.
Implications

It should be clear from the comprehensive definition above, that activities included in all the phases of theory construction (formulation, measurement, analysis, and inference) must be involved, and intimately related, in order to produce such a theory. Each part of the definition, then, relates not only to the generally ascribed activities in the formulation phase, but also to activities from the other phases as well. This can be shown graphically by means of the breakdown on the following page.

The first section of the definition can be said to relate to all four of the phases in so far as the final set of propositions should reflect the product of the total theory construction process. This means that it will probably be somewhat different from the originally formulated set of propositions (i.e., the set proposed in the formulation phase). It should reflect insights gained from the other phases. This is recognized in part by the next statement in the definition.

Formulation, measurement and analysis phases are all involved in accomplishing the second part of the definition, (i.e. the stating of definitions and linkages at both the theoretical and operational levels). The formulation phase defines them, the measurement phase acts upon the operational definitions and linkages (thus providing data), and the
Table 4.1: Demonstrating the phases involved in accomplishing the various aspects included in the definition of theory

<table>
<thead>
<tr>
<th>Part of the definition</th>
<th>Involves activities from</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &quot;Set of logically interrelated propositions&quot;</td>
<td>Formulation, analysis and inferential phases</td>
</tr>
<tr>
<td>2. &quot;Evolved from clearly stated definitions and linkages at both the theoretical and operational (empirical) levels&quot;</td>
<td>Formulation, measurement and analysis phases</td>
</tr>
<tr>
<td>3. &quot;Explicit statement of underlying assumptions&quot;</td>
<td>Involves boundary-setting activities from formulation phase which affect some measurement, analysis and inference decisions</td>
</tr>
<tr>
<td>4. &quot;leading to explanation and some degree of prediction.&quot;</td>
<td>Formulation, measurement Analysis, and inference phases.</td>
</tr>
</tbody>
</table>
analytical phase serves to relate the empirical level back to the theoretical level, since statistical tests are actually written in terms of the population or theoretical level.

The third portion of the definition of sociological theory refers to stating one's assumptions. This involves boundary-setting activities which will be described in the next two chapters.

The fourth and final section of the definition of theory states the goals expected from a theory (i.e. explanation, which refers to the assessment of relationships, and at least some degree of prediction, or "fit", between the theory and the empirical situation to which it has been applied). A theory is just supposition unless it can accomplish (to some degree at least) one or more of these goals. Obviously then this criteria requires the close integration of activities from all the phases. If one uses measures that are not closely linked to the definitions prescribed in the formulation phase, the analytical tests, regardless of their outcome, will have little meaning.

This rather extended discussion has been included in order to demonstrate further the necessity for using an integrated and comprehensive approach to theory construction.
Summary

This chapter has been concerned with developing a comprehensive definition of sociological theory. Definitions of various well-known writers in the field were considered, each emphasizing different aspects of what they called "theory." A composite definition, reflecting those various emphases was then formed. Since this composite includes many diverse elements, such as, definitions of concepts and measures, linkages, explicitly stated assumptions and a requirement that it provide some degree of explanation and prediction, in addition to the usual stipulation of a set of interrelated propositions; it therefore represents the kind of comprehensive definition of sociological theory referred to in the opening statement to this section.

Having established an explicit definition for a sociological theory, the remainder of this section will be concerned with examining various portions of that definition. These include stating one's assumptions, outlining the specific boundaries of the domain of interest (stating the nature of the problem, the unit of analysis, and the population), specifying concepts and measures (including their definitions) and discussing linkages, propositions and their systematization. While these constitute various aspects of theoretical formulation, they do not exhaust the list of activities within theory construction, at least not
the theory construction process that was defined in Chapter One. The process of theory construction as defined earlier, requires the testing of those propositions and systems, as well as their development. Such testing requires not only theoretical formulation but also measurement of data and analytical testing. Once again the discussion has demonstrated a need for the integration and interdependence of the various types of theory construction activities. Furthermore, this emphasis should become more and more apparent as the reader progresses through this text, beginning with the next chapter on assumptions.
CHAPTER FIVE:
THE STATING OF ASSUMPTIONS

This aspect of theoretical formulation is concerned with the stating of assumptions that result in narrowing the scope of interest. An assumption is said to be a statement, the veracity of which is not questioned. Theories are generally tested at least indirectly, but assumptions rarely are. Furthermore, while assumptions are always made, they are not always made in an explicit fashion. This chapter will be concerned with demonstrating the importance of making those assumptions explicit since the parameters of our assumptions set the boundaries for all future decisions.

Probably the most comprehensive, and thus important, assumption involves the selection of a theoretical orientation. Each orientation, and the perspectives within those orientations, carry with them sets of assumptions. Furthermore, adherents of orientations tend to approach the other general assumptions in this chapter from similar points of view. Therefore, while this chapter will be concerned with discussing the following four general assumptions of

1. Theoretical Orientation
2. Level of Abstraction
3. Type of Approach
4. The place of Values
it will hopefully become evident that these assumptions are all interrelated.

The four major topics within this chapter have been designated as assumptions. As such they are usually selected on the basis of personal preference rather than as a result of having met some stringent criteria. Furthermore, since none of the possible choices has ever been "proven" to be the only viable one, it is important for the theorist to at least make his/her choices known, as well as the reasons behind their selection. All too often, the assumption of a theoretical orientation by a theorist has opened the way for criticism from others out of a different, but equally, assumed orientation. Making an explicit statement concerning the assumed orientation would seem, then, to represent the best way to avoid the confusion and unproductive arguments that generally ensue when a theory is developed under one set of assumptions but evaluated on the basis of another.

Theoretical Orientation

Theoretical orientations are generally examined in the confines of substantive theory courses. Little effort is usually made to relate these perspectives to either theoretical formulation or methods. Unlike other sciences, sociology does not have a single overarching theoretical orientation to guide its adherents, although most discussions concerning theoretical orientations generally evolve into a
debate as to which one is the proper approach for sociology. It would seem, regardless of protestations to the contrary by adherents of the various orientations, that each approach tends to emphasize a different facet of social reality rather than an overall or comprehensive one. Selection, therefore, of a theoretical orientation has important implications for the decision concerning "what to study". An examination of the following definition of theoretical orientation should make these implications more readily apparent.

---

A theoretical orientation is

"...a fundamental image of the subject matter within a science. It serves to define what should be studied, what questions should be asked, how they should be asked, and what rules should be followed in interpreting the answers obtained.... It subsumes, defines, and interrelates the exemplars, theories, and methods and instruments that exist within it" (George Ritzer, 1974:7).¹

---

George Ritzer (1975) concurs with the statement above that no current sociological paradigm (i.e., no single

¹Note that Ritzer actually used the terminology of paradigm for what has been referred to here as theoretical orientation. The way in which he uses the term is synonymous with theoretical orientation as I will use it. This is not always the case and therefore, to avoid confusion, the more common term, theoretical orientation will be used here.
orientation or perspective) can claim enough support throughout sociology to be considered the sociological paradigm. Instead he has noted that there are three basic orientations, (or paradigms as he refers to them), within which one may categorize the many perspectives.¹ He further emphasizes the fact that orientations (and the perspectives within those orientations) are generally limited in both the subject matter they examine and the method which is characteristically employed. While there have been a few attempts to integrate some of the perspectives, those attempts have often been resisted by the adherents of the individual perspectives, each apparently concerned about infringement upon the uniqueness of their own particular approach. [Note that this uniqueness is usually seen by adherents as meaning that theirs is the only perspective with the correct approach for examining social reality.] As the next two figures illustrate, however

"...no aspect of social reality can be adequately explained without drawing on insights from all of the orientations (author's

¹The term perspective will be used here to denote approaches that fall within an orientation (or a paradigm as used by Ritzer). Thus, structural-functionalism and conflict will be considered as two perspectives within the social factist orientation.
emphasis" (Ritzer 1975:211).

Figure 5.1, An Ordering of Theoretical Perspectives in Sociology, attempts to graphically display the differential emphases of the various perspectives and orientations. This is done in terms of the position each takes with regard to (a) the structure of reality; (b) the structuring of events (units); and (c) the preferred methodology for collecting data.

Structure of Reality refers to the way the perspective examines social reality. A number of perspectives examine social reality in terms of a static, frozen cross-section. In addition, examination of reality can be further bound by the notion of a closed environment. In other words, there is no linkage between a system, event, or etc., and its environment. A closed system is a self contained unit. As such, a closed model contains all the relevant variables---one need not be concerned with outside influences or things that are occurring in the environment. Structural functionalism has often assumed this position.

A second approach to the way structure is organized is
Figure 5.2: An ordering of Theoretical Perspectives in Sociology

<table>
<thead>
<tr>
<th>Structuring of Events (Units)</th>
<th>Structuring of Reality</th>
<th>Orientation and Preferred Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Group or Individual (as an Organic Whole)</td>
<td>Open</td>
<td>Conflict</td>
</tr>
<tr>
<td></td>
<td>Procedural</td>
<td>General Systems</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Structuralists (Etzioni)</td>
</tr>
<tr>
<td>Individual or Small Group</td>
<td>Closed</td>
<td>Parsons</td>
</tr>
<tr>
<td></td>
<td>Static</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual</td>
<td>Folks Theory</td>
<td>Social Behaviorist (Psychological)</td>
</tr>
<tr>
<td>Psychological</td>
<td>(Structural</td>
<td></td>
</tr>
<tr>
<td>Feats</td>
<td>Ethnography</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethnography</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Partially Adapted from Sociology: A Multiple Paradigm Science by George Fitzger (1975).
the open, processual approach. The focus here is on an open linkage with the environment. People, organizations, systems and etc. are all affected by their environment; so, to understand social reality, one must examine the linkage between a system and its environment. Furthermore, reality is an ongoing ever changing structure that should only be examined in terms of ongoing processes.

Sociological perspectives also differ in terms of the unit of analysis which is thought to be most basic to an examination of social reality. Some emphasize a macro approach. Social reality is viewed as a system, having interdependent subsystems which are different from the system as a whole, (i.e., the whole is greater than the sum of its parts). This view tends to emphasize the examination of aggregates (groups, organizations, communities and etc.). Other perspectives take an opposite view. Systems, groups, and etc., are actually made up of individuals. It is their actions which are important. Social reality therefore, as it pertains to groups or aggregates of individuals, are understandable by means of examining the individuals (and relations between those individuals) that make up the group or aggregate (i.e., the whole is equal to the sum of its parts). As Figure 5.1 indicates under the label, "Structuring of events (units), the Social Factists tend to subscribe to the former view while the Social Definitionists
operate on the latter. There is a third approach which also takes the individual, or at least a part of the individual, as the basic unit. These are the social behaviorists who assume that everything can be reduced down to the mind of the individual. This approach is basically concerned with an attempt to reduce all social behavior down to its "psychological roots".

Finally, sociological orientations differ in terms of methodology. Ritzer has made the comment that any methods used in data gathering can be adopted within each of the various sociological orientations, however each orientation carries with it a preference for one over the others. Thus the Social Facists generally prefer the questionnaire or interview, whereas the Social Definitionists (Social Psychologists) usually prefer direct observation.

The full implication for theory construction of these differential approaches to the structure of reality and events (units) has perhaps, not yet been made clear. In fact, it is the position that each perspective takes concerning these basic questions that leads to the acceptance of other important assumptions. For instance, the Interactionists assume that man (the individual) is the basic unit of analysis and that he/she is by nature an active creator (vs. the product of norms and roles, as seen in Structural Functionalism). Working on the basis of this
The image contains a table and a diagram with the following content:

**Figure 5.2 Differential emphases to basic assumptions**

<table>
<thead>
<tr>
<th>ASSUMPTION</th>
<th>NATURE OF MAN</th>
<th>HOW SOCIAL ORDER ARISES</th>
<th>HOW SOCIAL ORDER IS MAINTAINED</th>
<th>HOW SOCIAL ORDER IS CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOCIAL STRUCTURAL</td>
<td>product of normative</td>
<td>continues as change</td>
<td>represents disturbance</td>
<td>is slow adjustment</td>
</tr>
<tr>
<td>FUNCTIONAL</td>
<td>roles and social structure to meet needs</td>
<td>long as needs are being met</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PARSONS</td>
<td>decision-making is shared values by ideas (norms) and situations</td>
<td>adapted and integrated</td>
<td>incremental (intrasystem changes)</td>
<td>radical (extrasytem changes)</td>
</tr>
<tr>
<td>CONFLICT</td>
<td>scarcity of resources of power into authority relations</td>
<td>inequality and new conflicts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYMBOLIC INTERACTION</td>
<td>active meanings and meanings</td>
<td>remain the new actors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXCHANGE</td>
<td>reward relations</td>
<td>marginal utility and distributive justice</td>
<td>new actors, new rewards, or costs</td>
<td>exceed profits</td>
</tr>
<tr>
<td>BEHAVIORAL SOCIOLOGY</td>
<td>positive response to external stimuli</td>
<td>reinforcement behavior changes</td>
<td>pattern are established changes</td>
<td></td>
</tr>
</tbody>
</table>

This table and diagram outline various perspectives on social structure and change, focusing on different aspects such as nature of man, social order, social change, conflict, symbolic interaction, exchange, and behavior sociology.
assumption, then, the Interactionist examines social phenomena in terms of the shared meanings and expectations of individuals as they (the individuals) define them. Structural Functionalists in keeping with their assumptions about the nature of man and social order, emphasize normative structures and how they meet the needs of the system or organism that is being examined. These and other examples of differential assumptions made by the various perspectives have been illustrated in Figure 5.2.

The differential emphases of the various perspectives has had its impact upon theory construction in sociology as Figures 5.1 and 5.2 attempt to summarize. Scott McNall (1969) describes succinctly the nature of these implications for theory construction, particularly how the selection of a specific theoretical perspective impinges upon the kinds of questions that are pursued by a theorist. He turns to a descriptive example of the neighborhood tavern to illustrate his comments.

McNall's example compares three perspectives, structural Functionalism, Conflict, and Symbolic Interactionism. Inherent in his discussion is the notion that each of these perspective differ as to the basic assumptions they make about reality. These differing assumptions often lead adherents of the various perspectives to examine the same social situation from differing points of view. Thus
according to McNall, the questions one asks about a social situation (such as those within his example of the neighborhood tavern) will differ according to the particular viewpoint stressed by one's perspective. Typical examples of the types of questions discussed by McNall are listed on the next page.

These examples have been taken from McNall and are meant to illustrate the differential emphases alluded to earlier. The first two perspectives are both included in Figure 5.1 under what Ritzer has called the Social Factist Orientation. Note that although the questions for these two differ in terms of their respective emphases on integration versus conflict, they are nevertheless similar in terms of the emphasis placed on roles and other such topics (i.e. function of the bar?...is the bar a source of conflict?) which are external to the individuals. Now compare their emphases to those of interest to the Symbolic Interactionist, (who has been placed under a different orientation, namely, the Social Definitionist). For the latter, the emphasis is no longer on structure but rather on meaning, i.e., meaning to the individual. It is the meaning the individual places upon external things that is of importance to the interactionist, rather than the external things impinging on the individual.
******QUESTIONS CHARACTERISTIC OF A FUNCTIONALIST******

1. What is the function of the neighborhood bar?

2. Why do people come...to meet others...as an extended primary group?

3. How does the bar serve to integrate their lives?

4. What is the social function of alcohol? e.g., Does it have a therapeutic function? Does the act of drinking serve the function of safety valve for potential psychotics?

******************************************************

******QUESTIONS CHARACTERISTIC OF A CONFLICT APPROACH******

1. What conflict does this bar represent?

2. Does the drinking pattern of specific individuals lead to a conflict with other roles, such as employment or marital life?

3. Is the bar itself a source of conflict in the community because of loud music, etc.?

******************************************************

****QUESTIONS CHARACTERISTIC OF SYMBOLIC INTERACTIONISM****

1. What are the implications of (meanings attached to) the drinking act itself?

2. What does the symbol world of the tavern consist of?

3. How do the regulars identify a newcomer?

4. How does the etiquette (shared expectation) of drinking in the neighborhood bar differ from those in a nightclub?

******************************************************

(Scott McNall, 1969)
It is clear from this example that each of these perspectives tends to lead the inquirer to ask different questions, to discover different "facts" about the same social situation. All may offer important insights about that social situation, but some aspects will be ignored or overlooked, depending upon the specific perspective that is utilized in the examination. Thus the Functionalist will look for integrative aspects in the social structure, tending to ignore or at least de-emphasize, the conflict generating aspects as well as interpretive meanings. Likewise, the conflict theorist tends to ignore those aspects which are not directly connected to the generation and or/resolution of conflict. Finally the symbolic interactionist often neglects to consider the real implications that structural factors contribute in addition to the interpretive meanings that he/she stresses. The reality then that a theorist "sees," and attempts to relate to, will inevitably reflect the perspective from which he/she approaches that reality.

Implications

The selection of a theoretical perspective is important to theory construction because it tends to limit one's view of reality and thus one's approach to that reality in terms of limiting the type of subject matter that one pursues. For example conflict theorists have largely confined themselves to examining conflict situations in large groups,
institutions and etc. The information that might be gained by examining individual motivation is generally ignored. Clearly then, the selection of a conflict orientation determines, to some extent, not only what is studied but also what is not. Likewise, the orientation affects the concepts and hypotheses that one examines (the structural Functionalist concentrates on functions or integrative structures, while the symbolic Interactionist looks at symbols and motivations).

The choice of theoretical orientation also has implications for the other phases of theory construction as well as for the theoretical orientation phase. Each orientation has a preferred methodology, the selection of a perspective within an orientation therefore means that the preferred method of data collection (in preference to the others) is usually the one that is selected. Thus utilization of a Symbolic Interactionist perspective generally means that the theorists will employ some type of observation technique to collect qualitative measures of concepts. If the measures are indeed qualitative, then the types of analysis procedures (if any) which can be utilized will generally be different from those that would have been used with quantitative data.

The implications of one's theoretical orientation have usually gone unrecognized or ignored. Perhaps this is
because adherents of particular perspectives tend to look upon their perspective as the only legitimate approach for explaining social reality. According to George Ritzer (1975) this same attitude is often taken by adherents of the various perspectives with regard to the level of abstraction that is thought to be proper for a theory. Thus the Social Factists are said to be more interested in developing grand, abstract theories than the Social Definitionists who generally prefer to stay on a much lower level of abstraction.

Just as the selection of a theoretical orientation was said to be important, selection of a level of abstraction also has serious implications for the theory that is developed. Nevertheless, little is generally stated by theorists with regard to the chosen level of abstraction, at least in an explicit fashion. Consequently, level of abstraction generally becomes yet another unstated assumption, which may or may not be clear to others and which may therefore result in faulty interpretations by those who mistakenly attribute a level that is different from the theorist's.

**Level of Abstraction**

The level of abstraction is important to the selection of "what to study" in so far as it relates to the question, "what is the theorist trying to explain?" If the object of the theorist is to develop a theory which is generalizable to
all classes of social phenomena then one must work on a highly abstract level (grand theory). This means that the theorist is interested in explaining the relationship between abstract concepts [which may or may not have direct correspondence to properties (referents) in the real world].

| Specific to empirical observation of single event | Involves some abstraction but specific linkage to real world counterpart (referent) is maintained | Highly abstract, reconstructed picture of reality |

| EMPIRICAL GENERALIZATION | MIDDLE-RANGE THEORY | GRAND THEORY |

Figure 5.3: Continuum describing characteristics of theories at the various levels of abstraction.

If on the other hand, the intent of the theorist is to explain or understand, perhaps even predict, more immediate events or phenomena in the real world, the theory must of necessity be somewhat less abstract and therefore less general than grand theory. While this can range from generalizations based on particular situations, on up, it usually applies to what Merton (1967, 1968), and others after him, have labeled middle range theory, theories for which
empirical (real world) referents are more clearly apparent than is usually the case with grand theory. Note, there are also those who prefer to deal on a primarily descriptive level, rejecting the notion of abstraction altogether, but this would not fit our definition of theory and so will be discussed only briefly within the next topic.

Kaplan (1964) discusses the distinction between grand and middle range theory at some length, (although he phrases it in terms of the difference between realism and instrumentalism). According to that discussion then, realism puts its emphasis on the brute determinants of theory. If a theory is essentially a picture of reality, then to arrive at a sound theory we must concentrate on discovering how things are, rather than on inventing ways to conceptualize them (pp. 307-308). One may liken this to Merton's charge that general theories of social systems are too remote from particular classes of social behavior, organization and change to account for what is observed (1967:39).

The instrumentalists, (and grand theorists), take an entirely different approach in terms of the proper function of theory. Theory is thought to be a symbolic construction rather than a description of fact. As such it serves to integrate or organize empirical generalizations, middle range theories, even "educated guesses", into a single deductive system. Because it abstracts from the material of
experience, it creates something which has no exact counterpart in the world of experience. In so doing, it reduces its ability to predict real world phenomenon, on a one-to-one basis. Therefore, although a theory, according to this view, should be grounded in facts, it need not constitute a direct representation of those facts.

"...to understand the world (primary purpose of theory in this sense) does not mean to hold in our hands the blueprints by which God created it, but some very human sketches by which we ourselves can find our way" (Kaplan, 1964:310).

There is no clear-cut criteria dividing one type of theory from another. As mentioned above, empirical generalizations and grand theory represent the extreme ends of a continuum, with middle-range theory more or less in-between. This section has not been included in order to proclaim one type of theory as being better than the other. Rather, it is to stress that judgment concerning theory should be made on the basis of its intended purpose (i.e., to generalize or to predict). General theory has usually been criticized for its lack of predictability. Middle range theory on the other hand, is often accused of being too situation specific. If social reality could be explained in terms of simple cause and effect relationships, the task to create a theory which could "be all things to all people (i.e., be general enough to apply to all situations, yet specific enough to predict in particular situations), would
be simpler. However, social reality is not simple and a
theory such as that just described would seem to be useless
as well as unrealistic. Kaplan notes that this fact has been
recognized (although apparently not by many sociologists) for
as far back as the conception of Jain logic in ancient India.
Thus, the doctrine of Syadvada proclaimed that

"...every proposition is true only up to a
point, in a manner of speaking, in certain
respects...even with the best of theories we see
through a glass darkly, and know only in part"
(pg. 310).

William Skidmore discusses this relationship or movement
away from predictability and towards generalizability in
terms of susceptibility to falsification. Thus the more
encompassing the theory, the less susceptible it will be to
direct falsification.

It is up to the theorist, then, to decide on the goal
he/she expects of the theory being developed [e.g.,
description; explanation, and/or prediction (middle-range
theory); or explanation and verification\(^1\) of a reconstructed
picture of reality (grand theory)]; and finally to choose the
level of abstraction which is most appropriate for achieving
that end.

\(^1\)Verification as used here is distinct from one-to-one
predictability. Thus verification is meant in terms of the
confirmation of lower level hypotheses which are deducible
from more abstract theory.
Implications

The level of abstraction intended by the theorist tends to influence both what is examined and how it is examined. Thus the concepts chosen by a realist (interested in middle range or lower level of abstraction) generally examines concepts that have direct empirical referents in the real world; or, as in the inductive approach, concepts which arise out of direct observation of their real world counterparts. The instrumentalists (grand theorists) on the other hand tend to be interested in concepts which are more removed from direct empirical referents and which usually subsume several less abstract concepts.

In a similar fashion, the preferred level of abstraction is closely related to the type of approach one takes in examining those concepts. We shall see in the next section, that the inductive realist's goals are most often accomplished with either a descriptive or an exploratory approach. The deductive realist may also choose one of these, although he/she usually does so only in terms of a preliminary step or stage. The predictive analytical approach then generally serves as the primary procedure. For the instrumentalist, explanation and verification are more important than prediction (although some degree of prediction is required to establish verification).
The selection of a differential strategy then, also represents an implicit assumption—namely, that the chosen strategy is the best one to meet the goals desired for the prospective theory. As mentioned above, these strategies can be of either an inductive (as in the exploratory strategy) or a deductive nature (the outcome of an analytical strategy). The choice will be shown in the next section to depend heavily on the choices made in the previous assumptions.

How to Approach the Subject Matter

There are three major strategies for approaching a problem or area. These consist of exploratory, descriptive, or analytical approaches. Neither the exploratory nor the descriptive depend upon theory, although both can offer insights in future theory building procedures. Analytical approaches on the other hand begin on a theoretical basis, so that theory contributes (often dictates) what is examined.

The descriptive approach is primarily interested in giving a description of "what is". No attempt is made to either "test" a previously formulated theory or to form new ones. Likewise generalizations are kept to a minimum since the descriptions are considered data bound. Many of the early studies in sociology and anthropology would fall under this approach. Generally speaking, methodological concerns about randomized sampling, normal distributions, and etc., are not required for this approach since the primary emphasis
is on description of a specific event or phenomenon rather than testing or making generalizations to a larger population. Thus the merit of this approach resides in the detailed information it provides about specific phenomenon which can later serve as raw material for future theory building attempts. Homans (1950) used a variety of such descriptive case studies in developing the theory for his book, *The Human Group*.

While neither the exploratory nor the descriptive approaches rely on theory, there is a basic distinction between the two. The goal of an exploratory approach usually involves an attempt to develop generalizations and/or theory. Descriptive approaches, on the otherhand, generally do not go beyond the mere description of facts, even though others may later use the findings in an inductive process of theory building.

The exploratory approach\(^1\), then, examines the empirical world in terms of an ongoing procedure whereby theory is developed, but only after considerable observation. Reynolds

\(^1\)This approach should not be confused with an exploratory stage within the analytical approach. The exploratory stage represents a preliminary part of the analytical approach rather than the approach itself. As a consequence only a rather loosely ordered, often incomplete, theory is developed in this preliminary stage as opposed to the gradual attempt of the exploratory approach to build an inductive grounded theory.
(1971) has described this approach in terms of the "research-then-theory" strategy. It is an inductive approach which includes a variety of operating strategies, such as the Baconian strategy ["deriving the 'laws of nature' from a careful examination of all the available data (which Bacon) calls the 'interpretation of Nature' (Reynolds, 1971:140""); Glaser and Strauss' (1967) Grounded theory; and Blumer's sensitizing concept analysis.

These strategies lead to what their adherents call the development of formal theory (which they distinguish from the usually deductive grand or middle range theory). In essence this approach involves a continuous interaction between empirical observation, interpretation, and theory construction. The chain-like sequence of this interaction has been depicted in Figure 5.4. Note that this process is depicted as an ongoing process, just as the real (empirical) world is always in a state of flux.

---

Figure 5.4: Chain of events characterized by the exploratory or "research-then-theory" strategy
Glaser and Strauss list the following advantages for this approach. Such a theory, they say...

(a) **fits the empirical situation** (since hypotheses and concepts arise systematically out of the data during the course of research).

(b) **is readily understandable** to sociologists and laymen alike (unlike attempts such as Parsons' "social action theory" which few sociologists are able to fully comprehend).

(c) provides us with relevant predictions, explanations, interpretations, and applications - in other words - **it works**. (The reason it works is because theory based on data is intimately linked to that data and thus usually cannot be easily be refuted by more data or another theory).

Glaser and Strauss cite Weber's theory of bureaucracy and Durkheim's theory of suicide as early examples of this type of approach. A more recent example would be Howard Becker's (1953) classic study "Becoming a Marihuana User".

While there is much to be said for this type of approach, it is not without its disadvantages and limitations. Speaking of the Baconian strategy but more or less applicable to all such exploratory strategies, Reynolds (1971: 142) outlines two such drawbacks.

1. The amount of data that can be collected is theoretically infinite.

2. The problem of finding substantively interesting patterns within the resulting data can be overwhelming.

In other words if social reality is constantly undergoing
change, where or when does one break off the study? How does one know that he has completely "saturated" (Glaser and Strauss) his evolving theory with all possible outcomes or categories? This would also seem to raise a correlate objection---to what extent is his theory so "intimately bound to his data" that "theory" becomes an infinite set of descriptions of unique phenomena?

The third type of approach is the analytical approach (or in Reynolds' terminology, the "theory - then - research" strategy). It differs from the other two approaches in its initial dependence upon theory. Thus in the analytical approach, theoretical thinking takes place prior to its application in the real world. Three situations characterize the level of development of this type of theoretical thinking. These can be summarized as follows:

**Situation A: Completely specified model or theory.** Occurs after considerable verification and reformulation based on research conducted under a variety of social situations. Analysis procedures usually confined to estimation (e.g. what is the size of the slope between two variables) rather than theory building or testing procedures.

**Situation B: Partially specified model or theory.** Theoretical development has progressed to the extent that some of the pertinent variables have been isolated, although others have not. Further research must be conducted in terms of both theory building and theory testing before a completely specified theory can be achieved.
**Situation C**: Have guidelines but otherwise unspecified model or theory. Has no more than an idea or perhaps has a whole "grocery list" of possible variables. Theoretical development is at a minimum. Analytical procedures are almost entirely confined to theory building techniques.

Actually there is no clear cut division between these types of situations, rather they represent idealized points on a continuum with theoretical development occurring anywhere along the continuum.

---

**Frameworks**

```
Level of Theoretical Development

<table>
<thead>
<tr>
<th>SITUATION A: Completely specified model or theory</th>
<th>Explanatory</th>
</tr>
</thead>
<tbody>
<tr>
<td>SITUATION B: Partially specified model or theory</td>
<td>Predictive</td>
</tr>
<tr>
<td>SITUATION C: Possess only a few guidelines</td>
<td>Exploratory</td>
</tr>
</tbody>
</table>
```

---

**Figure 5.5**: Levels of theoretical development and the frameworks in which they are applied

The analytical approach can be further divided into three overlapping frameworks, depending upon the goals involved (see Figure 5.5 above). The exploratory framework is usually confined to a theory building goal and is therefore utilized in a situation "C" or low "B". Very
little is usually known in a definitive sense although the possibilities are often quite extensive. Actually the techniques pursued in an exploratory analytical framework are not all that different from those discussed earlier under a "research - then - theory" framework. The theorist depends heavily on results from empirical research to narrow down the list of probable variables. The major difference seems to be one of degree. The exploratory analytical framework is considered to be no more than a preliminary stage within this approach.

The second type framework that is usually employed within the analytical approach is the predictive framework. In this framework the attempt is made to isolate those variables or systems of variables which "fit" the empirical situation the "best," (i.e. yields the closest prediction of a criterion variable), until recently, most of what was being done in sociology using an analytical approach was of a predictive nature. Maximization of $R^2$ has represented the critical test for most predictive studies. The emphasis is placed, therefore, on variables as measured rather than the structure of relations between true concepts. This means that the linkage between the measured variables and the abstract concepts, as well as between the empirical hypotheses and theoretical propositions, is not as important as whether the measures and empirical hypotheses predict.
While this type of framework has been useful in situations where one is still attempting to isolate important variables (i.e., "B" and "C" type situations), some sociologists have recently suggested that continued emphasis on prediction is detrimental to the further development of sociological theory.

Hubert Blalock (1969), has been prominent among those who hold that sociologists should be turning their attention to an explanatory framework where the emphasis would be on determining the structure of relationships between concepts rather than the ability of measures to predict. In terms of analysis this would mean examining slopes, (rather than correlations or $r^2$), evaluating whole models (rather than single predictive equations); and assessing the impact of measurement, specification and sampling errors in an attempt to adjust back from measured variables to "true" concepts, (rather than confining one's attention to measures only).

The ideal situation would seem to call for a combination of the explanatory and predictive frameworks. We have, however, already discussed some of the problems involved in expecting precise predictions from a general theory. It has been suggested that some of the problems could be alleviated by adding conditional statements to the general theory. This would tend to increase the predictive power of a theory, providing the conditions were met, while still allowing the
theory to be general enough to extend beyond a few situations or events. In a later chapter, the notion of causality will be introduced. It will be shown that a pragmatic revision of the traditional conception of causality, now calls for the kind of approach described above, namely the development of a general causal theory, which is accompanied by conditional statements describing when and in what manner it applies.

Implication

The approach one decides to take in terms of developing a theory represents one more boundary-setting assumption which in effect narrows the scope of interest delineated by the two previous choices. As with the others this represents an extremely important decision since it can affect all aspects of the remaining theory construction process. This means that the type of concepts, measures (whether to have any and if so, what kind), hypotheses, measurement decisions (design, data collection, and etc.), and the amount and type of analysis used will often depend upon the decisions made in regard to the three assumptions described above (i.e., choosing a theoretical orientation, level of abstraction, and approach to the subject matter). There seems to be no "real" evidence to support any one orientation, level of abstraction, or approach to the problem. As George Ritzer has pointed out, no single (present-day) orientation, level or approach can give a complete understanding of all social
phenomena. Such choices can be said, therefore, to reflect the theorist's value system and represent only some of the ways in which values impinge upon even the most "value free" enterprise.

But what does this mean in terms of building a "scientific theory" and a science of sociology? Can a "scientific sociology" tolerate theory which has been influenced by the values of the theorist? This and the related topics of ethics and pragmatics will be discussed next in terms of how and whether they enter into the development of sound theory.

Impact of Values

Values, ethics and pragmatics are three considerations which enter into practically all problem selection procedures, although they are not generally recognized as doing so. Since the time of Weber, students have been taught that a scientific enterprise must be kept free of the theorist's own value system or the product would become biased reflecting a preconceived notion rather than reality. William Skidmore has summarized this point of view as follows:

"In order to discover what 'is', and to properly conceptualize what "is", it is necessary for the sociologist to bring no personal prejudice or bias concerning proper social relations to his study. This does not mean that he ought not to be a moral man, but for purposes of description and theory, if one wishes to know what is, then
one must observe, describe, and theorize dispassionately. If disinterestedness is not maintained, what one believes ought to be may get in the way of what is; dogma would [then] interfere with thought" (1975:34).

This position with regard to values sounds fine on the surface, values which bias or cloud our judgement with respect to what is real have no place in scientific theory building. However, to accept the position above is to accept the equation that all values are generically biasing, an assumption which is rejected here as well as by others (Kaplan, 1964; Skidmore, 1975; Gouldner, 1970; and etc.).

In order to pursue the value free position further, perhaps we should examine the term, bias. What is meant when they say that values bias results?

"We may describe bias in general as a kind of inverse of genetic fallacy; a proposition is accepted or rejected, not on the basis of its origin, but on the basis of its outcome. It is believed or not according to whether our values would be better served if it were true than if it were false" (Kaplan, 1964: 373).

Thus values can be said to be biasing one's results if results are somehow modified to bring them into line with our preconceived notions of how things should be (a value judgement). Kaplan makes the statement that subscribing to values does not necessarily entail bias. What is important is not whether one adheres to values but how one conducts his/her inquiry (i.e., the effect those values may or may not have on the outcome).
"Everything depends upon the conduct of the inquiry, on the way in which we arrive of our conclusions. Freedom from bias means having an open mind, not an empty one. At the heart of every bias is a prejudice, that is to say, a prejudgement, a conclusion arrived at prior to the evidence and maintained independently of the evidence" (Kaplan, 1964: 375).

and again...

"Values make for bias, not when they dictate problems, but when they prejudge solutions" (Kaplan, 1964: 382).

Some have attempted to exclude values entirely, simply describing things as they are. However, because of the often unconscious and subtle nature of the part that values play, their impact is not easily excluded. "The attempt to eradicate biases by trying to keep out the valuations themselves is a hopeless and misdirected venture" (Myrdal, 1944: 1043).

What are some of these often unrecognized ways in which values intrude into our activities as social scientists? One of the primary ways in which values enter is in the selection of the problem (What to Study?) -

"Whatever problems a scientist selects he selects for a reason, and these reasons can be expected to relate to his values, or to the values of those who in one way or another influence his choice" (Kaplan, 1964: 381).

In addition to the assumptions mentioned earlier, theoretical orientations in sociology usually carry with them sets of values regarding who, what, where and how to study things. These are values which are generally held in common
by most of the adherents of an orientation. The results of selecting a certain orientation and/or set of values are biased only in so far as the scientist fails to recognize that there may be other ways of studying or conceptualizing the phenomenon that might be equally as effective, a failure which is often exhibited by fervent advocates of particular theoretical perspectives.

"Many sociologists are more committed to their paradigm (perspective) than the development of sociological thinking. Their major commitment [then] is to the 'victory' of their paradigm (perspective)" (Ritzer, 1975: 211).

The subtle effect of values also enters into the data collection process. The choice of collection technique, the very questions that we ask, as well as the way in which we ask them, are all influenced by our values. Furthermore, the interpretations we make with regard to our findings must always be concluded in terms of our field of knowledge and experience. Thus Kaplan states, "values enter into both the making of a prediction and into the conceptualization of what 'it' is [that is being predicted]" (Kaplan, 1964: 386).

The question, "Who is to benefit" is both a value oriented and a pragmatic one. The answer usually has implications for both the problem area selected and the methodology that is used. Most sociological studies are undertaken with one or more of the following groups in mind.
1. Science (the development of sophisticated techniques, measures, or general theory)

2. Sociology (the expansion of substantive knowledge)

3. Funding Agencies (concerned with obtaining solutions to problems on the lowest cost/solution basis)

4. Specific target groups and/or general public (often on the receiving end as subjects for study and as participants in resulting programs)

The third and fourth groups are usually not as interested in learning the "true relationships between abstract concepts." They are more interested in pragmatic results, (e.g. if health care services are allocated $X more, what, if any, will be the increase in the overall health status of a community?)

Likewise, questions which are of interest to the first and second groups often conflict with the interests of those in the third and fourth groups. As scientists, sociologists have an obligation to pursue knowledge but sometimes this results in unwelcome or unwarranted costs for those who are studied. All research must be weighed in terms of both economic and social costs and benefits.

Even where the desired information does not conflict, the theorist may find his/her interests being cut back in order to allow for the pursuit of specific requests by a funding agency which have little or no theoretical value. What the social theorist must learn is that ideal and real
conditions are rarely the same. The information that is finally gathered almost always reflects a compromise between the interests of two or more of these groups. To the extent, then, that the theorist must compromise, at least some of the data of interest will probably be replaced by data which are only of interest to the funding agency and/or the target audience. As such the question, "Who benefits?" represents another value and boundary setting issue which must be considered before the theorist even begins to formulate his/her problem statement.

It should be evident from the discussion above that values often have a subtle but nevertheless real impact on even the most ardent supporters of the "value free" doctrine. The question inevitably arises then, how does one deal with these kinds of value infringements? Myrdal (1944) believes the only alternative (since he does not believe that values can be eliminated) is to make ones assumptions known while minimizing the detrimental effects (bias) as much as possible.

"There is no other device for excluding biases in the social sciences (or any other) than to face the valuations and to introduce them as explicitly stated, specific and sufficiently concretized value premises".

Kaplan agrees with the above, but he stresses that those values must constantly be put to the test, "even values, it seems to me have their function in inquiry, if they are
explicit and empirically grounded" (Kaplan, 1964: 409). In the chapters ahead, we shall see that it is not enough to merely state one's assumptions. The social theorist must also assess the degree to which he is meeting those assumptions. This is what Kaplan is referring to when he writes that values must not only be stated but they must also be empirically grounded.

Summary
This chapter has been concerned with describing the importance of stating assumptions explicitly. Decisions regarding "What is the most appropriate theoretical perspective?"; "What is the proper level of abstraction?"; "What is the best approach for examining the subject matter?"; all serve to limit the scope of interest but are generally not taken into consideration once the decision is made. For instance, a theorist who selects conflict theory for a perspective will rarely test to see if conflict theory is indeed the most illuminating perspective that could have been chosen. He/she merely assumes, by virtue of its selection, that it is. A theory which has been developed under one set of assumptions may not hold however, when a different set of assumptions are imposed, leading to confusion rather than theoretical advancement. Such a situation need not occur if the theorist will only take the effort to state the assumptions on which his/her theory
An assumption which has become almost institutionalized over the years is one which calls for a "value free" approach to scientific theory construction. I have attempted to demonstrate some of the often subtle and unrecognized ways in which values and value judgments impinge upon any and all endeavors. To repeat a point made by Kaplan "being free from bias is having an open mind not an empty one." It is not values that need to be eliminated (an impossible task) rather it is their subtle unrecognized and detrimental effects. The best (and perhaps only) way to guard against such effects is simply to make one's values known so that their effects, whatever they are, can be effectively evaluated by others as well as by the theorist.

Assumptions have been discussed in terms of their implications for future decisions in narrowing down the area of interest. The next chapter will also be concerned with activities that reduce the scope of concern. These activities will however be of a far more specific nature than the ones considered in this chapter.
CHAPTER SIX:
STATEMENT OF THE PROBLEM

The previous chapter was concerned with outlining various assumptions (such as the choice of theoretical orientation, type of approach and etc.) that tend to narrow the scope of interest. For example, it was pointed out that adherents from different orientations emphasize different aspects of social life, to the extent that all could examine the same situation but investigate entirely different aspects (problems) within that situation. In this chapter I shall demonstrate how the statement of the problem continues to narrow the focus of the theorist's endeavor because it outlines in detail both the nature of the topic that serves as the focus of interest and the procedure (the how) one expects to follow in examining that phenomenon.

"Without some guiding idea we do not know what facts to gather. Without something to prove, we cannot determine what is relevant and what is irrelevant" (M. Cohen, 1956:148).

Clearly then, the statement of the problem is important in the sense that it offers the theorist a guide as to what to study (i.e., specifies the nature of the problem in clear and concise terms) and how to study it (i.e., details both the objectives of the study and the procedures to achieve them). Learning to state the parameters of a problem is not particularly easy and is often neglected. Nevertheless,
theorists have frequently found that difficulties which arise later in the theory construction process tend to be directly attributable to an insufficient attention to detail in the statement of the problem. A vague or ambiguous problem statement will generally lead to confusion or error in the ensuing steps. The importance of a comprehensive and clearly enunciated problem statement can perhaps be best summarized by the following quote from Kerlinger (1964:22).

"Though this statement seems self-evident, one of the most difficult things to do, apparently, is to state one's problem clearly and completely. In other words, you must know what you are trying to find out. And when you finally know this, the problem is a long way toward solution."

What are the elements then in a "good" statement of the problem? In answer to this question, I shall first propose a rather comprehensive definition, in keeping with those offered previously for both theory and theory construction. This will be followed by a discussion of the various parts contained within that definition.
THE STATEMENT OF THE PROBLEM outlines in specific terms the substantive area of interest, containing a selective review of previous work in that area, the importance of the problem, who (individuals, groups) or what (social roles) is involved in the problem, and finally the objectives and techniques for developing a theory to aid in understanding the problem.

The elements, therefore, of a "good" problem statement can be said to include the following:

1. Theoretical background leading to a precise definition of the problem.

2. Why the problem is important?

3. Whom or what the concepts in the problem refer to (i.e., the unit of analysis).

4. What is the composition of the "population at risk?" ...of the actual sample?

5. What are the major objectives

6. How will those objectives be achieved?

I shall now turn to an examination of each of these topics in the hope that their relevance for both the process and the product (the proposed theory) can be made clear.
Theoretical Background

An important goal of most sciences is to build a body of knowledge (theory) so that practitioners and people in general will not have to start from scratch to understand something (i.e., we don't have to keep discovering the wheel). Instead, of re-discovering, concentration can be placed on extending knowledge as was done in the field of rocketry which developed the power to take man into outer space so that we might learn more about the worlds beyond our own.

The same principle applies in sociology. By using the knowledge and theories that others have discovered, we should be able to extend that knowledge by modifying, reinterpreting, or developing alternative theories.

A "review of the literature" has generally been regarded as a compulsory element in articles, papers and theses. However, the link between that review and the

---

It would seem however, that this represents another example of how disjointed most approaches have been in terms of their overall approach to constructing theory. Very often in Masters and Ph.D. theses one will find a separate "review of literature" chapter. What is needed is not one separate chapter but a review of the literature at each phase. Thus there should be a review of the theoretical approaches that have been taken, a review describing the various definitions (both conceptual and operational), and a review of the analytical approaches that have been taken with regard to previous studies in that substantive area. For this reason, review of literature will be taken up at various points in the text.
analysis has not always been clear, as in cases where path models are introduced in the theoretical sections of papers but ignored throughout the remainder of the paper. At times then, it seems clear that the theoretical review has been included simply as an afterthought, because it was expected.

What are the implications of ignoring the careful development of a theoretical framework? Besides the fact that it puts us in a position where we must continually "re-discover the wheel," there are other consequences that also tend to impede the development of sociology as a science. The proliferation of terms for the same phenomenon is one such consequence. Multiple terms not only cause confusion but also make it difficult to relate to other works which use an alternate term. (Does the theorist mean the same thing?) It also makes the problem of synthesizing more difficult. Most important of all the consequences, however, is the fact that inferences based on a weak theoretical framework will at best also be weak or at the worst misleading (perhaps even false).

Including a review is not enough, though. It must be a critical review, one that is clearly and concisely stated, where the items included are all obviously closely related. Furthermore it should be designed to move from general positions or theories down to a clear and parsimonious statement of what is being sought (i.e., the nature of the
Beginning students particularly, tend to throw in everything (including the kitchen sink) whether it adds to the understanding or explication of the desired problem or not. This is perhaps only less desirable than no review at all. Selectivity therefore should be open enough to recognize competing theories or points of view, but restrictive in the sense of narrowing the scope to only those sources which are important to a definition of the problem. For instance, a theorist might be interested in examining the relationship between socio-economic status (SES) and the desire to enter college. The theoretical review need not include all SES studies that were ever conducted. Only those which are pertinent to the relationship in question, (or offer measurement or analytical alternatives leading to a more desirable explanation of the relationship) should be included. The fact that SES has been found to be related to a longer life need not be of interest (unless the tendency for a longer life somehow affects the relationship between SES and the desire for college).

Importance of the problem:

If the ultimate goal is to develop a theory relevant to this problem, the problem should be significant in terms of
at least some of the following:\(^1\)

a. timeliness  
b. practicality  
c. of concern to either a wide population or a critical population  
d. fills a theoretical gap  
e. permits generalization to broader principles of social interaction or general theory  
f. sharpens the definition of an important concept or relationship

The next topic to be discussed concerns the objectives and how the theorist intends to meet those objectives.

**Objectives and How to Meet Them**

What are the goals that have been set for the theory (e.g. pure vs. applied application, middle vs. grand theory, and who or what will benefit from the development of the theory)? What are the proposed steps for developing the kind of theory outlined above. Note that the objectives and how the objectives are to be met are actually restatements (or rather written versions) of many of the thinking processes that were discussed throughout most of the previous chapter on assumptions. It is in the statement of the problem, therefore, that the theorist has the opportunity to spell them out explicitly. This explication however takes place in

\(^1\)The criteria included here were adapted from a list outlined by Delbert Miller, 1970: 3.
terms of the specific problem defined in the theoretical framework as opposed to the more general approach taken earlier in the chapter on assumptions.

Implications

In addition to a review of literature leading to a precise definition of the problem, the problem statement also outlines why a theory is needed, the goals set for that theory, and how the theorist intends to go about meeting those goals. As such the statement of the problem develops the theoretical framework within which the remaining steps of theoretical formulation will take place. By stating why the topic needs to be examined and the goals set for the resulting theory, the problem statement offers a convenient vehicle for stating many (or all) of the assumptions discussed in the previous chapter. It likewise has a direct effect upon measurement and analysis when it includes a statement as to how the goals will be met. This should emphasize the importance that can, and should, be attached to the statement of the problem. Unfortunately this advise is not always taken. A statement of the problem is not unlike a road map in so far as when a theorist fails to fully develop his/her problem statement (i.e., uses an old or a faulty map), he/she is likely to run into a lot of dead ends before the right road is found.
A complete statement of the problem involves several correlate problems which may be included within the problem statement or discussed separately. These involve specifying the unit of analysis, the population and the sample. These issues tend to finish the boundary setting activities which have been discussed in this and the previous chapter. In other words, they delineate to whom or what the problem statement will apply.

**Specifying the Unit(s) of Analysis**

Specifying the unit of analysis has been described by Peter Abell (1971) as the most fundamental decision to be faced at the outset of any sociological investigation. Concepts may be said to be the "building blocks of theory" but the concepts themselves represent properties of, or relations between, units of analyses.

**Types of units of analysis**

There seems to be no real concensus as to what can be called a unit of analysis. Inkeles (1964) says that units of analysis consist of social acts, social relationships social roles, institutions, households, social groups, communities, and societies. They constitute what he calls the basic elements of social life. As such he would apparently include folkways and customs among his list. Peter Abell (1971) presents a framework focusing on the individual as the most
basic unit which in turn make up the more complex units of aggregates, structures, and collectivities,

"The vast majority of empirical studies start with individual people, though, to be fair, these are often supposed to represent more abstract entities—such as when somebody is interviewed to determine the 'nature of his role'" (Abell, 1971).

It would seem from the quote above that Abell is reluctantly acknowledging the possibility of an even more basic unit (i.e. social roles). Unlike Abell, Bertrand (1972), specifically includes social roles in his list of units. He divides units into micro units (norms, roles, status positions, situses, and stations) and macro units (social relationships, small groups, complex organizations, communities, and society). Still another classification system for units of analysis was proposed by Teodor Shanin (1972). The figure below summarizes his notions concerning types of units of analysis. Clearly, this framework

<table>
<thead>
<tr>
<th>System of Sociological Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Analysis</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td>Sociological Unit</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Individual Unit</td>
</tr>
</tbody>
</table>

Figure 6.1: Theodor Shanin's basic units of sociological analysis
is different from the others, although some of the same units are included. It should perhaps be noted than Shanin prefaced his discussion with a statement that he realizes there are such things as complex or formal organizations, but preferred to ignore them.

This is only a small sample of the variety of phenomena that have been labeled "unit of analysis". For the purpose of discussion throughout this text however, the following definition will be implied.

---

A UNIT OF ANALYSIS is the basic entity that is being examined. It is the answer to the question "To whom or what do the concepts refer?"

---

This definition could encompass most of the classification schemes discussed above but is particularly relevant to the first three. Certainly it should be obvious from the wide assortment of alternatives, that it behooves the theorist to clearly state or specify the nature of the unit of analysis being examined.
The specification problem

The diversity of opinion represented by these examples of units of analysis need not lead to confusion or dismay. It is not the availability of alternative frameworks that causes problems, rather, it is the failure on the part of theorists to clearly specify the unit of analysis at the outset and then to maintain that unit throughout the various phases of theory construction.

But specification involves more than a simple statement that "X" is the unit of analysis. The nature of the unit, as understood or assumed by the theorist, must be clearly stated. Even commonly used terms such as small groups, organization, community and etc., convey different meanings to different people. Community studies contain many good examples of the type of confusion that exists with regard to the use of the term "community." George Hillary (1955) conducted a survey on definitions of community and found no fewer than 94 such definitions whose only area of agreement was that they all dealt with people.

"There is one element, however, which can be found in all of the concepts and (if its mention seems obvious) it is specified merely to facilitate a positive delineation of the degree of better heterogeneity: all of the definitions deal with people. Beyond this common basis, there is no agreement" (1955:117).

It should perhaps be noted, lest one make the erroneous assumption that the situation has improved since 1955, that
Bell and Newby in a review of work on the community since 1955, concluded that "It should be apparent by now that it is impossible to give the sociological definition of community" (1972:31).

In conjunction with the specification problem cited above is the problem of distinguishing between unit of analysis and a setting within which the examination is occurring. This has been a particular problem with the terms, community and society, and to a lesser degree, organization. Thus many so called "community studies" in reality are studies about groups or organizations within a community setting or context.

The consistency problem

A second problem that often arises in connection with specification of the unit of analysis is that of consistency. It is not unknown for the data to be collected from individuals even though the theory and inferences refer to organizations. Theory and data can be said to be on two different levels of abstraction since the data refer to specific concrete phenomena whereas the theory is more general and not restricted by time of place. While the level of abstraction changes, care must be taken that the unit of analysis remains consistent. The area of effectiveness offers a good example. In each box of Figure 6.2, the subject matter is effectiveness. However, the content
<table>
<thead>
<tr>
<th>Unit of Analysis</th>
<th>Individual</th>
<th>Organization</th>
<th>Small Group</th>
<th>Community</th>
<th>Society</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Abstraction</td>
<td>Degree to which he accomplished prescribed tasks</td>
<td>Achievement of organizational goals, i.e., making a profit</td>
<td>Maintaining cohesion of the group</td>
<td>Providing necessary services to members</td>
<td>Adequate growth</td>
</tr>
<tr>
<td>Theoretical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Rate of economic growth</td>
</tr>
<tr>
<td>Operational</td>
<td>Number of completed jobs</td>
<td>Profit = Income - Costs</td>
<td>Number of mutual interest groups</td>
<td>Number of services provided</td>
<td>Income or gross national product</td>
</tr>
</tbody>
</table>

Figure 6.2: Effect of changing either unit of analysis or level of abstraction on the meaning of the concept effectiveness.
differs from box to box depending upon the unit of analysis and the level of abstraction that is being examined. (Note the theoretical and operational definitions in the squares are not the only choices which might have been selected. Furthermore, the importance of this example lies not in the exact content chosen but rather that the nature of the concept effectiveness changed depending on whether one changed the unit of analysis, the level of abstraction, or both.) In describing the theoretical background of effectiveness, many theorists will remain entirely on the organizational level of analysis. Then, when they switch to the operational level of abstraction, the unit of analysis will suddenly change to individuals. Finally, the unit of analysis to which their inferences are made will sometimes switch back to the organizational level.¹

Most sociologists have found that the more complex their intended unit of analysis gets, the more difficulty they have with the issues discussed above. Specification and consistency become more and more problematical. An individual can usually be observed and/or asked questions, but it is harder to observe an entire organization, not to mention an entire society. This is the major reason that

¹Riley (1963) and Zito (1975) describe this as an atomistic fallacy, see the discussion in the next section.
sociologists turn to aggregated data (i.e. data that is collected from individuals or a lower level and combined in some fashion to represent data for the more complex unit to which the lesser unit belongs) as means of measuring the more complex units. The indiscriminate use of aggregated data has led to problems, however. These problems will be discussed in the next section.

Ecological correlation or fallacy?

There has been a great deal written about the uses and misuses of both ecological correlations and aggregated data. An ecological correlation is an attempt to assess the relationship of a group variable on another variable. In his classic article on the misuses of the ecological correlation, Robinson (1950) was not suggesting that we do away with ecological correlations. The point he was emphasizing is that too often sociologists attributed group characteristics or properties to the individuals who belonged to that group. For example, Zito gives the following illustration of ecological fallacy and its opposite, atomistic fallacy. These are two squares. Each is divided into 4 equal parts.
Figure 6.3: Illustrating ecological and atomistic fallacies

The ecological fallacy consists of attributing (inferring) the group characteristic, "squareness" of the whole B to each of its parts, b1, b2, b3, and b4. In the case of the "A" square this inference would have been appropriate. A point that Meltzer (1950) and Hammond (1973) try to demonstrate. However, in the case of the B's this inference is fallacious.

The atomistic fallacy occurs in the reverse. This type of fallacy happens when one attempts to infer the triangleness of b1, to the whole figure B, or attributing individual characteristics to the group.

It was Robinson's contention that ecological correlations are generally utilized "only when the individual data is not obtainable" (1950). In other words, most ecological correlations do become ecological fallacies because of the way sociologists tend to apply them. Perhaps this is why some less perceptive writers have mistakenly
utilized the two terms, ecological correlations and ecological fallacy, as if they were synonymous. (For an example of the latter error see Felson, 1974). Most writers have not gone to Felson's extreme position of advocating the elimination of ecological correlations. As Laumann and House state:

"No one has seriously argued that one must inevitably be led astray by the ecological approach" (1974).

Generally these writers agree with Robinson that ecological correlations can and have been misused at times in the past. However, they disagree that this is inevitable (as Felson states) or that ecological correlations are generally used because of a lack of individual data (Robinson).

"It is the inference that (may be valid or) invalid, not the correlation at the group level. We might well be interested in the relationship at the group level for its own sake" (Abell, 1971:25).

Needless to say, the last word concerning the relative merits or demerits of utilizing aggregated data has not yet been written. What is clear from this ongoing discussion, which began with Robinson in 1950 and continues today, is that the term "ecological correlation has often been misconstrued as being necessarily bad (i.e. equated with the term "ecological fallacy"). Ecological correlations may indeed have their place as suggested above by Abell and others. However, care should be taken whenever such
inferences are made because the line between ecological correlation and ecological fallacy can be a thin one.

**Relationship between unit of analysis, population and sample**

A final point related to the selection of a unit of analysis concerns what many would designate as a methodological issue—the delineation of a population and sample(s). The selected population should reflect the unit of analysis of interest within the area of interest. On the surface this would seem like an unnecessary comment to make. However, in practice, a majority of theorists fail to make any reference to the population of units to which his/her theory will apply. This is particularly important in the verification stage where the population represents the theoretical level, whereas the sample represents the empirical level. This means that if the empirical hypotheses hold up, inferences can be made regarding their veracity in the population and thus indirectly adding evidence in support of the theory. The theory building process that is being discussed here assumes that some attempt will be made by the theorist to carry through the process rather than stopping after the formulation phase. Therefore it is necessary for the theorist to describe both the population and the sample with which he intends to "test" his theory.

Defining a population could also be relevant even where a direct test was not intended. If the theorist attempts to
write a conditional theory or attempts to define conditions for a general theory then defining the population(s) under which the theory should hold would, of necessity, be one of the conditional statements.

The interrelationships between unit of analysis, population and sample are illustrated in both Figure 6.4a and Figure 6.4b. As one moves down from unit of analysis to the sample, each step represents a greater restriction of the domain under consideration.
In Figure 6.4b, two units of analysis, individuals and organizations, have been selected. Note first of all, that the unit of analysis remains the same as one moves down the level of abstraction (individuals and local coordinators vs. organizations and local agencies). Another interesting point.

---

**Figure 6.4b**: Examining the interrelationship between unit of analysis, population, and sample using two different units of analysis concerns the use of the term "normative organizations" under both sets. On the right side they serve to describe the type of unit of analysis (i.e. organizations) that is being
examined. On the left side, however, normative organizations are used in a contextual framework only. In other words, it is the individuals within such organizations that are important for the theorist not the organizations per se.

Unless the population selected is very small, it is difficult to deal analytically with an entire population. By using a sample of the population, however, it is possible to analyze relationships with relative ease. The assumption is generally made then that the relationship exhibited in the sample analysis will be similar to one in the population as a whole. There are a number of problems involved in making such an assumption. As Figure 6.5 illustrates, there are an infinite number of samples which can be drawn from a population. Some of these will be more reflective of the population than will others, even though the same selection techniques may be employed.

It should be plain then that some sampling error (error caused by the nonrepresentativeness of the selected sample) will be inevitably involved when inferences are made to the population. Those papers or articles which say that they are assuming their sample is the population or fail to specify the population are ignoring the possible impact of this type of error.

There are two other types of errors which may also cause problems in making inferences from the sample to the
population. These are measurement error and specification error. Specification error results when one fails to include the most appropriate set of variables or fails to utilize the appropriate functional relationship (for instance, when one uses a linear model to represent a logarithmic relationship). This may also involve the inclusion of inappropriate variables or the exclusion of important ones. Measurement error on the other hand lies in the empirical instrument used to measure the theoretical concept. Arriving at a reliable and valid empirical instrument for measuring a theoretical concept is not an easy task, particularly when the concept is quite complex.
For example, suppose a theorist has selected the concept, socio-economic status. How does one measure this concept? Some have said that one's social status is the amount of power one can exert in a community. Hunter in his book, *Community Power Structure*, describes the reputational method which measures status by means of asking people who they think gets things done in the community. Others have said that social status is based on one's income, education, and/or occupation. Each of these have been used separately and in combination to measure social status. Warner's Index of Status Characteristics (Warner and Ells, 1949), combines five factors in a single index or scale. The precision with which these various measures capture the meaning of the theoretical concept, varies. The amount of measurement error...
then will depend upon the degree to which the empirical measure is not isomorphic with the meaning of the theoretical concept. The problem of measurement error will be discussed periodically throughout the text, since it is relevant to nearly all of the ensuing topics.

Sampling error, measurement error, and specification errors represent the major problems encountered in generalizing or inferring from sample results to the entire population. Care should be taken to assess their impact before any major conclusions are drawn with respect to the population. The attention that has been paid to the impact of these errors has been less than adequate in the past, since few articles address these issues, except perhaps to make the assumption, in passing, that these errors are minimal.

Up to this point, the discussion has revolved around a general discussion of domain setting activities (i.e. statement of the problem, unit of analysis, population and sample). I shall now turn to the empirical example, alluded to in Chapter Three, which will be utilized throughout the remainder of the text, as a means of illustrating the various points with real data. This means that a theoretical foundation will be laid out in this and the remaining chapters concerned with theoretical formulation. Furthermore, a set of data, which is consistent with that
theoretical foundation, will be utilized to illustrate the measurement and analysis techniques that will be discussed. This approach should enable the reader to see immediately how the various procedures are interrelated as well as how they can be applied.

A Causal Model of Effectiveness

This section will be concerned with examining a causal model of effectiveness that was developed at Iowa State University by Charles Mulford, Gerald Klonglan, Richard Warren, and Paul Schmitz (1972). This model, and the data on which it was based, will be used to illustrate the issues which will be discussed in the remainder of the dissertation.

Using a single set of concepts and data should help the reader gain a feeling of continuity and integration from one issue to the next. It is hoped that this integration of tasks and issues can then be implemented when the reader attempts to develop his own theory.

Theoretical background of effectiveness

In 1961, Amitai Etzioni noted that most research on effectiveness had emphasized either structure or motivation as determinant. It was his contention that neither approach was sufficient by itself to gain a clear understanding of effectiveness. Thus his own theory, based on the notion of
compliance, allowed for both structural and motivational variables within a general systems framework. Some ten years later, Tausky (1970) described three major approaches: the classical (what Etzioni had called structure); the human relationist (motivational); and the structuralist (general systems theorists, among whom he placed Etzioni). Note: this represents the theoretical perspective assumed in the example.

The classical approach views organizations as closed systems which means that the system is assumed to be unaffected by its environment. Explanations of effectiveness, both on organizational and individual levels, are confined to structural variables within the organization. For instance, organizational effectiveness is generally defined by this approach as accomplishment of formal goals (where formal goals are those outlined in official organizational charters).

The closed model approach above was rejected by the human relationists (motivationists). So too, was the examination of structural variables. Formal goals were said to have little or no effect in the real world. Taking a more social psychological perspective, the informal goals of the workers were stressed rather than the formal goals of the organization.
As was pointed out by Etzioni, neither of these two approaches were found to be entirely satisfying which then led to a third approach, General Systems. While this too has had its problems, it seems to offer a more complete approach to the study of both organizational effectiveness and individual role performance (effectiveness). The general systems approach has been said to have arisen, at least in part, from the work of Talcott Parsons on what he called the social system. Parsons viewed organizations as semi-closed social systems which are acted upon by both motivational and structural factors. Organizational effectiveness was defined as the degree to which the organization was able to solve four necessary functional problems - adaptation (A), goal attainment (G), integration (I) and latency (L). Yuchtman and Seashore (1967) refer to this approach as the derived goal approach because the functional goals are independent of the intentions and awareness of the organizational members. In other words, although these problems must be solved in order for the system to survive, rarely if ever is there a conscious realization of this need on the part of the organizational members. Thus the formal goals are conspicuously different from the derived, functional goals (to solve the problems). Parsons indicates that most organizations tend to emphasize one type of problem over the others, depending on the orientation of the organization. In
fact he has proposed a typology based on the type of problem which receives most of the organization's attention. For instance a church is ostensibly more interested in pattern maintenance problems (socializing its members). This should not be interpreted as meaning that an organization which emphasized one type of problem solution can disregard the others. What it does mean is that for each organization there is an optimum relationship between these functional goals. Thus it is not necessary to stress the solution of all problem types equally, rather the proper interrelationship between these functional goals should be maintained. The church is a good example. Although the pattern maintenance problems may be the most prominent, few churches can ignore economic (adaptation) problems. Likewise, particularly during the 1960s, many churches found themselves involved in goal attainment problems (Viet Nam, civil rights, etc.).

Parsons tends to be primarily concerned with elaborating social system effectiveness rather than individual effectiveness but in his description of the "unit act" he says that action involves individual decision-making, constrained by situational conditions, in the pursuit of goals. Thus Parsons' system of action includes both subjective or motivational elements (voluntary decision making) and structural components (situational constraints
such as norms, values, etc.).

Etzioni's own approach to the study of organizations is also said to combine both structural and motivational considerations. His model of organizations is based on compliance relations. Compliance was defined as "...a relationship consisting of the power employed by superiors to control subordinates (structural) and the orientation (motivation) of the subordinates to this power (Etzioni, 1961:XV). Etzioni has suggested that compliance is related to other important organizational variables, particularly as they apply to the behavior of lower participants in the organization.

One of the first attempts to place some of Etzioni's propositions into a causal framework was done at Iowa State University. The article "A Causal Model of Effectiveness in Organizations" by Mulford, Klonglan, Warren and Schmitz (1972) is an example of the work that has been done in this area. This particular article emphasized Etzioni's propositions which dealt with lower participants and their effectiveness (role performance) in a normative organization. Note: the level of abstraction is obviously middle range not grand theory since it applies to lower participants in a normative organization. Furthermore, the term causal framework will be shown in a later chapter to relate to an explanatory, analytical approach.
Using Etzioni's propositions (and inferences from current research where Etzioni did not specify a relationship), the authors constructed and tested a causal model of individual effectiveness in organizations. The results of their analysis tended to support many of Etzioni's original propositions.

This is then the general framework with which the remainder of this monograph will be concerned. An example data set, which relates to this framework, will serve as the basis for the empirical examples that will be utilized throughout the remainder of this text. The following variables represent the subset of nine variables on which those examples will focus.

1. **Selectivity** - "ratio of potential participants over actual ones" (Mulford, et al., 1972:64).

2. **Socialization** - "the acquisition of the requisite orientations for satisfactory functioning in a role." (Mulford, et al., 1972:63).

3. **Communication** - "a symbolic process by which the orientations of lower participants to the organization are reinforced or changed." (Mulford, et al., 1972:64).

4. **Scope** - "the number of activities in which the participants are jointly involved and the extent to which activities of the participants of an organization are limited to other participants of the same organization." (Mulford, et al., 1972:66).
5. **Pervasiveness** - "the number of activities in or outside the organization for which the organization sets norms" (Mulford, et al., 1972:66).

6. **Salience** - "the relative emotional significance that the participant associates with membership in the organization" (Mulford, et al., 1972:66).

7. **Tension** - "the emotional strain resulting from comparatively high levels of rationality, discipline and effective neutrality found within the context of an organization" (Etzioni, 1961:161-163).

8. **Job Satisfaction** - "a particular type of attitude or sentiment associated with a member's participation in an organization" (Mulford, Klonglan and Schmitz, 1971:104).

9. **Role Performance** - "actual behavior judged relevant to the pursuance of one's job" (Mulford, et al., 1972:69).

These will be discussed at greater length in the next chapter on concepts and measures. They constitute some of the major variables included in Etzioni's Compliance Theory.1

The problem then which will be examined as an illustration throughout the remainder of this dissertation concerns an examination of the correlates of effectiveness (role performance) as outlined above from Etzioni's compliance theory. More specifically, a causal model of

---

1For further information, see "A Comparative Analysis of Complex Organizations: An Enlarged Edition," (1975), by Amitai Etzioni.
effectiveness¹ will be applied to a specific situation 
(involving local civil preparedness director/coordinators) to 
see whether the hypothesized relations are supported.

The sample data used for illustration in this text was 
taken from a population of local civil preparedness 
director/coordinators in the three states of Georgia, 
Massachusetts and Minnesota. The data was originally 
gathered in connection with a pilot study which was later 
followed up by a larger national study. The selection of 
these states was purposive in that they reflect (1) different 
internal structures, (2) different areas of the country and 
(3) different types of expected disasters. The civil 
preparedness organization is highly hierarchical and 
intimately related to the military in Georgia. In 
Massachusetts the organization is based on townships; whereas 
in Minnesota, it is on a regional basis. Different sections 
of the country, the Midwest, North-East, and the South were 
selected because they represent varying historical and 
cultural backgrounds which might be reflected in divergent 
orientations on the part of the local director/coordinators. 
Also, the type of disaster which is typical of an area 
differs depending upon where it is located. Minnesota and

¹Actually variations of the causal model proposed by 
Mulford, et al., (1972), will be examined, in order to best 
illustrate the different topics under discussion.
Georgia are subjected to frequent tornadoes, Massachusetts is not. Minnesota and Massachusetts experience heavy snows, blizzards, etc. All of these states have been confronted with flooding. However, hurricanes are characteristic of Georgia rather than Massachusetts or Minnesota. Another difference between these three states which has policy implications is the type of general housing construction that characterizes the different areas of the country. Homes in Minnesota and Massachusetts are more prone to have basements than homes in Georgia, thus the problem of stocking and securing shelters is probably more relevant in Georgia than in the other two states.

As noted above the internal authority structure varies from state to state. Figure 6.7 below illustrates some of the more common arrangements. The local civil preparedness agency receives half of its funding from the federal government and the other half from local funds. A regional field officer provides the orientation guidelines at the local level, often without consultation with the state level. The local director/coordinator (the unit of analysis throughout this example) usually serves as a volunteer or on a part-time pay basis. In many cases he is the entire staff of the local agency. It seems reasonable to regard civil preparedness—-at least on the local level---as a normative organization since most director/coordinators, being neither
"employees" nor "inmates," must have at least some degree of commitment to the organization or they wouldn't have assumed the responsibilities of the position.

The sample consisted of a random sample of 90 local civil preparedness director/coordinators drawn from the master rosters from each of three states: Minnesota, Georgia and Massachusetts. Actual respondents totaled 76, 80 and 84 respectively which included both paid and volunteer director/coordinators. Response rates ranged from 84% to 93%, yielding samples which are thought to be representative of the chosen population.

The population and sample for the examples has been described at some length. Unfortunately, little is said in many journal articles about either, but this is particularly
so about the population. One of the points that will be demonstrated later, is that information about the characteristics of the population are necessary to fully interpret one's results.

Summary

This chapter has been concerned with discussing the activities that set the substantive boundaries for the theory that will be developed and tested. These activities include stating the problem, specifying the unit of analysis, and describing a population (and sample) to which the theory will apply. These activities were shown to involve more than is commonly acknowledged and to be highly interrelated, with each activity serving to restrict the area of interest more than the previous activity(s). As such they have inevitable consequences for all other activities in the theory construction process. An empirical example has been included illustrating some of the points described earlier in general terms. This example is based on Etzioni's compliance theory and will serve as the basis of illustration throughout the remainder of this dissertation.
CHAPTER SEVEN:
SPECIFYING CONCEPTS AND MEASURES

One of the most hotly debated questions in philosophy concerns the nature of social reality and the way social actors constitute that reality. It is not within the province of this text to pursue the philosophical aspects of this question in depth. However, it can be stated that part of the problem concerns the process of abstraction (a process whereby real world properties, and the relations between them, are turned into mental images or constructs) and whether or not this process creates anything with real meaning. Social theorists generally use the process of abstraction to create concepts (and propositions) as a means of verbally "capturing" reality. However, this means that the social theorist's concept can never be the real world property or thing, only stand for or symbolize it. Thus the philosophical problem, "What is reality?", must be confronted to the extent that the social theorist must examine whether his/hers concepts do bear a relationship with reality. This chapter will be directly concerned with how one attempts to do just that.

---

1The reader is referred to such sources as Nagel (1961), Hempel (1966) or Berger and Luckman (1966) for a more complete exposition of this debate.
The first thing to establish then is "What is a concept?" What a concept is, often depends upon who is doing the defining. For some, the term, "concept," is a concept. These people would probably not accept the comment in the last chapter which said a concept describes attributes or properties of a unit of analysis. To them, unit of analysis is a concept. In other words, any abstraction is a concept. The definition by Gould and Kolb (1964) would tend to fit this attitude.

"An abstract or psychological thing presupposing conscious minds which at least potentially have the concept, i.e., understand it, operate with it, apply it."

On the other hand, there are those who say as Abell (1971) did in the previous chapter, that a concept is a special kind of abstraction. Denzin (1970) and Kaplan (1964) offer definitions which seem to differentiate in this way between a unit (social object in Denzin and actor in Kaplan) and a concept. Thus Denzin (1970) refers to concepts as "images of reality," which "designate and suggest a plan of action toward some social object." This same theme can also be seen in the following statement by Kaplan (1964)

"The meaning of terms results from a process of conceptualization of the subject matter. In this process things studied are classified and analyzed ...grouping together a set of actions, verbal or otherwise as the case may be, and without regard to the actors performing them" (p. 50).

It is important that the theorist understands fully the
meaning of the term concept, in terms of what it is and the process involved in creating and utilizing it, for without concepts we have no propositions or theories. This text will view concept in terms of the second approach discussed above, as reflected in the following definition:

A CONCEPT is a mental abstraction which represents a characteristic or property of some social unit or object.

Considering their importance then, two examples will be given which will not only demonstrate what a concept is but also the processes involved in moving from a real world property to the conceptualization of that property and then back again to the real world, by way of the operational measure or indicator. The first example will concern this process as it relates to a simple object, an apple. The second will be more sociological in nature and deal with a phenomenon called socialization.
When one looks at an apple what he sees is an apple, but when he describes that apple, he does so in terms of its properties or characteristics, such as: size, shape, color, taste, etc. Likewise when he attempts to measure those properties, he is never able to grasp the property itself, only an estimate or indicator of it. For instance, if size is said to be the circumference of the apple at its widest point, one person might measure the size to be 5 1/2 inches. Another person, however, might find it to be closer to 5 3/8 inches or 5 7/16 inches. There is no way to establish, for certain, which, if any, is the "true" size although one may be able to show that one estimate is better than the others.

This illustration summarizes the major objectives of this chapter i.e., to discuss the relationship between properties (concepts), measures (indicators) and the reality that both are supposed to reflect, as well as ways to assess the quality of that reflection (the degree to which the concepts and indicators correspond to one another and to reality). The act of defining size as the circumference of the unit (apple) at its widest point represents the conceptual definition. The decision to use a tape measure reading represents another type of activity, i.e., the operationalization of the concept. When one operationalizes a concept he/she attempts to construct an indicator or measure of the concept as defined by the conceptual definition. The method of constructing the measure represents the operational definition. The measure that is
constructed must be closely related to both the concept and the reality it is supposed to represent. However, it is neither a concept nor reality, but a thing apart, similar in meaning but not in form. The following definition summarizes what will be meant by the term, operational (or empirical) measure or indicator.

AN INDICATOR OR OPERATIONAL MEASURE is an instrument (question, index, scale, and etc.) which is developed in order to capture the qualitative and quantitative aspects of a concept and the real world property which that concept represents. Furthermore, it is defined by an operational definition which should specify the following: the meaning sphere of the concept that will be included, type of measure, response and scoring frameworks (i.e., range of values and the level of measurement).

Figure 7.1 below represents a graphic attempt to demonstrate the relationship between concepts, measures or indicators and reality. The relationship shown in the figure is actually that of a process—the process of abstraction that one goes through in attempting to "grasp" something mentally and analytically. If we relate the previous example of the apple to the figure, we can see that the "thing" was the apple. It represents the thing, object, or idea being
Figure 7.1: Relationship between real world property, concept and indicator
examined, as such it would be labeled unit of analysis according to the terminology in the figure. The abstract process enables us to "see" things which do not have a physical embodiment (e.g., gravity or role) as well as those that do (apple). On the basis of observation then, a mental (cognitive realm) image is formed. This image can be described as empirical, despite its being in the cognitive realm because it is an impression or image of the thing itself (apple) and not yet an impression of the properties that make up the thing or unit. This step in the process is often overlooked when people discuss the relationship between concepts, measures and the real world. It is included here because observation is not precise, different people "see" differently. Therefore, it is important that this step be recognized because properties (concepts) are defined on the basis of this observed image not necessarily the real unit itself.

The shift from image to properties occurs in the cognitive realm in an attempt to understand or "grasp the unit mentally". It is achieved through the process of abstraction, so that properties of the unit are identified, labeled, and meaningfully defined. Thus size, described earlier as a property of the apple, represents an abstraction or concept since "in the broadest sense (concepts) are properties of, or relations between, units of analysis"
(Abell, 1971:5). Actually, size would only be a word or label until coupled with a shared meaning (Zito, 1975 and Zetterberg, 1966). The conceptual definition serves the useful purpose of establishing this shared meaning. As such the conceptual definition can perhaps be described as the bridge between the cognitive image and the abstract perceptions of its properties. In the case of the apple, the concept size was defined as the circumference at its widest point.

Having established the meaning of the abstract concept, the next step is to attempt to project that meaning into the concrete realm. For this, one devises an indicator. An indicator (as defined above) is an estimate, an attempt to capture (measure) the quality and/or quantity of a property or concept in the concrete setting. The bridge between the abstract concept and its indicator is the operational definition. As such the operational definition states explicitly the set of procedures one should use to develop the measure which will be used to indicate the concept.

Again it may seem as strange to label the indicator abstract as it did to label the image as empirical. However, it can be said to be abstract in the sense that it represents only an estimate of the true property. In order to attempt to grasp the real or true nature of a property, one must turn to inference. The process of inference constitutes the
bridge between the abstract indicator and the real dimension of the concrete unit that the indicator is supposed to represent. Thus we infer that the size of the apple is 5 1/2 inches. Conclusions about the unit (the apple) can be made, then, on this inference that the measured indicator (5 1/2 inch circumference) is consistent with the real property (size).

It should be apparent from the diagram and the discussion above that there is a continuing need for integration and reassessment, as one moves from one step in the process to the next, if the resulting concepts and indicators are to truly reflect the real world properties they are supposed to represent. For instance, a concept based on a faulty image is unlikely to produce an isomorphic indicator. Likewise an indicator which covers only part of the meaning sphere of the concept will not serve as an adequate measure of the property which the concept represents. This is particularly true of more complex concepts which contain a number of dimensions within their meaning sphere. The discrepancies which exist between studies using the same concepts may be an artifact of too little attention being given to the correspondence of indicators and the meaning spheres of complex concepts. In other words, they may deal with different aspects of a complex concept, (See Figure 7.2 on page 182).
Unfortunately, too little attention is being paid to substantive consistency. Conceptual definitions are often accepted without question, simply because they have been used before. This may lead to the adoption of an inappropriate (may be referring to a different level of unit, i.e., organization rather than individual) or an incomplete definition (includes only part of the meaning sphere you are trying to examine). The same thing can be said about indicator selection procedures. The fact that an indicator may have been used before does not mean that it refers to the same meaning sphere of the conceptual definition nor to the same image of the unit that you observe.

Perhaps a sociological example would be appropriate at this point. Let us consider the concept socialization as it was used in the example data set. The conceptual and operational definitions will be followed by explanatory statements that attempt to put the definitions within the theoretical context from which they were taken. This is done because many definitions are incomplete and one must then go to the contextual information in which they were placed in order to fully understand the theorist's meaning and intent. For instance, is it clear from Etzioni's conceptual definition below that he visualizes socialization as a multidimensional concept? In the context, however, one sees that he means for socialization to include the acquisition of
both orientation and skills, making it contain at least two dimensions.

Conceptual Definition: "Socialization is the acquisition of the requisite orientations for satisfactory functioning in a role" (Mulford, et al., 1972:63).

Operational Definition: Measured by the amount and quality of job orientation that the local coordinator had received from other local coordinators nearby, from local government officials and from state agency personnel. An additional item was used to assess his knowledge and level of understanding of his responsibilities and commitments. Scoring for each item was done with an abbreviated certainty scoring technique ranging from 0 to 9. A single composite score was formed based on the summation of the four item scores. This resulted in an indicator of socialization with a possible range of zero to thirty-six.

According to Etzioni (1961, 1975) normative type organizations tend to emphasize socialization more than other types of organizations. Furthermore they tend to emphasize instrumental socialization or job orientation rather than relying on more indirect external social units.
for that socialization. The lower participants or local coordinators in the civil preparedness organization are generally volunteers for positions which are filled with relatively little selectivity. This has created the need for formal attempts within the organization to develop the type of orientations and skills that the organization feels a local coordinator needs.

"...if the organization can recruit participants who have the (necessary) characteristics through training or education, it does not have to develop these characteristics through (its own) training or education. On the other hand, if the organization has to accept every member who wishes to join, or every member of a specific but larger and unselected group, it has to turn to socialization to produce the desired result" (Etzioni, 1961).

Relating the concept socialization to Figure 7.1 and the example data set, the unit of analysis is the local director/coordinator who has undergone, in the concrete realm, some degree of early socialization. The amount of socialization he received can be viewed as a property in much the same way as size represented a property of the apple. In the cognitive realm one no longer thinks in terms of specific individuals but rather in terms of a generalized notion of the local coordinator. The property, "socialization experienced" is generalized also. Next it is labeled and defined conceptually so as to clearly demarcate the meaning sphere of the term or concept. Finally an indicator is
developed, a socialization scale, according to the procedures described in the operational definition. This then represents the general process involved in the development of the concept socialization and its measure or indicator. But how does one assess the adequacy of what has been done?

There are a number of criteria which should be considered when examining the question of adequacy of a concept and its indicator. Note that only the first two sets (A and B) of the criteria will be examined in detail here. The third set, (C), requires the input of actual data and will therefore be discussed in Phase Two: Measurement, Chapter 9: Preliminary Assessment Techniques.

A. Adequacy of concept: isomorphism of meaning sphere (conceptual definition) with the concrete or real phenomenon

B. Correspondence between the conceptual and operational definitions in terms of
   (1) Unit of analysis
   (2) Substantive coverage
   (3) Attribute or variable
   (4) Level of information
   (5) Dimensionality

C. Adequacy of indicator
   (1) Isomorphism of meaning sphere (indicator) with concrete or real phenomenon
   (2) functional unity
   (3) reliability
   (4) validity
   (5) amount of measurement error
   (6) descriptive statistics
Adequacy of Concept

It is imperative for the theorist to set forth the meaning sphere of concepts in an explicit manner, preferably with a conceptual definition. Clear communication regarding the theorists intended meaning is the only way to avoid the confusion that otherwise inevitably results. There are many reasons why such difficulties develop in the absence of clear definitions, some of which were outlined years ago by Goode and Hatt (1952). Their comments, extending over several pages (pp. 44-48), have been summarized and elaborated on below.

1. Concepts develop from a shared experience
To understand their intended meaning, it is often necessary to understand or take part in that experience in much the same way that one learns another language. Thus American sociologists use the terms Gemeinshaft and Gesellschaft in conjunction with paragraph long definitions and descriptions rather than their English counterparts (community and society) which do not communicate the exact meaning conveyed by the German terms.

2. Terms used to denote scientific meanings may also have meanings in other frames of reference. Scientists, and social scientists in particular have frequently borrowed terms from other branches of science or the vernacular. Examples include "stratification" (taken from geology and applied to societies rather than rocks); "organic" model (borrowed from biology but applied to societies, communities, organizations and etc., instead of a human or plant body); and "culture" (representing the totality of social heritage in a society) although it is commonly understood as an awareness of fine music, art literature and "good" manners.
3. **A term may refer to different phenomena** The term "variable" represents a good example of this. Some use it interchangeably with the term "concept." Others reserve the term for an empirical indicator or measure. Still others use the term "variable" when both the concept and measure are desired, as in the case of Abell (1971), "Concepts are 'turned into' variables by mapping them into a set of values." In this text, the term "concept" will be used when the conceptual meaning is all that is desired. Likewise when the measure (indicator) alone is needed, that term will be used. However, when the notion of both are required (as is the case when one examines "variables" in a regression analysis or a path model) the term "variable" will be used. A fourth usage is completely different and will be used also. It refers to whether a concept and/or measure (a property) is conceived of in terms of mutually exclusive categories (an attribute) or continuous categories or functions (a "variable"). Obviously the term variable has been used many different ways, to the extent that its usage is being curtailed by many.

4. **Different terms refer to the same phenomena.** Concepts, constructs, variables, and even units (Dubin, 1969) have all been used by writers to refer to what has been designated here as a concept.

5. **A term may have no immediate empirical referent at all.** A concept such as social structure can be difficult to grasp empirically. It can't be weighed, touched, or even "seen" in the same way that age, sex, education, income and etc. can be measured. In order to gain a complete understanding of what is meant by the term it is necessary for the theorist to describe in detail the complex series of referents (other less abstract concepts) that one must go through before one reaches the empirical referent.

6. **The meaning of concepts may change.** In the next chapter on Linkages, I shall discuss the changes that have occurred in the meaning of the
term causality. In the past, causality has included the notion of determinism. Recently, writers such as Abell (1971) and Rhoads (1971) have suggested a different meaning for causality, one divested of its deterministic aspects. The problem is, many still operate under the old meaning, offering criticisms on the basis of that old meaning which are not relevant to the new one.

The Goode and Hatt discussion should help to illustrate the necessity for clearly defining one's concepts as a means of communicating the essential points of interest and as a means of avoiding confusion or misunderstanding by others. The adequacy of a concept depends heavily upon the clarity of its definition as a means of establishing its meaning sphere. Since definitions (operational as well as conceptual) are so important, perhaps we should examine the characteristics of a "good" definition.

There are a number of approaches that a theorist can use to examine the adequacy of either conceptual or operational definitions. I shall discuss two of them, the Aristotelian framework and Ruby's (1950) five "criteria for an adequate analytical definition. While I am well aware of the fact that the Aristotelian framework is rarely applied, at least on a conscious level, I have included it here because it offers an excellent model for beginning theorists to follow. Perhaps if more theorists applied it, we would not be as plagued by the many vague and ambiguous concepts that are currently found in sociological theory.
The Aristotelian framework divides a definition into two parts, the **definiendum** (the word, concept, or measure) being defined and the **definiens** (the defining part of the definition).

\[
\text{Man} \quad \text{is a rational animal} \\
\text{definiendum} \quad \text{definiens} \\
\text{differentia} \quad \text{genus} \\
(rational) \quad (animal)
\]

The definiens is then broken down into the **genus** (the general class of things to which the definiendum belongs) and the **differentia** (the special characteristics possessed by the definiendum). If we apply this to one of the definitions given earlier (see page 172) we can break it down as follows:

\[
\text{socialization is} \quad \text{the acquisition of the requisite orientations for satisfactory functioning in a role} \\
\text{genus} \quad \text{differentia}
\]

Thus socialization represents the word being defined (the definiendum) and the rest is the definition (definiens), of which "acquiring the necessary orientation" represents the
general class of activity (genus) and "for satisfactory functioning in a role" is the qualifying or special characteristic (differentia) which sets it apart from all other types of orientation that might fit the general classification.

This represents one way of examining a definition. The reader should be reminded that although some definitions are difficult to break down according to this framework, it can serve as a fairly good technique for constructing definitions which will avoid the pitfalls of vagueness, ambiguity and synonymity that often obscure the theorist's intentions.

While the Aristotelian framework above offers a model for constructing definitions, Ruby (1950) has listed a set of five criteria that can be used to test the adequacy of a definition. These consist of the following:

1. The definiens should be equivalent to the definiendum. This is what was meant above by saying that the adequacy of a concept should be judged according to the isomorphism (one-to-one correspondence) of the meaning sphere (definiens or definition) with the concrete or real phenomenon.

2. The definiens should state the essential characteristics of the definiendum. Social phenomenon have many characteristics, but only some that set them apart from other phenomena. Thus it is important to include only the essential characteristics.

3. The definiens should clarify the nature of the definiendum. The definition should not utilize
vague or unfamiliar terms that need to be defined themselves before the meaning of the original term can be made clear.

4. An analytical definition should avoid word substitution (circularity). The use of synonyms to define a term may meet the first criterion above but would fail to meet either the second or third.

5. A definition should be positive rather than negative. If one attempts to define a term in a negative fashion (X is not Y) it will be necessary to define the rest of the world (i.e., Y must be all non X's).

If these criteria are met, the theorist should be able to avoid the pitfalls of vagueness, ambiguity and synonymity alluded to above by Ruby. A definition is vague when the definiens does not state the meaning clearly (criterion 3). This often occurs when the theorist has not fully grasped the meaning himself/herself. The only way to avoid vagueness is to have sufficient conceptualization on the part of the theorist. Another problem that is encountered is that of ambiguity. Ambiguity occurs when there are two or more clear meanings for a term but it is not clear which is being applied (alluded to in criterion 2). In order to avoid this problem, the theorist should list the alternative meanings and then specify the one which is most appropriate for the situation. The question of synonymity has been dealt with above under criterion 5. Its solution is easier than the others, since it only requires the theorist to eliminate the
use of synonyms in his/her definitions. One should, therefore, decide on one term and avoid using all other circular terms.

The meaning sphere delineated by Etzioni in his conceptual definition of socialization would seem to be fairly inclusive of the concrete process it is supposed to represent with the possible exception of time. In his discussion of socialization, Etzioni (1961:127-150) makes the distinction that socialization takes place early in one's career in the organization. He also makes an analogy between socialization and job orientation. However, the conceptual definition does not address this point directly. It could be argued that orientations are constantly being formed or reformed so that socialization (as defined) does not refer exclusively to early socialization as Etzioni seemed to have intended. Perhaps he should have said "for satisfactory functioning in a (new) role." Certainly this would have established his meaning sphere more precisely.

Another conceptual definition which is perhaps more clearly inadequate is Etzioni's definition of selectivity which he defines as the ratio of actual participants over potential ones. This definition deals only with quantity and not with quality. If 100 unqualified people apply for one job whereas 5 or 10 people (who are qualified) apply for a second job, in which case is the "selectivity" higher?
Clearly, the meaning sphere included in Etzioni's conceptual definition lacks the most crucial aspect (the applicant's qualifications for the role) of the concrete phenomenon it is supposed to reflect.¹

This has been an attempt to stress the importance of assessing the quality of one's concept (and conceptual definition) in terms of its correspondence with reality and the clarity of its meaning. Sociologists have too often, in the past, accepted someone else's definition without questioning either of these points. This causes problems for the next set of criteria to be discussed (i.e., the correspondence between concept and measure). If the measure reflects a faulty concept, then the analysis will be meaningless or misleading. If, on the other hand, the measure is constructed on the basis of reality, ignoring the faulty conceptualization, there would be no basis for making inferences back to the theoretical level. The proposition involving the faulty concept would not be the one that was tested. It should be clear then that the theorist should not just assume that any conceptual definition is adequate.

Furthermore, the adequacy of the conceptual definition should

¹It should be noted that although Etzioni discusses criteria of recruitment in the context, he nevertheless fails to explicitly include such criteria in the conceptual definition for selectivity.
be taken into consideration when one is assessing the correspondence between the concept and the measure.

Correspondence Between Conceptual and Operational Definitions

One of the first points that should be considered when comparing the conceptual and operational definitions is their agreement on the unit of analysis. Do both definitions refer to the same unit? On the basis of the theoretical or conceptual definition alone one could infer that its referent was individual as opposed to organizational. However, Etzioni was not entirely consistent in his theoretical discussion particularly when he discussed the substitutability of organizational selectivity and socialization. At best one must say that the unit of analysis is somewhat unclear in the conceptual definition, although it is clearly on the individual level in the operational definition where the actual concept measured was "perceived early socialization rather than socialization" per se. This shift from organizational referents on the conceptual level to individual "perceived" variables on the operational level is characteristic of most of the variables in the Mulford, et al. (1972) study. This arises in part because of the nature of local civil preparedness agencies (often a one-man outfit) but also because it is easier to talk theoretically about organizational level variables than
it is to measure them. (See the earlier discussion in Chapter Six related to unit of analysis issues).

A second point which can be examined here is the question of **substantive coverage**. To what extent does the indicator cover the meaning sphere of the conceptual definition? As illustrated in Figure 7.2 below, the indicator may be isomorphic or inclusive of the meaning sphere of the concept, it may cover more than the meaning, contain information outside the meaning sphere of the concept, or it may be less precise including both the concept and other meanings. As far as the example of socialization is concerned, the indicator seems to be inclusive of the meaning sphere intended by Etzioni (early socialization) but slightly less than that covered in the conceptual

---

Figure 7.2: Substantive coverage of indicator
A third point of comparison would be whether the definitions describe attributes or variables. An attribute is a property which differs according to what is possessed. There are two types: (a) all or none categories, such as male or not male, alive or dead and (b) mutually exclusive categories such as religious affiliation (Protestant, Catholic, Jew). A variable, on the other hand, is a property that differs according to the amount or degree to which it is present. Some examples of variable properties are age, amount of socialization, level of education and so forth. This distinction is important because it has implications for the type of measure and statistical tests that are appropriate, as well as the statement of the propositions.

Associated with the notion of attribute versus variable is the notion of level of information. An attribute type concept assumes a nominal level of information whereas a variable type concept may be ordinal, interval or ratio. It may at first seem strange to see this discussion in terms of concepts. Generally level of information is referred to as

---

1Note: It was suggested earlier that the conceptual definition might not be completely isomorphic with the concrete phenomenon it was supposed to represent. So, if the indicator includes only the meaning sphere of the conceptual definition, to what extent can we expect the indicator to represent reality?
level of measurement and is confined to the operational measures or indicators. However, there is increasing recognition that the proper level of measurement for an indicator should not simply be an artifact of the measure or procedure used but rather should reflect the way the concept is used theoretically.

As an example, Carter (1971) uses the concept of education. It is a fairly easy task to measure education on an interval level—just ask the number of years of education. The question that Carter raises however, is whether or not the resulting measure has relevance to either theory or the real world? The difference between a 10th grade education and an 11th grade education can not be said to have the same importance as the difference between the 11th grade (where no diploma is received) and the 12th grade (where one is). Whether an individual has a diploma is more important in the real world than how many years he attended high school. Likewise quantitatively 4 more years of schooling does not portray the qualitative difference between a high school graduate and a college graduate. In other words the concept of education is more meaningful substantively as an ordinal variable then an interval one. Thus the indicator for education should reflect the level of information that is implied by the conceptual definition rather than allowing measurement considerations to make the
Borgatta and Bohrnstedt (1972) are making a similar point (perhaps for a different reason) when they emphasize that recent "attention on (developing) nonparametric techniques as an alternative to good measurement has often been misplaced." In this case, Borgatta and Bohrnstedt are pressing for the development of truly interval or ratio indicators to measure concepts specifying amount, degree, proportion, etc. From their point of view sociologists too often settle for developing nonparametric techniques to use with poor measures (ordinal ranking) rather than concentrating on the development of indicators which would truly measure the amounts, degrees, proportions, etc., that many concepts specify.

What is common to both of these articles is the demand that sociologists examine the level of information contained in the meaning sphere (conceptual definition) of the concept and develop indicators with the same level of information. Figure 7.3 describes some of the major characteristics of five levels of information. The first three characteristics have relevance for both concepts and indicators whereas the last three are more directly related to the indicators.

With reference to the example of socialization the level of information is not clearly spelled out by the conceptual definition, although one might infer the notion of amount or
<table>
<thead>
<tr>
<th>Scale</th>
<th>Simple</th>
<th>Nominal</th>
<th>Nominal</th>
<th>Ordinal</th>
<th>Interval</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Naming objects uniquely</td>
<td>Classification of objects</td>
<td>Levels or ranks of objects</td>
<td>Amount or Degree</td>
<td>Rate or Proportion</td>
<td></td>
</tr>
<tr>
<td>Structure</td>
<td>None</td>
<td>Equivalence rel. internal to classes</td>
<td>Linear order</td>
<td>Metric Scales</td>
<td>Metric Scales</td>
<td></td>
</tr>
<tr>
<td>Use of Numbers</td>
<td>As names for objects only</td>
<td>As names for classes only</td>
<td>Any serotonically inc. sequence</td>
<td>Order &amp; dist. prop of numbers</td>
<td>Order, dist. &amp; equal dist.</td>
<td></td>
</tr>
<tr>
<td>Arithmetical Operations</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>$Y = bX$</td>
<td>$Y = bX$</td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td>John</td>
<td>Sex, Religion, denomination</td>
<td>Friendship (Enmey, Acquaintance, Friend, Best Friend)</td>
<td>Performance Scale</td>
<td>Height, Weight, Pop.</td>
<td></td>
</tr>
<tr>
<td>Single Variable</td>
<td>List</td>
<td>Proportions</td>
<td>Median, Quartiles, Zeciles, Quartile Deviations</td>
<td>Means, Medians</td>
<td>Means, Meilarr</td>
<td></td>
</tr>
<tr>
<td>Procedures</td>
<td>PerCentages</td>
<td>Ratio</td>
<td>Standard</td>
<td>Deviations</td>
<td>Standard Deviations</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 7.3: Level of information**
degree from the word "acquisition." This is in fact the inference that is made in the operational definition which measures the amount and quality of job orientation with a summated likert type interval scale.

Another point which might be used for a comparison between the conceptual and operational definitions has to do with dimensionality. A unidimensional concept is one in which a single aspect encompasses the entire range of the meaning sphere. The meaning sphere of a multidimensional concept, on the other hand contains more than one aspect and in many cases is defined in terms of other less abstract concepts. Figure 7.4 below contains two such examples. Socialization is an example of a multidimensional concept which contains various aspects within its meaning sphere. SES (socio-economic status) is also a multidimensional concept. It reduces first to other concepts, then to the meaning spheres within each of them. This multidimensional quality should also be reflected in the measure as illustrated in the figure.

The criteria for assessing the correspondence between concepts and indicators have been summarized in Figure 7.5. The concept socialization has been defined in three different ways in order to demonstrate the various alternatives of the criteria. For instance, one of the definitions for socialization is "whether or not an individual received any
Conceptual Definition

Multidimensional; SES

Unidimensional: Occupation Education Income

Socialization

Job

Understanding of

Orientation General Organizational Obligations Cognitive

Responsibilities

Values goals Perspectivest

Means or tactic Participation

Operational Definition

SES Composite

North-Hatt Highest

Occ. Prestige + Annual + Level

Scale Salary Attained

Socialization Composite

3 measures of adequacy of + Understanding of

2 measures of Orientation Responsibilities

σOcc σSal σEduc

Figure 7.4: Dimensionality of concepts and measures
### Theoretical Level of Abstraction

(refers to concepts)

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Information</td>
<td>Socialization = Nominal</td>
</tr>
<tr>
<td></td>
<td>Whether or not individual received job orientation</td>
</tr>
<tr>
<td>Unit of Analysis</td>
<td>Socialization = Ordinal</td>
</tr>
<tr>
<td></td>
<td>Level of job orientation received</td>
</tr>
<tr>
<td>Dimensionality:</td>
<td>Socialization = Interval</td>
</tr>
<tr>
<td></td>
<td>Acquisition of the necessary skills and orientations associated with effective participation in any organization. Assumption of degree of acquisition.</td>
</tr>
</tbody>
</table>

**Theoretical Population**

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual</td>
<td>Individual</td>
</tr>
<tr>
<td></td>
<td>Social = Orientation</td>
<td>Social = Orientation</td>
</tr>
</tbody>
</table>
OPERATIONAL LEVEL OF ABSTRACTION
(Refers to measures or indicators)

<table>
<thead>
<tr>
<th>ATTRIBUTE</th>
<th>VARIABLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal</td>
<td>Socialization: Received job orientation YES 1 NO 0</td>
</tr>
<tr>
<td>Ordinal</td>
<td>Socialization: Level of job orientation received NONE 0 SOME 1 ADEQUATE 2 EXTENSIVE 3</td>
</tr>
<tr>
<td>Interval</td>
<td>Socialization: Four items: amount and quality of job orientation plus understanding of responsibilities and commitments.</td>
</tr>
</tbody>
</table>

- Probable range of values: 0 or 1
- Range: 0-3

UNIT OF ANALYSIS
- Unidimensional soc = orientation
- Multidimensional two aspects (amt and qual) of orientation plus 2 aspects of skills (knowledge and commitment)

EMPIRICAL (actual data) LEVEL
- Validity
- Reliability
- Properties

Figure 7.5: Correspondence between concepts and indicators
job orientation." As such this definition refers to an attribute (either the individual has or has not received socialization); the unit of analysis is the individual; the level of information is nominal (it exists or it doesn't exist); and it is unidimensional (socialization equals job orientation). These same characteristics are reflected in the responses of the operational measure or indicator. There are two legitimate responses: Yes or No. Thus the indicator is an attribute (since he did or he did not); the level of measurement is nominal (it exists or it doesn't); the unit of analysis is the individual (the question is directed, at individuals who have or have not received job orientation); and the measure is only inquiring about job orientation per se, not types of orientation nor other possible aspects of the concept socialization.

The discussion to this point has dealt with the attribute type definition only. Two types of variable definitions have also been given. The first, an ordinal type definition, views the phenomenon in terms of ordered levels or ranks such that one level can be said to contain more (or less, depending upon the direction) of the characteristic in question than the preceding level. However unlike the interval level of information (which generally assumes degree rather than level) the ordinal type makes no assumption concerning the distance between breaks or levels. Carter's
example of a concept which should be defined theoretically on an ordinal level of information, was education. The difference between 10 years of school and 11 does not operate the same as the difference between 11 years of school and 12 (i.e. graduation from high school). Therefore, the theoretical level of information should be ordinal and theorists should write about education in terms of levels (i.e. 11 years of school or less, high school graduate, some college or technical training beyond high school graduation, received technical diploma, received bachelor's degree and etc.)

The interval type definition assumes that each degree is equidistant from the next. This means that the difference between four socialization units and five socialization units is the same as the difference between one socialization unit and two socialization units.

Another difference between the ordinal and interval type definitions in this illustration concerns dimensionality. The ordinal type definition, given here, like the nominal is only concerned with one aspect of socialization, job orientation. The interval type definition, on the other hand, deals not only with job orientation but also with

---

1Note: This discussion is not intended to mean that ordinal level definitions must be unidimensional, only that this particular one is.
acquiring necessary skills, thus this definition can be said to be multidimensional (rather than unidimensional) since it covers more than one aspect of the concept socialization.

Figure 7.5 also ties the concepts into the previous discussion concerning the interrelationship between unit of analysis, population and sample. The concepts (which are on the theoretical level) refer to characteristics or properties of the units of analysis that fall within the theoretically specified population. The empirical test which in conjunction with the empirical statistical hypotheses (as prescribed by the empirical hypotheses) is performed on data drawn from a sample which is supposed to be representative of the population (and thus could be said to be the operational "measure" of the population in a similar fashion as the indicators or measures were described as an operationalization of the concepts).

It should, however, be emphasized, that the process of operationalization of concepts into indicators occurs entirely in the abstract, even though the indicators are on a much lower level of abstraction than the concepts. One must put those indicators into operation (i.e. gather data from a sample of units) before the theorist can really speak of "empirical" indicators. Thus the process of operationalization (the act of converting concepts to measures) can be extended to imply the act of converting
measures or indicators into data. Similarly it can be said to imply the act of converting a population of units into a smaller, more manageable representative abstract sample of units which must be further "operationalized" before it becomes the actual set of units from which data was extracted.

---

**Figure 7.6: The two levels of operationalization**

Operationalization can be said then to occur on two levels, as shown in Figure 7.6 above. The first level of which takes place entirely on the abstract level before any data has been collected. The indicators are different from the concepts although they are trying to "capture" the same thing (i.e., the true property that they both represent). Likewise, the scale with values that is used in the
statistical testing procedure is different from the abstract indicator. Problems of misinterpreted questions, failure to establish the legitimacy of the data collection, coding errors, and etc., can enter into make the data scale differ from the abstract indicator. These and other problems can also result in differences between the actual and the abstract samples. These topics will be pursued later in greater depth in Chapter Nine, Measurement.

Adequacy of Indicator

In discussing concepts and indicators, two sets of criteria have been discussed so far. The third, adequacy of indicator, will be discussed in greater depth in the chapter concerned with measurement, since most of the criteria require manipulation of data. However, one criterion within this set can be assessed before data has been collected. This criterion is concerned with assessing the correspondence between an indicator (measure) and the real world property it is supposed to represent. It is similar to the notion behind face or content validity since the latter is concerned with the question: "Does it measure what it says it measures?"

Therefore, an attempt to assess the correspondence between indicator and reality before the data is collected should help to assure a good fit with reality (in terms of content validity) once the data has been collected.
Implications

This section of theoretical formulation has dealt with concepts and indicators (or measures). They have been examined together because they attempt to represent the same thing, a real world characteristic or property of a unit of analysis, but on two different levels of abstraction. In the past, theorists have usually concerned themselves with concepts only. This has led to considerable confusion, particularly since even concepts not always clearly specified. When the theorist states explicitly what he/she means by a concept and how that concept should be measured, the confusion can be largely eliminated, and thus allow others to make more direct tests of the theory, while keeping interpretations to a minimum. For instance, Bonjean et al. (1967) list no less than 133 indicators of socioeconomic status, the meaning of which include the following categories: objective composites, subjective and objective composites, occupational, and reputational. Granted that few concepts have as many indicators in usage as this one, the point still remains that the theorist must specify the meaning and measure he/she has in mind in order to avoid the confusion (and often conflicting results) which occurs when a theory is tested with a variety of indicators, not all of which are compatible.
It has already been mentioned that the selection of concepts included in a theory, depends upon prior choices, including among others: the selection of a theoretical perspective, the intended purpose of the theory, the unit of analysis, and even reality itself. Likewise concepts and indicators impinge upon other facets in the theory construction process. The definitions of the concepts and indicators determine the level of information and measurement which in turn influence the type of statistical tests that can be conducted. The type of indicator chosen (sensitizing versus an index) have implications for the data gathering technique that will be employed.

Concepts are supposed to represent abstractions of real world properties. As such they are not the properties themselves—just sense impressions of the properties. However the more accurate the sense impression, the more valid the inferences based on those impressions will be. Care must be taken therefore to establish a clear correspondence between the concept, indicator and real world property. Inferences are made on the basis of data analysis results and are concerned with the linkages between sets of concepts. Thus statistical tests do not prove or disprove, validate or invalidate concepts or indicators. However, the assumption is made that a strong link does in fact exist between the concepts, indicators and real properties. And, as with all
assumptions, the violation of this assumption affects the amount of confidence one can place in the inferences that have been made. In other words, if one of the concepts in a hypothesis is poorly chosen or defined (such as the selection of education to represent socio-economic status), no matter how well one measures education, inferences made on the basis of say a correlation test will not adequately describe whether a relationship exists between socioeconomic status and the other concept. It might reflect a relationship between education and the other concept, but not SES per se since education can be said to represent only one aspect of the overall concept SES. A breakdown in the link at either point in Figure 7.7 causes a subsequent loss of

---

![Diagram](image)

Figure 7.7: Two points at which breakdowns occur in the link between concepts, indicators and real world properties

credibility in the inferences that can be made from the test results to the general hypotheses and the theory.
Summary

This chapter has been concerned with delineating the importance of a clear and concise specification of concepts and measures. It began with a discussion of what is meant by the terms "concept" and "measure or indicator" as well as how they are related to each other and to reality. It was stated that conceptual and operational definitions contain the meaning sphere of concepts and measures and that a clear explication of those meanings is necessary if the remainder of the theory construction process is to have any meaning. Therefore, since these definitions are so crucial, a number of criteria were discussed by which these definitions could be assessed, in terms of their relationship to each other (i.e. the conceptual definition with the operational definition) and to the real world phenomenon they are supposed to represent.

Considerable attention has been given to the content in this chapter as well as the next because the concepts and linkages represent the basic elements of a theory. Concepts have, in fact, been described by many as the "building blocks" of a theory. As such the theory building process requires that they be carefully selected and defined. This analogy can, however, be carried one step further to illustrate the importance of linkages and their relationship
to both the concepts and propositions of a theory. Thus, if concepts serve as the building blocks of a theory, then the linkage words must surely represent the mortar which binds or connects those blocks (concepts) together to form the propositions and hypotheses of a theory.
CHAPTER EIGHT:
LINKAGES, PROPOSITIONS AND THEIR SYSTEMIZATION

This chapter will be concerned with how concepts (and measures) combine to form the propositions and hypotheses of a theory. Special emphasis will be placed upon specifying the nature of linkages since, like concepts in the previous chapter, they occupy a crucial position in the development of a theory. The notion of propositions will also be introduced, in terms of the various types of propositions (and/or hypotheses) and in terms of differing methods of systematizing those propositions into a theory.

Linkage words are concerned with capturing the nature of the relationship between concepts. The linkage chosen should reflect the actual relationship that exists between the real world phenomena that the concepts represent. Traditionally this has not been emphasized in sociology. Thus the most common linkages stated little more than the fact that a relationship was thought to exist. I hope to demonstrate in this and succeeding chapters that more care must be taken in forming propositions and hypotheses to specify the exact nature of the relationship to be tested.
Stating the Nature of the Relationship

Language is perhaps more important than many social theorists have recognized. Linkage words in hypotheses are supposed to state the exact nature of the relationship between concepts. However, many propositions, hypotheses and theories never get more specific than to state that a "relationship exists," even though the theorist may actually have a far more specific linkage in mind. The following hypotheses all deal with a relationship between socialization and effectiveness (role performance), only the linkage words are different.

1. Socialization is related to effectiveness (role performance).

Note: what does this tell us about the expected relationship? Which, if any, of the following graphs or diagrams satisfy this proposition?

\[\text{Graphs A, B, C, D}\]
2. Low socialization is more associated with low role performance than high socialization.

<table>
<thead>
<tr>
<th>Socialization</th>
<th>Role Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Low</td>
<td>More</td>
</tr>
<tr>
<td>High</td>
<td>?</td>
</tr>
<tr>
<td>?</td>
<td>More</td>
</tr>
</tbody>
</table>

How much do we really know about the relationship (linkage) between these two concepts?

3. There is a positive relation between socialization and role performance.

How is this different from the two hypotheses above?

4. If socialization increases then role performance will also increase.

The graphs for Hypothesis 3 and 4 are similar but Hypothesis 4 is more explicit. Why? Because it not only proposes a positive relationship (as in Hypothesis 3) but it also designates socialization as the determinant (independent or explaining) variable and role performance as the result (dependent or explained) variable. As such it implies, at least faintly, a causal type relationship.
5. Role performance is a linear function of socialization, the slope of which is significant.

\[ RP = f(Soc) \]
\[ RP = \beta_0 + \beta_1 (Soc) + \xi_i \]

Certainly the degree of explicitness has increased in this hypothesis, as well as the degree of implied causality.

6. Role performance increases 5 units for each unit increase of socialization.

This hypothesis is attempting to state the nature of the relationship between the two variables quantitatively, that is, to spell out the extent to which a change in socialization will affect role performance.

A number of criteria can be examined in the course of selecting the proper linkage term(s). I shall discuss three of these: (1) qualitative versus quantitative (2) specifying the type of linkage and (3) linear versus nonlinear relationships. The kind of linkage chosen will, of course, have an effect upon subsequent activities. Careful attention to issues such as those listed above, will help to prevent mistakes in the other phases.

The six general hypotheses listed above all refer to a relationship between socialization and role performance, however, the specificity concerning the nature of that
relationship tends to increase as one moves from the first to the last hypothesis. In the case of the first hypothesis, little or nothing is gained in terms of information regarding the specific nature of the relationship. To state a hypothesis in this form, then, is to say nothing either qualitatively or quantitatively. On the other hand, Hypothesis Six states the exact (quantitative) nature that the theorist expects to find between the two variables.

**Qualitative vs. Quantitative**

In recent years a number of writers have begun to make a distinction between qualitative and quantitative propositions or hypotheses. Actually it is the linkage that largely determines whether a proposition is one or the other. According to Coleman (1964) a qualitative proposition is one whose linkage tells how or why one concept (and/or indicator in an empirical hypothesis) is related to or influences another. The linkage in a quantitative proposition, on the other hand, tells the extent to which one concept (or indicator) influences another. Which of the above hypotheses then could be called quantitative using Coleman's distinction?

Most sociological hypotheses today tend to be written in the form of 3 and 4, which Coleman would designate as qualitative linkage. However, more explicit relationships (i.e. more quantitative linkages) such as those in Hypotheses
5 and 6 are often implied. I will attempt to demonstrate, in a later portion of the chapter that this is due in part to the fact that many sociologists have been taught to think in terms of covariation rather than causation.

Covariation certainly has its place in examining social reality. It is, nevertheless, insufficient to portray all types of social relationships. As Rhoads (1971) and Abell (1971) point out, social theorists must learn to specify the linkages in their propositions and hypotheses so that they correspond to the real world linkages they supposedly represent. In other words relationships may be covariational or causal, linear or nonlinear, simple or complex, and it is up to the theorist to specify which, depending upon the type of linkage he/she chooses.

Types of linkages

Practically every writer of a theory construction text has proposed a set of criteria for classifying linkages. I shall discuss Zetterberg's because it is generally thought of as the first definitive set and the one which has most often been elaborated on.

Zetterberg (1966) has outlined a set of five criteria which can be used to classify propositions and hypotheses based on the type of linkage that has been used. These criteria will be discussed specifically in terms of Hypothesis Six above, although general comments will be made
relating to the other five hypotheses, where they are appropriate.

**Hypothesis 6:** Role Performance increases 5 units for each one unit increase of Socialization.

**Criterion 1:** **Reversible** vs. **Irreversible**

- If X then Y;
- if Y then X.
- If X then Y;
- If Y no conclusion.

The linkage in this hypothesis would be classified as irreversible, since the implication in the hypothesis is that a change in socialization occurs first and is followed by a change in role performance. (In most of the organizational literature, socialization is considered to come early in one's career, prior to role performance.) Because of the specificity of this particular linkage, the irreversible designation has been made, however, this would not necessarily be true of Hypotheses 1, 2, or 3. They might as easily be classified as reversible, since the linkage is not clearly defined.

**Criterion 2:** **Deterministic** vs. **Stochastic**

- If X, then always Y.
- If X, then probably Y.
One must be careful in designating a proposition as deterministic. The word itself carries pejorative overtones, at least in sociology. In fact, Zetterberg makes the statement "Deterministic relations (linkages) seem very rare in sociology" (1966:70). Judging by his discussion, the key words to look for in deciding upon deterministic vs. stochastic seem to be terms such as always vs. sometimes, likely, or etc. Although Hypothesis 6 above contains neither of these types of terms, the language does seem quite explicit "1 unit increase in socialization leads to a 5 unit increase in role performance." It is probable then that this proposition, as stated above, would be classified as deterministic. (One should remember, however, that this proposition was constructed for purposes of illustration, only, and it is unlikely that it would find any empirical support in the real world). On the other hand, Hypothesis 5, which is similar in meaning to 6, though not quite as explicit, would probably be classified as stochastic.

This example illustrates some of the problems involved in attempting to classify a single proposition. Generally propositions represent condensed statements summarizing pages of theoretical discussion. As such it is extremely difficult to classify propositions without considering the theoretical context in which they were developed. For instance,
Hypothesis 5 would more clearly fit the stochastic classification if placed within the background of what Etzioni (1975) has designated as the Iowa State Studies. It would then be clear that socialization, along with other variables, can be used as a predictor of role performance, although only to a degree; and certainly not to the degree specified in Hypothesis 6, nor that required by a deterministic linkage.

Criterion 3: Sequential vs. Coextensive

If $X$, then later $Y$.

Hypothesis 6 could be designated as having a sequential type of linkage for at least two reasons. The first concerns the language chosen in the linkage, i.e., for each unit increase in socialization there will then be an increase of 5 units of role performance. Note that the word "then" is not actually used but seems to be implied, particularly in view of the second reason, namely that theoretically, socialization is generally regarded as coming prior to role performance in one's career.
Criterion 4: **Sufficient** vs. **Contingent**

- If X then Y, regardless of anything else.
- If X then Y, but only if Z.

Because of the explicit nature of the linkage, Hypothesis 6 will be designated as a sufficient statement.

A possible example of contingent linkage can be provided by a statement taken from Etzioni (1961) which proposes that the effect of communication on role performance depends on the scope and effectiveness of the socialization process. Thus one should find that communication has an impact on role performance but only if socialization meets certain conditions, (if communication then role performance but only if socialization).

Criterion 5: **Necessary** vs. **Substitutable**

- If X and only if X, then Y.
- If X then Y but if Z then also Y.

The designation of Hypothesis 6 on this criterion is difficult, just as it was on criterion 2. It does not contain the keywords that one associates with either of the choices, namely "socialization and only if" nor "but if something else then also" role performance. In terms of placing the statement (or rather Hypothesis 5) in a context such as the Iowa State Studies, mentioned earlier, one would
choose substitutable since socialization was found to be only one of a number of variables which can affect role performance.

Using Zetterberg's criteria, then, Hypothesis Six can be said to reflect an irreversible, deterministic, sequential, sufficient, and substitutable relationship (linkage).

As pointed out above, however, hypotheses and propositions are sometimes difficult to classify according to this criteria so that one sees little real application being made of this or similar delineations. In practice, the above generally boils down to a distinction between covariational vs. causal relationships where the criteria outlined by Zetterberg can be easily subsumed under these headings as illustrated in Table 8.1 on the next page. Note that in doing so a distinction is made between two views of causality, the traditional (deterministic) approach versus a relaxed (contingent) view.

**Classical Deterministic Causality**

The deterministic brand of causality which gained prominence during, and for some time after, the era known as the scientific revolution, was based on the notion that everything follows natural laws (and is therefore determinable). Nothing arises out of nothing. Everything must therefore be the result of some cause. The essential elements of this conception of causality were described as
Table 8.1: Relationship between Zetterberg's criteria and the covariation vs. causality framework

<table>
<thead>
<tr>
<th>Zetterberg's criteria</th>
<th>Type of Linkage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covariation</td>
<td>Relaxed View</td>
</tr>
<tr>
<td>Reversible</td>
<td>Irreversible</td>
</tr>
<tr>
<td>Stochastic</td>
<td>Stochastic</td>
</tr>
<tr>
<td>Coextensive</td>
<td>Sequential</td>
</tr>
<tr>
<td>---</td>
<td>Contingent</td>
</tr>
<tr>
<td>Substitutable</td>
<td>Substitutable</td>
</tr>
</tbody>
</table>

necessary (E will be produced only on the condition that C is present); unique (for any one C there can be only one result, e); asymmetric (dependence of effect, E, must be strictly one-sided, i.e., if C causes E, E cannot "cause" C); invariability (C will always cause E); sufficient (C's being present is adequate for E's being produced); productivity (that which creates, as in the statement "cause (is) that by whose action an effect is born"). Written in propositional form, classical causality can be stated as follows:

"If C happens, then (and only then) E is always produced by it" (Bunge, 1959:47).

If we compare these criteria to the ones specified by
Zetterberg we find that both sets contain necessary and sufficient elements; the asymmetric criterion above corresponds to Zetterberg's irreversible, and productivity is fairly close to Zetterberg's sequential. The unique and invariability aspects can be compared to Zetterberg's deterministic criterion. Thus theories which can be classified as irreversible, deterministic, sequential, sufficient and necessary, according to Zetterberg's criteria, can be described as causal (in the classical deterministic sense).

It should be noted, however, that few, if any, relationships (linkages) in the real world could meet this set of criteria which is precisely why it has fallen into disuse. As far back as the 1890's empiricists such as Hume expressed growing skepticism with the ability of science to establish empirical connections between cause and effect. The following comment by Hume provided the essence of the empiricist's argument against what has been called deterministic causality.

"When we look about us towards external objects and consider the operation of causes, we are never able in a single instance to discover any power or necessary connection; any quality, which binds the effect to the cause, and renders the one an infallible consequence of the other. We only find that one does actually, in fact, follow the other...consequently there is not, in any single, particular instance of cause and effect, any thing which can suggest the idea of power or necessary connection" (Hume, 1894:24).
Thus Hume and his followers never rejected the actual existence of cause and effect relations, only man's ability to demonstrate the necessary empirical connections between them.

The Empiricists' View

"The causal connection is not given in sensory experience but is supplied by the experiencing mind. Causality, therefore, is not a fact of the world, having ontological status but is rather a contribution of the mind and thereby has only epistemological status" (Rhoads, 1971:27).

What is observable then is succession of facts, sequences of antecedents, and consequents. The why or connection between things can only be a mental construct or a metaphysical entity which unfortunately is outside the realm of man's ability to establish conclusively. This inability "strikes at the heart" of the necessary and sufficient aspects of the earlier description of classical causality. The ascendence of quantum physics with its emphasis on indeterminancy, chance, or probability, directed its attention toward the uniformity aspects of causality. Thus one was supposed to speak in terms of probability or chance rather than uniformity (i.e., for any one C there is an X probability that the result will be E). The requirements of antecedent causes (or the asymmetric characteristic) was said to be inadequate to deal with either reciprocal relations between variables, multiple antecedents, or joint effects all
of which had come to replace the former unidirectional assumption of single cause, single effect. To reiterate Hume, then,

"The original idea of necessary connection, since it does not arise from a single impression must arise from a reputation of similar instances which produces the habit or customary transition of the imagination from one object to its usual attendant. One comes to feel the events to be connected though they are not, for all that, shown to be actually connected" (Madden, 1969:73).

The best one can hope for according to this point of view, then, is to establish a sequence of events. For this reason, cause became defined as "an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second" (Hume, Treatise of Human Nature). This is the view to which later associationists such as Pearson, would ascribe in their emphasis on covariation rather than cause, (which almost disappeared in terms of explicit use). But there were those who held views on the other side of the argument.

Immanuel Kant, a contemporary of Hume, feared that Hume's views would provide a death-blow to science. He argued that:

"No empirically conditioned practical principle can serve as the basis of a law, for experience can never be the source of universality and necessity. A law, whether it is practical or theoretical must originate, if it is to originate at all, a priori, in pure reason" (Bassart, 1968: 337).
Thus he felt that the link through reasoning, between cause and effect in the real world could be discovered through the building of deductive systems. Kant, and others after him, recognized that the ubiquity and importance of causality, not only for scientists and philosophers, but for the "common man on the street," can be seen in the vast number of verbs for causation, that are present in every day language, e.g., "to push, to bend, to corrode, to cut, to make, to ignite, to transport, to convince, to compel, to remind, to irritate, to influence, to create, to motivate, to stimulate, to incite, to mislead, to induce, to offend, to effect, to prevent, to facilitate, to produce, and etc., (Ducasse, 1966:141). What was needed then was a revision of the notion of causality, a revision devoid of determinism, but one which still allowed for causal connectives.

A Pragmatic Revision

It has already been stated that Kant attempted to counter Hume's skepticism by advocating the building of deductive systems as a way of scientifically capturing the notion of cause. A more current attempt of Hubert Blalock (1962) seeks to combine to some extent the notions of Hume and Kant. Blalock accepts Hume's supposition that cause can never be empirically proven. However, he agrees with Kant that science needs the notion of causality if it is going to develop theory. Thus he postulates that cause rests in the
conceptual model rather than reality and can therefore, be verified (though not proven) through empirical techniques. Francis (1961) also speaks to the notion that causality is a necessary aspect of science. However, he advocates that sociologists dismiss the metaphysical questions raised by Hume and Kant, and simply treat cause as an undefined but necessary tool of science. Peter Abell has constructed a third alternative which he calls a "relaxed notion of causality":

"Cause is a T-concept. Causal connectives are not observed between variables, they are inferred from observations or, rather more precisely, certain sets of observations are compatible with a particular causal account" (Abell, 1971:116).

This notion of "relaxed causality" as outlined by Abell corresponds to one suggested by John Rhoads (1971). According to Rhoads, the revised notion of causality which is gaining momentum has the following characteristics:

1. Cause is ontological, although it may also have some epistemological implications.

2. The causal nexus is characterized by production which distinguishes it from noncausal connections. Thus causality does not seem to be exhaustive of all scientific knowledge.

3. Statements of causal relationships are conditional i.e., they state that an effect will occur when the causal conditions appear.

4. The notion of probability replaces the previous notion of uniformity.

Stated in propositional form: if C then probably E will be
produced. NOTE: as stated above, the proposition assumes an isolated or closed system (no disturbances from other factors, i.e., "all things being equal"). Accordingly, the "cause" of an event is generally neither a sufficient nor an absolutely necessary condition for the events occurrence, rather it is what may be described as a contingently necessary cause (Nagel, 1961). However, in sociology as well as in the other social or even natural sciences, we seldom encounter uncontaminated relationships. This can explain why we often get differential results or less than one-to-one prediction from our theory (Rhoads, 1971), but it need not vitiate the utility of causality for development of sociology and/or the social sciences.

On the Question of Determinism and the Notion of Causality

The notion of classical cause, rightly or wrongly, generally brings to mind the notions of necessity, force, and determinism based on univeralistic laws. In its extreme conception, this view resembles one described by Matson as "...the universal pattern of rigid determinism, or mechanistic causation where behavior is a pawn of forces set in motion at the beginning of time, where the concept of freedom was no more than self-deception" (1964:14). In a less extreme, but no less determined, description of this view;

"Surely if we can say an individual's action is caused, we imply it is determined and thereby restrict the notions of freedom and rational
decision in human affairs; the individual could not have done otherwise!" (Abell, 1971:130).

Thus the notion of classical causality somehow becomes entwined with the notion of men as machine, one who performed not by his own choice or reason but as the result of the "structure of the atoms that form him, and their motion (which) propels him forward" (Matson, 1964:14).

The relation between cause, reason and freedom became one of Kant's central preoccupations. In the critique of pure reason Kant argued that freedom and natural causality do not necessarily conflict. It is not cause which must be avoided but rather the notion that causality through freedom requires a freedom that is absolutely unconditional. "But (the unconditional can never be found) since objects are given to us only through the mediation of sensuous intuition, all sensuous objects are (thus) subjects to natural causality" (Bassart, 1968:337). To be totally free in Kant's mind then is to have no freedom at all! That would mean that social and natural forces would operate at random. Choice, particularly rational choice, would not exist. What is free then to Kant is the choice between alternatives, not the freedom to choose what those alternatives will be. There need be no basic incompatibility then between causality and freedom, when regarded in this more relaxed notion of a causality freed from determinism.
This is the conception which recent formulations of causality have adopted, and which I have labeled "a relaxed notion of causality" as originally coined by Abell.

"To show that an action is caused does not imply that the actor could not have done otherwise. What it does show is that given a set of initial conditions (usually beliefs, values, etc.). The actor does, as a matter of fact, adopt certain purposes, frame intentions which he implements in behavior. Indeed if we recognize our behavior as caused in this way, then it enables us to alter it; it enhances the idea of freedom and rational decisions" (Abell: 1971:130).

Herbert Simon has also commented on the move back to a reformulated causality.

"The viewpoint is becoming more and more prevalent that the appropriate scientific model of the world is not a deterministic model but a probabilistic one in which causality is reconceptualized as probability rather than necessity" (Lerner, 1965:6-7).

If we return to the formulations of causality described earlier, it is clear that the notion of determinism, certainly in its extreme sense as outlined by Matson, has been replaced with the notions of conditionality (probability). When viewed as contingent rather than absolute, causality becomes free of its deterministic qualities and resumes its useful place as a tool for describing social phenomena.

In his discussion on causality in the social sciences, Rhoads cautions that not all relationships are causal, although it is his contention that "relations of cause and
effect make up an important segment (of the universe studied in the social sciences) and must therefore be identified if understanding (an explanation) of these types of phenomena are to be achieved". This same point of view has been expressed by others such as Nagel (1961), Bunge (1959), and Abell (1971).

"For some philosophers, intentions and actions are logically related and it would be a mistake to seek a causal connection between the two. But note that this interpretation is still open to a causal explanation of why people have the intentions they do. So we could ask for a causal explanation of their action in the sense that action means intended behavior" (Abell, 1971:127).

He goes on to state that physicists are often able to ignore causal accounts but:

"The social scientist, however, cannot avoid such questions; whereas the physicist can remain content with kinematics the social scientist is much more impelled in the direction not of social kinematics but social dynamics. This is because causal links in the social domain are made by man. A failure to recognize this is a sort of reification" (1971:131-132).

It should be obvious from the above discussion that scientific explanations in sociology (and the social sciences) encompass both causal and noncausal explanations and that a complete understanding of social phenomena requires the formulation of both. It is up to the theorist to observe those relationships (linkages) in the real world and then fashion his propositions and/or hypotheses accordingly. The real world should therefore be the guide
rather than following a preconceived formula that supposedly applies to all situations. If the theorist approaches reality from a predetermined formula, reality may become defined in terms of the approach rather than the other way around (what you see is what you get). Nowhere has this been as true as it has in terms of specifying the form (linear vs. nonlinear) of a relationship. The assumption of linearity has become so routine that it usually remains not only unstated but also unexamined.

**Linear vs. Nonlinear Relationships**

Most of our thinking and propositional writing in sociology is in terms of linear (points fall on a straight line) relationships. All of the example hypotheses above would fit this type of thinking. However, there is no reason to believe (in fact, there is considerable evidence to the contrary) that all relationships in the real world are linear. We as sociologists need to begin to recognize this fact, by testing our assumptions of linearity and by deriving new methods of thinking or stating our propositions (hypotheses), and of testing them in terms of possible curvilinear relationships.

**Types of nonlinear relationships**

There are probably as many different types of nonlinear relationships as there are curves. A few examples will be discussed in terms of the type of curve it represents and an illustration
of its possible application.

A. Accumulation effect. There are some things (such as inter-group conflict) which we tend to tolerate in our social structural environment, as long as it remains at a relatively low level of intensity. At this level it has no influence on our behavior. Thus a group can "stick together" even though members, as individuals, may not always agree with one another. However, day-to-day aggravations eventually add up and at a certain point or range, the group finds it must do something (e.g., expell a trouble-maker) or break-up. Troubled marriages often exhibit this type of relationship between the husband and wife. The famous saying "The straw that broke the camel's back" is a vernacular expression for this kind of relationship.

B. Saturation effect. This is characterized by the situation occurring in a number of newly oil-rich Arab countries. The governments of these countries complain that money invested in development reaches a saturation point beyond which the country can no longer absorb any more benefit (in terms of development) no matter how much more money is invested. Others have described this kind of relationship in terms of a
threshold effect.

C. Both accumulation and saturation effect in the same model. This is a case where a relationship holds, but only in the middle range of the independent variable(s). An example of this type of relationship, according to some writers, is that which exists between Role Performance and Job Satisfaction. It has been said that low levels of job satisfaction have an inhibiting effect on role performance. A dissatisfied worker is not likely to care about improving his/her performance. Likewise, though, a highly satisfied worker, may not feel any compulsion to exert him/her self either. They are satisfied with what they have and therefore the stimulus for achieving more is missing. Thus only within the middle range of job satisfaction, will a change (positive or negative) in job satisfaction affect role performance.

D. Reversion effect. This type of relationship is similar to a parabolic curve and will be found in reality as portrayed here or \( \sqrt{ } \). It was found in the 1950's that large families were characteristic of those in both low socio-economic status and high socio-economic status, while median socio-economic status families tend to have fewer
children. The higher economic status families could afford
to have more children and the low status families couldn't
afford not to. There are of course many other examples which
might have been given but this should suffice to alert the
reader to the possibility of nonlinear relationships and to
the need for examining the traditional assumption of
linearity. (Procedures for recognizing nonlinear
relationships will be discussed in a later section on
preliminary analysis procedures.) The problem with ignoring
or not recognizing a nonlinear relationship is that the
result of a linear test on a nonlinear relationship usually
leads to a false conclusion.

Implications

The linkage term states the nature of the relationship
between the concepts in a proposition or hypothesis. Careful
specification of that linkage is therefore extremely
important. The linkage chosen must reflect the actual
relationship which is thought to exist in the real world. If
the wrong linkage is chosen, the theorist will inevitably go
to a lot of expensive trouble to prove nothing. Or what is
worse, proclaim evidence of a nonrelationship where a
relationship (different from that hypothesized) actually does
exist. This is particularly true when linear linkages are
mistakenly hypothesized for relationships that in reality are
nonlinear.
Even when a relationship in the real world is linear, sociologists have had a tendency in the past to use vague, noncommittal terms as linkages. What does it mean to say "There is a relationship between socialization and role performance"? There should be a direct relationship between the theoretical linkage, the operational linkage and the statistical linkage (chi square, correlation, slope or etc.) that is actually tested. To test this proposition one would probably have to confine it to a chi-square (which tests no more than that there is a relationship). But a great deal more is known about the relationship between these two concepts. This proposition is therefore terribly inadequate. If we hope to construct sound sociological theory, theorists are going to have to "stick their necks out" and hypothesize the linkages that they see in the real world.

It has been said that sociologists are too heavily influenced by authorities. This is often the reason that is given for the continued use of vague linkages such as the one in the example above. The big names in sociology (many of whom were writing at an earlier time when less was known and measurement and statistical techniques were less sophisticated) such as Merton, Parsons, Etzioni and etc., rarely, if ever, used linkages that posited much more than a directional relationship. Thus, although in most cases knowledge and techniques have improved beyond this level.
social theorists, to a great extent, still follow the lead of these "big names" and write their propositions in terms of vague nonspecific linkages.

I might suggest that perhaps this is one of the reasons many sociologists have been content with the gap between theory and methods. The theory has really only been "window dressing" for some researchers. What really mattered were the indicators (not the concepts) and the null hypothesis. The rest was just padding. This dissertation is based on the assumption that good solid theory will only be constructed when theory and methods are brought together in an integrated fashion. When that happens, the linkage chosen by the theorist will determine the kind of statistical testing that will be done in the analysis phase of constructing the theory.

Much of what is often discussed under the heading of propositions and hypotheses in other textbooks has been discussed in terms of linkages. An emphasis has purposely been placed on the parts of a proposition (linkages and concepts) rather than on the propositions themselves. The primary reason for taking this approach is to get theorists to focus more attention on these parts, particularly the linkages since they seem to be the aspect which is most often ignored or overlooked. There are however, a number of issues which pertain to the propositions (and hypotheses) as a
whole. These will be discussed in the next section.

Propositions and Hypotheses

A proposition can be described as "a statement about the relationship between two or more concepts. It connects concepts in relational form" (Lin, 1976). The basic elements of a proposition therefore consist of at least two concepts and one or more linkage terms which state the nature of the relationship between those concepts.

The term proposition is often used interchangeably with the term hypothesis. This has led to a certain amount of confusion since other writers use the term proposition for the abstract level of analysis and hypothesis for the empirical level of analysis. In order to eliminate this type of misunderstanding, Table 8.1, on the next page, contains the terms that will be used throughout this dissertation. These have been listed according to type.
General Hypothesis: Socialization is positively related to Effectiveness

Theoretical Statistical Hypotheses: Null: $\rho_{\text{SOC} \cdot \text{RP}} = 0$
Alternative: $\rho_{\text{SOC} \cdot \text{RP}} = 0$

Empirical Hypothesis: Socialization scores will be positively related to Effectiveness scores

Empirical statistical test (Decision Rule):
Null: $r_{\text{SOC} \cdot \text{RP}} = 0$
Alternative: $r_{\text{SOC} \cdot \text{RP}} \neq 0$

---

Figure 8.1: Suggested format for presenting hypotheses

In the past, theorists usually placed their general hypotheses or propositions together, their empirical propositions together, and their statistical hypotheses and tests (which have sometimes been written in one form or the other rather than both) together. It will be suggested here, however, that continuity will be easier to obtain between the levels or types if the theorist integrates them into sets of

---

1This format was originally presented by Dr. Richard Warren in a class on causal model analysis. Dr. Warren is currently a professor in both the sociology and statistics departments at Iowa State University, Ames, Iowa.
statements such as the example in Figure 8.1.

In essence what this arrangement represents is an attempt to find a more meaningful way of presenting one's hypotheses. Thus the actual statistical test which is calculated on the basis of sample data is the one specified by the empirical statistical test, which in turn depends upon the measures (or indicators) and the relationship (or linkage) specified in the empirical hypothesis. Direct tests are never made on either the general hypothesis or the theoretical statistical hypothesis (unless the sample and the population are the same). Rather, they are inferred on the basis of the results of tests specified by the empirical statistical test. The theoretical statistical hypothesis in turn depends upon the concepts and relationships (linkage) specified in the general hypothesis. As such the empirical hypothesis represents the operationalization of the general hypothesis while the empirical statistical test represents the operationalization of the theoretical statistical hypothesis. Figure 8.2 illustrates the direction of flow from one hypothesis to the next, in terms of the discussion above. The reader should note in particular the linkage in the general and empirical hypotheses, i.e., "is positively related to"). The statistical hypotheses and statistical test reflect this directional aspect of the linkage by predicting
that the relationship, (in the statistical hypothesis and \( r \) in the test) will be positive (greater than zero). Note: this set of hypotheses and test will be examined in greater depth in Chapter Ten.

A final comment on propositions concerns the difference between highly theoretical general propositions (hypotheses) and ordinary propositions (hypotheses). Ordinary propositions are on a lower level of abstraction than general propositions. As such they are more tied to specific times and space and are thus considered to have lower informative value, where informative value depends upon the variety of events which can be accounted for by a proposition (Zetterberg, 1966).

Theories contain sets of interrelated propositions (general hypotheses) rather than single propositions. Ordering those propositions according to some sort of
systemization makes those interrelationships more apparent and, according to one source (Zetterberg 1966), easier to verify. The next section will consider a number of alternative methods for ordering propositions within a theory.

Systematization of Propositions

Propositions can be ordered on basis of either the independent variable or the dependent variable, arranged in a matrix; placed into an axiomatic form; or depicted in a causal model framework. I shall discuss each of these alternatives briefly in the remainder of this chapter.

Ordering on the basis of the Independent Variable

Suppose the Civil Preparedness Agency described in Chapter Four decided to up-grade their socialization program. One of the questions they would be most interested in answering would pertain to the ultimate effect of such a change. In other words what are the variables which would most likely be affected by the change. They might also wonder what the effects would be if they increased the amount of communication at and between the various levels within the
SOCIALIZATION (JOB ORIENTATION) IS RELATED TO

SCOPE (COLLEAGUE INTERACTION)

JOB SATISFACTION

ROLE PERFORMANCE

PERVASIVENESS (PRESSURE TO CONFORM)

COMMUNICATION IS RELATED TO

PERVASIVENESS (PRESSURE TO CONFORM)

JOB SATISFACTION

TENSION

ROLE PERFORMANCE

Figure 8.3: Ordering of propositions on the basis of the independent variable

organization. In order to answer questions like these the theorist might want to order his/her propositions on the basis of similar independent variables, such as those in Figure 8.3. A somewhat similar approach can be taken with the dependent variable.
Ordering on the basis of the Dependent Variable

In Chapter Thirteen, I shall examine a causal model with two dependent variables. Another approach which could have been taken, consists of ordering propositions on the basis of the dependent variables (job satisfaction and role performance) and examining the propositions separately, (see Figure 8.4).

Figure 8.4: Ordering on the basis of the dependent variable
Matrix approach to ordering propositions

It is possible that the theorist might want to examine the interrelationships of a group of variables. The matrix approach, as illustrated in Figure 8.5, offers a suitable framework for this, particularly in a correlational or covariation framework. Note: The designation of dependent and independent variable is not really meaningful in this case although Zetterberg (1966) uses the terminology. I am suggesting that a matrix approach be confined to a covariational framework rather than a causal one (as implied by Zetterberg's use of determinant and result) since it would not make sense substantively for a result to cause its determinant except in a feedback type of situation in which case of course, we would be dealing with a different variable.

(i.e., SOC → COM and COM → SOC)

Actually, the feedback situation would represent what Zetterberg has described as a Chain pattern of propositions (which we will place within the causal model approach). He makes the statement that the matrix approach is not restricted to sequential propositions, but, as was illustrated above, this statement only makes sense if one is
<table>
<thead>
<tr>
<th></th>
<th>Soc</th>
<th>Com</th>
<th>Sco</th>
<th>Perv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socialization</td>
<td>---</td>
<td>Com•Soc</td>
<td>Sco•Soc</td>
<td>Perv•Soc</td>
</tr>
<tr>
<td>(Soc)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>Soc•Com</td>
<td>---</td>
<td>Sco•Com</td>
<td>Perv•Com</td>
</tr>
<tr>
<td>(Com)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scope</td>
<td>Soc•Sco</td>
<td>Com•Sco</td>
<td>---</td>
<td>Perv•Sco</td>
</tr>
<tr>
<td>(Sco)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pervasiveness</td>
<td>Soc•Per</td>
<td>Com•Perv</td>
<td>Sco•Perv</td>
<td>---</td>
</tr>
<tr>
<td>(Perv)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 8.5: Ordering of propositions by means of a matrix
using the matrix approach in terms of a covariational framework.

The number of possible propositions can get quite large in some theories so that the listing techniques above can get extensive. The next approach to the systematization of a set of propositions offers a method of reducing the overall possible set to a minimum number of independent propositions, called axioms.

**Axiomatic form**

Reynolds (1971) has defined an axiomatic theory as one which contains the following items:

1. A set of definitions for both concepts and measures
2. A set of existence statements that describe the conditions under which the theory operates
3. A set of relational statements, composed of axioms, independent statements from which propositions are deduced
4. A logical system to relate all concepts and deduce propositions.

Using the same set of concepts and data that was described by the two previous chapters, I shall attempt to demonstrate this type of systematization. Criterion one was discussed in Chapters Six and Seven. In terms of existence statements (criterion 2), I will for the moment, only say that the theory concerns individuals who work in normative type of organizations, as described by Etzioni (1961, 1975).
normative type of organizations, as described by Etzioni (1961, 1975). [There are of course other conditions which could be considered such as the possibility of moderator variables, (variables which serve as the basis for meaningful subgroups within a population of units and which cause differential results in terms of reliability and statistical analysis), however the topic of moderators will be touched on in later chapters and therefore this should suffice for the purpose of illustrating existence statements].

The third criterion is a set of axioms from which propositions can be deduced. Let us consider the following set of propositions (which, for the moment will be considered axioms):

A. COMMUNICATION $\rightarrow$ SCOPE
E. SCOPE $\rightarrow$ PERVASIVENESS
C. PERVASIVENESS $\rightarrow$ ROLE PERFORMANCE

The next step according to this approach consists in deducing new propositions such as

IF 1. SCOPE $\rightarrow$ PERVASIVENESS
AND 2. PERVASIVENESS $\rightarrow$ ROLE PERFORMANCE
THEN 3. SCOPE $\rightarrow$ ROLE PERFORMANCE
It should also follow, in terms of logic that

\[
\text{IF } 1. \text{COMMUNICATION} \rightarrow \text{SCOPE} \\
\text{AND } 2. \text{SCOPE} \rightarrow \text{PERVASIVENESS} \\
\text{THEN } 3. \text{COMMUNICATION} \rightarrow \text{PERVASIVENESS}
\]

But the path from communication to pervasiveness fails to receive empirical support, at least with the measures and sample being used here as an illustration. What could have caused this breakdown? If the logic had held, the relationships would be depicted as in Figure 8.6a. However, Figure 8.6b represents an equally good possibility for the

![Figure 8.6a](image)

![Figure 8.6b](image)

first two relationships to hold while the last (the deduced one) would fail. Obviously this approach has certain drawbacks as well as benefits. A listing of both is included on the next page in Table 8.2.
### Table 8.2: Advantages and disadvantages of the axiomatic approach

<table>
<thead>
<tr>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Since some statements are derivable from others, it is not necessary that all the concepts be measurable.</td>
<td>1. Selecting the best set of propositions is often problematical (i.e., how do you know which ones to select)?</td>
</tr>
<tr>
<td>2. Reduces the complexity of the scientific statement of the theory since you only need the set of axioms and its logical system. The rest can be deduced.</td>
<td>2. Invalid conclusions can sometimes follow, even though the premises (statements from which the conclusion has been derived) are true.</td>
</tr>
<tr>
<td>3. Research may be more efficient since support for any one statement lends support for the theory as a whole and for each of its other constituent parts.</td>
<td>3. Valid arguments (i.e., reasoning follows correct rules of logic) but yield faulty results</td>
</tr>
</tbody>
</table>
A couple of nonsociological examples, taken from Ruby (1950) would perhaps illustrate disadvantages 2 and 3 above more succinctly.

+--------------------------------+----------------------------------+
| Disadvantage #2:               | Disadvantage #3:                  |
| (invalid argument based on true premises) | (valid argument resulting in false conclusion |
+--------------------------------+----------------------------------+
| All Muscovites are human.     | Holy Rollers are chain-smokers |
| All Russians are human        | All Muslims are Holy Rollers   |
| Therefore all Muscovites are Russians | Therefore all Muslims are chain-smokers |
+--------------------------------+----------------------------------+

(Ruby, 1950:153)

It should be evident from the discussion above that the axiomatic approach requires both validity (i.e., follows the rules of logic) and verification (i.e. testing to ascertain whether the statement(s) is (are) true) if scientific sociological theory is to be the result.

This approach has not been without its critics. Movahedi and Ogles (1973) quote Einstein (1959) Reichenbach (1953) and Popper (1963) as advocating against the axiomatic approach ever accomplishing much in sociology. Much of their criticism resides in the problems that occur in trying to adapt the rules of logical deduction to ordinary language.
Movahidi and Ogles propose that quantification and the use of mathematical or formalized systems would perhaps be more beneficial than Zetterberg's axiomatic deductive approach with ordinary language.

Reynolds suggests that the best evidence against the use of the axiomatic approach is the scarcity of examples which have used it.

"The lack of social science theories in axiomatic form suggests that it has either been impossible or inconvenient for social scientists to put their ideas with this format" (Reynolds 1971:97).

He suggests further that the causal process form has proven a much more convenient approach to the development of sociological theory. Berger et al. (1962), Blalock (1969), and Abell (1971) advocate the causal approach as well. Blalock even suggests that perhaps axiomatic theories should be converted into causal theories as a means of bringing them more in line with reality and the ability to test that reality.

The causal model approach

According to Reynolds (1971) the causal process approach is characterized by several criteria the first two of which are similar to the first two criteria of the axiomatic approach above. The third is dissimilar, however, since it does not rely on logical deduction, rather, he says that the causal form consists of
1. Definitions

2. Existence statements

3. A set of causal statements, with either deterministic or probabilistic relations that describe one or more causal processes or causal mechanisms that identify the effect of one or more independent variables on one or more dependent variables.

This is not unlike the discussion on causal model linkages raised earlier in the chapter, especially when all three criteria are considered. The terminology is not always the same but the meaning of causality constructed by Reynolds' three criteria are very close to those described by both Rhoads (1971) and Abell (1971). I shall not attempt to repeat their discussions here; rather, a final comment will be made regarding causality and causal models, particularly with regard to the way causality is viewed by those who use the method.

**Causality and causal models.** The notion of cause has been previously discussed in this chapter. However, the important topic of causal models has not yet been placed within the nexus of that discussion. In doing so, an important question which must be asked is whether it is really necessary to understand what causality is in order to apply causal modeling techniques? The criticisms that have surrounded the term, cause, have led many sociologists to
divorce the method from the theoretical meaning.

Cause, according to Francis (1961) is a necessary aspect of science, but one which is not thoroughly understood. Consequently he suggests that sociologists simply dismiss the metaphysical aspects, accept it as an undefined term and proceed to utilize it. This is in fact what many researchers do when they "draw up" a causal model. As a result, causal models are sometimes superimposed on social relations that are really noncausal.

Simon (1957) and Blalock (1962) have suggested that cause should be considered to be an attribute of the conceptual model, which again serves as a sort of justification for not directly addressing the question of cause.

"First, the concepts to be defined all refer to a model --- a system of equations - and not to the "real" world the model purports to describe. Hence both Hume's critique and the determinism - indeterminism controversy are irrelevant to the question of whether these concepts are admissible in scientific discourse" (Simon, 1957:12).

For Blalock, also, cause can be discovered within the model, by examining sets of real and predicted correlations. Thus although this view represents a more sophisticated approach than that suggested by Francis, both Blalock, and Simon are nevertheless still attempting to use cause without addressing the notion theoretically.
A point which has and will continue to be emphasized throughout this dissertation, is the necessity for the adoption of an integrated approach to the study of social phenomena. If this is a worthwhile goal, and if it is to be achieved, can assumptions like causality merely be accepted without examination as advocated by Francis? Or must we conclude as Blalock did that both Hume and Kant are right, i.e., we need the notion of cause in order to develop as a science (Kant) but cause does not exist in the real world, only in the realm of the mind (Hume)? If we accept either of these positions but want to use the notion of cause, we must do so artificially and without much theoretical scrutiny.

An alternative approach has been offered by several writers such as Rhoads (1971) and Abell (1971). According to this approach, Hume is correct that deterministic causation does not exist in the real world. Human relations are too complex for the occurrence of uncontaminated simple cause and effect relations to occur. However, recognition of this complexity led the proponents of this approach to a reformulation of the notion of cause rather than its dismissal or mere assumption. According to that reformulation, cause is viewed as conditional or contingent rather than deterministic. Furthermore, causality is said to have ontological status, that is, it is seen as an attribute of reality rather than merely an attribute of a model.
representing reality. As such it represents an important source of social order, although not the only source. Rhoads (1971:34-35) discusses a number of noncausal relationships that can also operate, such as taxonomic generalizations, functional explanations and etc. Reality then is characterized by both causal and noncausal linkages which means that social researchers should retain the concept of causality and continue to search for causes while at the same time they must also recognize the existence of noncausal linkages when these are encountered. In order to distinguish between causal and noncausal situations, however, the researcher must have a clear understanding of what is meant by "cause".

"Without a clear idea of what causality means it is impossible to distinguish causal relationships from other connectives, the consequence being a lack of clarity about the character of knowledge acquired through empirical research...if the sociologist is merely content to report relationships, the issue of cause and effect does not arise. If, however, he is curious to inquire more precisely into their character, he must face the problem of causation" (Rhoads, 1971:35).

Causal model analysts have been criticized for their "willy nilly" application of the techniques to social situations without due consideration or justification (Short, 1974; Miller and Stokes, 1975). Knowledge of how to use a tool is only beneficial to the extent that one recognizes "when" to use it. We have examined three perspectives on the relationship between causality and causal models. The first
(Francis) emphasized the tool aspect, (don't worry about what it means, just use it). While the second approach (Simon and Blalock) represents a more sophisticated argument for the tool (cause is an attribute of the model), it still does not address the question of appropriateness. Only in the case of the third view is a link made between the researchers model and reality, between the tool and when to use it. According to this view, some, not all, social relations are causal in the real world. These then represent the occasions (the "when") to apply causal modeling techniques; for it is then, and only then, that results truly reflect the reality under investigation. This reflects the view supported by Rhoads and Abell, as discussed above. It would be difficult to ascertain from journal articles which view has the most adherents since the subject is rarely discussed. In that sense, the third view would probably lose by default. On the other hand, perhaps the real significance arises from the fact that causality is not usually addressed when causal models are analyzed. The very lack of consideration tends to substantiate the claim that a gap exists between theory and methods, even sometimes with a method that is supposed to depend on theory.
Implications

The form in which propositions are presented offers yet another opportunity in which theoretical formulation, measurement, analysis and integration can be explicitly brought together (or as with current procedures, kept separate). The form of presentation suggested here practically requires the former (i.e. integration), since the general hypothesis determines both the theoretical statistical hypothesis and the empirical hypothesis which in turn both determine the empirical statistical hypotheses that is actually tested.

Systematization or ordering of propositions within a theory is necessary to provide a clear understanding of the interrelationships about which the theory is concerned. Organizing on the basis of independent or dependent variables is beneficial but generally only when the number of such variables is small. For more complex theories, axiomatization or causal modeling are better. Because of the problems involved with the logical system of the axiomatic approach, the number of sociologists who have utilized it remains small. The causal approach on the other hand has become the "in" approach.

The recent growth in popularity of the causal approach can be contributed to some extent to improved techniques of measurement and causal analysis techniques. However, it
would also seem that this approach has found "fertile ground" because it seems to more closely approximate everyday experience. Reynolds seems to be recognizing this in the statement below:

"No matter what type of theory a scientist claims to be dealing with, when he explains 'how' something happens he usually refers to a description of one or more causal processes" (Reynolds 1971:98).

The use of a path (or causal) model has generally taken the place of an explicit set of hypotheses. This is probably due in part to the fact that it takes less space in an article to present a causal model than it would to list all the necessary hypothesis, especially since the model takes the place of general empirical, and statistical hypotheses. Furthermore it is extremely difficult to state verbal hypotheses that would demonstrate the multiple causality as clearly or concisely as can be done in a causal model. For example, the following model would require eight two-variable propositions or a set of two three-variable and one five-variable propositions to describe the relationships depicted. Even then, the reader would not get the feel for the model as a whole that can be obtained when the relationships are explicitly illustrated in a causal model diagram.
The causal approach will be adopted throughout the remainder of this dissertation since it seems to be an approach that most readily facilitates explanation, one of the primary functions of theory.

Summary

This chapter completes the discussion on theoretical formulation. It has been primarily concerned with linkages in terms of various issues that should be considered in selecting a theoretical linkage and carrying it through the theory construction process. Some of the issues included examining the meaning (nature) of the chosen linkage; whether a qualitative or quantitative linkage was intended; whether the linkage in the real world (and therefore in the proposition) was linear or nonlinear; and specifying the type of linkage (sequential vs. coextensive or causal vs. covariational).

The emphasis was placed on linkages in this chapter because the author feels that they are all too often taken
for granted. Theorists may take great pains to define and operationalize their concepts so that they reflect the real world but rarely do they extend the same effort to the selection of linkages. Instead most theorists follow a rather standard form: X is related to Y; the more X the more Y; and so forth. I am not saying that these are "wrong", only that they don't usually represent the specificity of meaning that the theorist really has in mind. If the relationship in the real world is indeed covariational, as some are, then a covariational linkage, such as X is positively related to Y, should be selected. However, if the relationship in the real world is causal, a covariational linkage is inappropriate and should not be selected.

Propositions and hypotheses were discussed in terms of the various types (general, theoretical statistical, empirical and empirical statistical) and their interrelationships. A format was also suggested for presenting the various types of hypotheses in a way that emphasizes their interdependence.

Finally different modes of systematizing the sets of propositions that make up a theory were discussed. These include ordering on the basis of similar independent or dependent variables, matrices of relationships, axiomatic reduction, and a causal modeling framework. After reviewing the other types of alternatives, the causal modeling
framework was chosen as the one to utilize in illustrations in the remainder of the text.

This concludes the discussion of theoretical formulation activities. Many of the topics raised in this section will have implications for activities in the other phases. It is important that the reader understand that this is an integrated procedure and that the activities discussed here should not be forgotten in the measurement, analysis and inferential phases. Choices made during any one phase, will have implications for whatever choices come later. An attempt has been made to explicitly state some of these in the implication section following each topic. This will be continued in the next chapter on measurement.
CHAPTER NINE:
MEASUREMENT

The previous section of this dissertation was concerned with theoretical formulation activities such as assumptions, units of analysis, concepts and measures, linkages and propositions. An attempt was made throughout that section to demonstrate how these activities are related to activities in the measurement, analysis and inferential phases. The present chapter will continue to emphasize this interdependent nature of activities both within and between phases.

Measurement considerations serve as a bridge between the theoretical formulation phase and statistical testing within the analysis phase. Activities within this phase are not directly involved in the formulation or testing of hypotheses as the theoretical formulation and analysis phases are. (Even operational measures and the operational hypotheses are specified within the theoretical formulation phase. Furthermore both are abstract, though to be sure, on a low level of abstraction).

Measurement issues are all directly involved in providing the mechanism (data) for analyzing the theoretical framework developed in theory formulation. As such measurement concerns are all empirically oriented, and in a way unlike that of either formulation or analysis activities.
This is probably easier to grasp in the former case than it is in the latter or analysis phase since many writers have treated analysis as if it were on the empirical level (due to the fact that data is being manipulated). However as Blumer (1969) and others have so rightly observed, once the data has been extracted from the real world, it too becomes an abstraction, to the extent that it is supposed to represent the attitudes and behavior of the population of units, and to the extent that measurement error is involved. Note: the point was made in Chapter Seven that an operational measure is only an approximation of its real world counterpart, never the thing or characteristic itself. As such it is an abstraction. Clearly, then, measurement represents the linkage between two types of processes, deductive and inductive, within the overall theory construction process. Figure 9.1 on the next page attempts to illustrate this "bridging" function of measurement activities. Furthermore, in so far as it involves both abstract and empirical (as Blumer uses the term) activities measurement also serves the purpose of linking the abstract theory construction process to the real world.
Measurement issues are, consequently, important. The test of a proposition and/or theory can only be as good as the data that is used to test it. However, theorists often find themselves in a position where they must use data that has already been collected. The rising costs of data collection coupled with shrinking sources of funding, will probably pragmatically speaking, tend to make this position even more common in the years ahead. Recognition of this situation (along with the need spoken of in Chapter Three to omit something) led to an emphasis in this chapter on measurement assessment techniques which can be applied after the data has been collected. Even so, a number of activities which come prior to data collection will also be discussed,
since some of these have implications for the analysis and inference phases that follow. This means that even when one uses previously collected data, he/she should be aware of what went into (i.e., the choices that were made in terms of design, sample, collection technique and etc.) the compilation of that data.

Considerations Prior to Data Collection

The collection of data involves a variety of activities which under **ideal** conditions will be influenced by activities and decisions that occurred earlier in the theoretical formulation phase and will, in turn, influence activities within the analysis and inferential phases. The research design serves as the "roadmap" between these earlier and later activities and, to some extent, may even be seen as a sort of operationalization of the problem statement. It will be the first topic covered in this section of the chapter since it serves as the organizing factor for the other types of activities involved in data collection. As pointed out above, it is not always possible to collect data first hand. Therefore, the last topic within this section will attempt to deal with the question of integration when it is not possible to carry out one's own data collection.
Research Design

The research design describes how the theorist hopes to turn the concepts, and measures (indicators) into data that can be used to test the hypotheses which form the proposed theory. As such it outlines such things as the type of study required (longitudinal versus cross-sectional); the type of sample needed (if any), and the type of data collection technique that is desired. The decisions concerning these must of course all be moderated or influenced by decisions made earlier in the theoretical formulation phase as well as by certain pragmatic considerations.

The selection of a theoretical orientation probably serves as the most influential decision of the theoretical formulation phase upon choices made within the research design, especially in terms of whether one needs qualitative versus quantitative data, taken at one time or extending over a period of time. The various theoretical orientations were shown in Chapter Five to have differential preferences in terms of the types of study designs and data collection techniques that were advocated. The orientation of the ongoing example, which is being utilized for purposes of illustration throughout this dissertation, would fall into the category labeled by Ritzer as "Social Factist." An examination of the research design does, in fact, reveal that most of the preferences of the Social Factist orientation
were adopted. Thus the study design was cross-sectional; the sampling technique was random (within three purposively selected states); and the method of data collection were of a survey-type nature. Furthermore, the collected data was quantitative in nature rather than qualitative.

Pragmatic considerations also must influence the decisions made in the research design. These include such things as the size of one's budget, the amount of personnel and their skills, the time required, and etc. (Miller, 1970). The interests and/or goals of the funding agency are also important in terms of the type of data desired by the agency (do they require merely a descriptive study of what is or an analytical study of how or why), as well as specific information, of interest primarily to the agency alone, which reduces the total amount of information that can be extracted for the theorists purposes. The purpose of the example study was preliminary in nature, a prelude to a much larger national study; but was, nevertheless of an analytical nature. Having stated the nature of the example data, I shall briefly discuss these choices and some of the alternatives that could have been selected.
Study design

There are five basic designs that a study can follow, although this list can be extended by modifying one or more of the designs listed below. These consist of the following:

1. Classical Experimental
2. Longitudinal
3. Cross-sectional
4. Single cell
5. Ex-Post Factum

![Figure 9.2a: Classical Experimental Design](image)

![Figure 9.2b: Longitudinal Design](image)

![Figure 9.2c: Cross-sectional Design](image)

![Figure 9.2d: Single cell Design](image)
The controlled experiment  The experimental design is usually held up to be the ideal-type of study design. It allows for an experimental group which is subjected to some form of intervention or manipulation and a control group which is not. Both groups are measured or examined before and after the experimental group has received the intervention. The difference between time one and time two in the experimental group that is not accounted for by the time difference in the control group is said to be the effect of the intervention or independent variable, since both groups were similar in the beginning. This approach tends to be more expensive and time consuming than the others since it requires two groups (longitudinal only needs one) and at least two time periods (cross-sectional only needs one). Furthermore, its use is generally restricted to a laboratory-type situation, in order to obtain the necessary type of control over extraneous influences. Finally the subject matter itself often makes control, in the classical sense, out of the question. Ethical research behavior requires that another type design be substituted when adverse consequences can result from either the intervention (go to a cross-sectional design) or the withholding of the intervention (go to a longitudinal design). A number of alternative, but less rigorous, designs have, therefore, arisen in response to the previously stated problems.
The longitudinal design This design allows for changes over time to be examined, either in terms of the same sample at two points in time (called a panel study) or different samples from the same specific population taken at two points in time (cohort study). It is thought by many to be the "best" compromise of the five designs, in terms of attempting to capitalize on the good points of the experimental design without incurring many of the cost and ethical problems involved in the latter. Nevertheless, it is not without problems either. While it is true that the time element is one of its best features, it is also the source of many of the problems involved in using this approach. Without the control group, it is more difficult to establish whether a change occurred because of the intervention (or independent variable) or for some of the many other possible reasons. For instance, many people just simply change their attitude or opinion about something. The fact that an issue is raised in an intervention or questionnaire may raise a respondent's interest in a subject so that his/her responses at time two may reflect more information rather than a change "caused" by the independent variable. When the respondents are interviewed or re-examined with too short of interval in between, there may be a problem of re-call, (the individual remembers how he/she answered the last time and so gives the same response). If much time is involved, there will
generally be a problem with loss of respondents. People move away, die, or simply change their minds about being involved in a study. Another problem which has in the past turned people away from this design is the fact that analysis procedures were usually quite involved. With the advent of high speed computers, new methods, and an increase in the statistical expertise of many sociologists, this problem has lessened so that there has been an increasing demand within the discipline for greater use of this type of design.

The cross-sectional design

Of the first three designs listed above, this is probably the one which is used most often, even though it is considered to be somewhat less desirable than the other two. Unlike either the experimental or the longitudinal designs, this design involves only one point in time. The purpose of this type of study is to examine relationships between concepts and/or groups from a larger population at one point. Thus the original study on which the example for the text was based, was called a cross-sectional design because it attempted to examine variables with a cross-sectional sample (randomly selected from three purposively selected states). The popularity of this design stems primarily from the ease of its utilization. It avoids the time-related problems of the two procedures above. Statistical techniques such as partial correlation and analysis of covariance often take the place of a control
group. This practice does eliminate the ability for active manipulation but it also reduces many of the ethical problems that active manipulation entails. Furthermore, since most sociological questions involve some notion of change over time for which there is no direct measure, an attempt has been made to correct this situation through the use of recall and temporally (time-ordered) causal models similar to the example introduced at the end of Chapter Eight. In that example, socialization and communication were said to occur before scope (colleague interaction) or pervasiveness (pressure to conform) which in turn occur before role performance. Thus, even though the data was collected at one point in time, time is presented as an implicit factor in the model. Logical interpretation in place of a longitudinal design has, as illustrated above, been the pragmatic concession to this short-coming of the cross-sectional approach. Whether this approach gives the theorist a view of one static moment or a representative picture of most moments is still being debated. This is of course especially true in the case of social change.

The single cell design. This design is considered by all to be the least desirable. It is simply the study of one group at one moment in time and is most generally associated with what some call an accidental sample. As such it is primarily concerned with a description of a unique event. A
study involving the hijacking and subsequent detention of hostages at Entebbe, Uganda would be an example of this type of design. Neither the sample, event nor the variables could be selected in advance. This type of situation eliminates practically all methods of control, including most statistical control procedures, which means that little can be legitimately done with this design beyond a purely descriptive level.

Ex-post-factum design The final alternative has been described as ex-post factum (after the fact) because the researcher or theorist uses data that was previously collected for some purpose other than the present study. As such, no manipulation is possible. The theorist "starts with observations of the dependent and retrospectively studies independent variables for their possible effects on the dependent variable" (Kerlinger, 1964). This has led some to equate it with the single cell design and to write disparagingly of its use, (see Merton's comments in Chapter Four with regard to this approach). These writers have stressed a number of important weaknesses that are inherent in the ex-post facto design, namely lack of control to manipulate or to randomize and a higher risk of misinterpretation. (Note: the latter problem can be abated somewhat by the testing of alternative hypotheses.)
Despite its weaknesses, however, ex-post-facto designs continue to be frequently utilized. Kerlinger attributes its popularity to the fact that the subject matter does not always lend itself to manipulation as required by experimental type designs. While this is undoubtably a factor, high costs of initiating and collecting new data, have led many to the pragmatic conclusion that such designs, when utilized with care, can be beneficial for the process of constructing and verifying sociological theory.

"It can even be said that ex-post facto research is more important than experimental research. This is, of course, not a methodological observation. It means, rather, that the most important social scientific and educational research problems do not lend themselves to experimentation, although many of them do lend themselves to controlled (in terms of testing alternative hypotheses) inquiry of the ex-post facto kind" (Kerlinger, 1964:373).

Evidence of this is further substantiated by the increasing interest that has been shown in the establishment of data banks (the collection in a central location of previously collected data from diverse sources) for the purpose of further utilization in future studies.

The use of ex-post factum (already collected) data does present some problems in terms of integration. The items making up the operational measures or indicators may not be the most advantageous. The theorist must make certain adjustments when this is true in terms of his/her theoretical development and the final inferences that can be drawn. The
preliminary assessment techniques should, however, facilitate interpretation by making comparisons between the conceptual and operational (empirical) definitions and with the reality that both are supposed to represent. Problems of differences in meaning can then be made apparent and taken into consideration when inferences are made. In this way, at least, an attempt can be made to tie the theoretical formulation activities into the analysis phase by way of measurement assessment.

Implications

The design of a study is particularly relevant to the statement of the problem (what you are trying to discover or prove) and the kind of analyses that can be legitimately accomplished. For instance if the theorist wants to examine differences between communities on some variable at one point in time then a cross-sectional design would be most appropriate. But, if the intent is to assess changes in variables over time within a community, the theorist should use a longitudinal design where time can be built into the study in an explicit manner. The statement of the problem, and the hypotheses that reflect it should, in effect, be operationalized by the study design. This is the best way for maintaining the proper continuity between hypotheses and data analysis. If an inappropriate design is chosen, it may not allow the type of analysis technique that is necessary to
test the hypotheses. Clearly then, the selection of a study design should take place within the context of an integrated approach to theory construction.

**Sampling**

The importance of sampling and its relationship to the population and unit of analysis were discussed already in Chapter Six. I shall not discuss sampling techniques here, except to list a few of the more common ones in Table 9.1 on the next page.¹ In order for the reader to get enough of a grasp to apply most sampling techniques, it would be necessary to devote considerably more space to it than the scope of this text will allow.

¹For a more in depth but sociological treatment of sampling, the interested reader is advised to see E. R. Babbie, Survey Research Methods (1973). For a more statistical treatment, see Kish, Survey Sampling (1965).
Table 9.1: Characteristics of some commonly utilized sampling techniques.

<table>
<thead>
<tr>
<th>SAMPLING METHOD</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Sampling</td>
<td>Drawn in such a way that every member of the population has an equal chance of being included. Composition of the sample however is unknown.</td>
</tr>
<tr>
<td>Stratified Random Sampling</td>
<td>Sample is randomly drawn from representative sections with known characteristics so that they are proportionately represented in the sample.</td>
</tr>
<tr>
<td>Cluster Sampling</td>
<td>A multistage sampling technique which involves the initial sampling of groups of elements---clusters---from which the final sample of units are drawn.</td>
</tr>
<tr>
<td>Judgmental or Purposive Sampling</td>
<td>A nonprobability type sample---used when practical considerations preclude the use of probability sampling. Sample is selected on the basis of the theorist's knowledge of a population, its elements and the nature of his/her research needs. Especially beneficial in pretest situation where one wants the widest variety of respondents to test his/her questionnaire.</td>
</tr>
</tbody>
</table>

It should perhaps be emphasized, however that sampling too should be done in the context of an integrated approach to theory construction. No one type of sampling is appropriate for all problems or populations. Furthermore, the types of analyses that can be legitimately carried out vary according to the kind of sample one is working with, just as it did with the type of design that is selected.
Data collection techniques

There are three common techniques of data collection; participant observation, survey (questionnaire or interview), and experiment; each with advantages and disadvantages, but the major difference between them consists in the type of information that results from their utilization. Thus the selection of a data collection technique is as dependent upon the problem statement (i.e. what the theorist is trying to discover or prove) as the type of design and sample are.

Participant observation

The following two definitions illustrate the two positions that are generally held with respect to the use of this particular collection technique. The first definition is positive toward the method, the second is not.
FAVORABLE DEFINITION: "A method in which the observer participates in the daily life of the people under study, either openly, in the role of researcher or covertly, in some disguised role, observing things that happen, listening to what is said and questioning people over some length of time" (Becker and Geer, 1975).

UNFAVORABLE DEFINITION: "An unavoidably retrospective process of registering, interpreting, and recording. The process and the kinds of data are influenced by continuing observed-observer transactions. The role of the observer may be passive or active. In either case affective involvement with the observed develops inevitably and may range from sympathetic identification to projective distortion" (Schwartz and Schwartz, 1955).

Adherents of participant observation stress that it is only through an intimate knowledge (gained from day-to-day observation) of a subject that the theorist will be able to understand a phenomenon through the eyes of the participants themselves. Any other type of understanding, such as that gained by survey methods is, according to this approach, an artifact or construction by the theorist rather than a representation of reality.
Critics, as reflected by the second definition above, contend that the participant observer is merely substituting a human instrument of gathering data in place of the questionnaire or interview. The problem with this, they emphasize, is that the data (observations) which is reported is a function of the observer's experiences, awareness and personality. Furthermore, the observer, as part of the context being observed, both modifies, and is influenced by, that context (Schwartz and Schwartz 1955).

Survey technique The survey approach generally consists of a systematic collection of data from a sample of a population by use of a questionnaire and/or interview. Its supporters suggest that it has the following characteristics:

1. it permits the rigorous, step-by-step development and testing of logical explanations

2. it facilitates the search for causes and effects

3. it aims at generalizability (rather than description of a particular phenomenon) which furthers the development of a scientific body of knowledge.

4. it facilitates the development and testing of alternative explanatory models so that the most parsimonious explanations can be revealed.

5. it forces the user to be specific as to what and how things were done, thus opening it up to more criticism, replication and revision (Babbie, 1973).
Critics on the other hand point out that the survey approach makes a number of assumptions which may not always be valid, such as:

1. Descriptions given by respondents are factual—many responses are in fact "colored", only through observation can the "true" facts be revealed (Becker and Geer, 1957).

2. We assume that the respondent will be able to talk about the things we are interested in and that we understand the meaning of everyday words (Becker and Geer, 1957).

3. Scales have different implicit meanings (Brown and Taylor, 1972).

4. Respondents may mean different things although they give the same answers:

"Whereas men might agree that values are to be allocated on the basis of merit, they may diverge as to the meaning of merit" (Brown and Taylor, 1972).

Most theorists and/or researchers tend to be very adamant in their choice between the two data collection procedures above. Perhaps this is true because the participant observation approach is the one which is often advocated by theoretical orientations labeled by Ritzer (1975) as "Social Definitionists" whereas the survey is the preferred approach of the "Social Factists." It is the position of this text, however, that both techniques have advantages and disadvantages and it is up to the theorist to decide which will offer him/her the type of information which is needed to accomplish his/her goals. Participant observation tends to focus on behavioral type data whereas
the survey emphasizes perceptual and attitudinal data (since behavioral data can only be recalled by the respondent).

**Experiment** The experiment differs from both of the two procedures above. It consists of a control group and an experimental group with data collected at two points in time. An intervention or manipulation is introduced with respect to the experimental group between the two data collection times. Attempts are made to control for the effects of all extraneous variables so that differences between the control and the experimental groups at time two can be attributed to the intervention.

In order to achieve the desired results (i.e. control for extraneous variables) experiments are generally conducted under laboratory-like settings. The "natural setting" observed by the participant observer is therefore missing. In addition, the contrived situation itself may yield answers which are not easily interpreted in terms of real world relationships.

The chief advantage of the experiment relates to the notion of control (control over extraneous variables and control over the introduction and variation of the predictor or intervention variable). However, it is often difficult to choose the "right" control variable so that all extraneous variables are excluded. Kish (1959) cites the example of the man who drank too much on four different occasions: scotch
and soda, bourbon and soda, rum and soda, and wine and soda. Having experienced extreme discomfort on each of the four occasions he consequently arrived at the following **logical** conclusion: "I'll never touch soda again!"

Another problem with the experiment is that the emphasis on randomization (the primary method used in experiments for obtaining control) results in a loss of representativeness. Thus "the **statistical** inferences derived from the experimental testing of several treatments are restricted to the population(s) included in the experimental design" (Kish, 1959:333).

As with the two previous procedures, the experiment has certain advantages and disadvantages depending upon the kind of information that is desired. It also has been associated with a particular orientation, the "Social Behaviorists." The final choice then must rest with the theorist and again should be made within the context of an integrated approach to theory construction.

The example which I am using to illustrate the various aspects of this integrated approach to theory construction utilized a survey technique. The comments throughout the remainder of the text will therefore refer to this type of data collection (unless otherwise stated). The purpose of a survey type data collection technique is to obtain some "measure" of the concepts within the proposed theory so that
analytical tests can be made on the hypotheses representing that theory. In this sense then, measurement can be said to be "the process whereby concepts are mapped into a set of values" (Abell, 1971:49). In the next section I shall discuss the types of measures that are usually employed with survey type techniques.

**Types of operational or empirical measures**

There are three types of measures which will be discussed here. These are single item indicators, indexes or scales, and multiple indicators. The single indicator represents the most simplistic and also the most vulnerable of the three approaches, especially since most sociological concepts are multidimensional (include more than one aspect within each of their meaning spheres).

The index and/or scale approach is the most common. This consists of forming a composite by combining items (to form an index or Likert scale) or using one of several other scaling techniques to arrive at a score for the scale which is said then to represent the meaning sphere of the concept. The table on the following page describes some of the scaling techniques that are used to form scales and indices.

There are of course many others which could have been
<table>
<thead>
<tr>
<th>Scale</th>
<th>Type</th>
<th>Description</th>
<th>Difficulties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likert</td>
<td>Scale or Index</td>
<td>One that has a number of items that are summed together. The sum is supposed to correspond to the measure of the concept.</td>
<td>loss of information; does not give any information about individual answers or their spread. Person whose answers are coded as 1, 1, 1, 1, 1, would get the same total score as one who had 0, 0, 0, 0, 5 but substantively there might be considerable difference between them.</td>
</tr>
<tr>
<td>Thurstone</td>
<td>Scale</td>
<td>Attempts to approximate an interval scale and thus overcome the difficulties of the Likert scale. It is a method of specification of items dealing with a positive vs. negative orientation—deals principally with an attitude. Large number of items are selected and submitted to judges who sort items into piles varying from most to least favorable. Items for scale selected from this and administered to respondents who</td>
<td>Requires a lot of work to develop it. It is possible to get a judge bias in selecting the items. Only have one chance (single response) to try to get individual's &quot;true&quot; attitude about something. If you use more than one response and sum them you again have the problem</td>
</tr>
</tbody>
</table>
choose his/her one best response. One may also sum the items (a kind of combination of the two Likert and Thurstone Procedures)

<table>
<thead>
<tr>
<th>Cumulative</th>
<th>Items are arranged in the scale according to increasing levels of intensity. Respondent selects the highest level of intensity. Bogardus Social Distance Scale is an example of this type.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guttman Scalogram</td>
<td>Also a type of cumulative scale in so far as questions are arranged in increasing difficulty but respondent answers all questions and then an attempt is made to see whether the behavior as indicated by their responses will fit</td>
</tr>
</tbody>
</table>

of loss of information as noted above under Likert

respondents are asked to respond in terms of generalities (stereotypes) does not allow for specific situations which may affect one's general mode toward a subject.

high degree of intuition is involved in deciding whether a set of data forms a Guttman scale (Upshaw, 1968). scale may have limited generalizability. It is an ex post factum scale since it is constructed after the data has been collected. another group of respondents may not react the same as the group on which it is based. Summated scales using the same responses have been shown to have similar results with much less effort (Lin, 1976).
included,¹ but these are the most well-known. Of these, the Likert is the most widely utilized. It combines two or more items using a summated score to represent either a multidimensional or a unidimensional concept. In so doing however, there is an accompanying loss of information related to the individual items making up the composite. The Thurstone scale attempts to approximate an interval scale and probably comes the closest to satisfying this assumption with regard to the use of parametric statistics. However, it requires a great deal of "a priori" (before the fact) work to develop it. Furthermore, you only have one chance to get the individuals "true" attitude about something. Many are uncomfortable about depending on a single response even one selected with such precision. They are recommending that the original items in a scale be developed on the basis of the Thurstone technique (thus insuring a closer match to an interval scale) but that the items should be combined as in the Likert format so that one is not forced to depend on a "single best response" to attempt to capture a multidimensional meaning sphere.

The third approach, multiple indicators, recognizes the

¹For a classic text on scale construction, see Edwards, a., Techniques of Attitude Scale Construction (1957). A shorter, but less comprehensive treatment is contained in Chapter Three of Blalock and Blalock, Methodology in Social Research (1968).
need for multiple items, especially for multidimensional concepts, but also attempts to overcome the loss of information inherent in composites by allowing the multiple indicators or items to enter at the same time but remain separate, i.e., one examines the effect of blocks of variables. This method is still in the developmental stage with a number of problems that are still to be worked out. Nevertheless it does seem to hold promise for the future. For further information, see Blalock (1969), Curtis and Jackson (1962), Jacobson and Lalu (1974) and Sullivan (1971 and 1974).

The measurement approach used in the example was generally of the Likert type. A number of different response frameworks, scoring, and weighting techniques were used. For further explanation with regard to the various response and scoring techniques, see Warren, Klonglan, and Sabri, The Certainty Method (1969).

This completes the topics that will be covered in this section of the chapter on measurement. There are of course many more which could have been included (especially those connected with the actual administration of a questionnaire). However, space limitations do not allow coverage of everything. Furthermore, considering the fact that many theorists will be working with data that has already been collected, this omission is not terribly detrimental since
the important issues are being covered which are required for the mental integration of the formulation, measurement, and analysis phases. The first section of this chapter was concerned with measurement decisions and activities prior to data collection, the second will be examining techniques which can assess the quality of the measures once they have been obtained by the data collection process.

Preliminary Assessment Techniques

This second section will examine various procedures which can be utilized after the data has been collected to assess the quality of the measures in terms of their functional unity (the degree to which they hang together), reliability and validity, and how well they meet the assumptions that are made when various analysis procedures are used. Preliminary assessment techniques can be utilized after a pretest to assess whether changes need to be made in the data gathering instrument before data collection, and/or they may be used prior to data analysis to help assess the quality of the measures for inferential purposes. Note that these techniques are concerned with examining one measure at a time whereas the analysis procedures in the chapters that follow will be concerned with examining relationships between two or more concepts.
Clarity of the items

The first thing to be examined is the clarity of the language used in the items. Are there terms in the items that may be misinterpreted by the respondents? (Does "job orientation" mean the same to respondents?) Do items contain multiple terms? The fourth item in the socialization indicator refers to knowledge, understanding, responsibilities and commitments. This could allow for differential responses - knowledge of responsibilities? Knowledge of commitments? Understanding of responsibilities? Understanding of commitments? or Some weighted combination?

Another substantive problem involved in this measure concerns the propensity to elicit socially desirable answers. It is doubtful that many of the respondents would admit a lack of knowledge or understanding of his responsibilities. These are substantive issues which relate to the correspondence between the socialization indicator and "empirical fact." Other criteria which can be used to assess the quality of the indicator include functional unity, reliability, validity, measurement error, and descriptive statistics.
**Functional unity**

This refers to the degree of consistency that exists among the various measures. In other words, if the measures of an indicator really reflect the same concept then they should be consistent or "hang together." This is, in fact, the assumption (i.e., the notion of homogeneous concepts) on which another criterion, reliability, is based. One way of assessing the consistency of a measure is by examining the intercorrelations of the measures \( (r_{ij}) \), the average intercorrelation \( (\bar{r}_{ij}) \), and item-total correlations \( (r_{it}) \). Examining the intercorrelations gives one a feel for the degree of similarity between the items or measures. The item-total correlation can be compared to a minimum item-total correlation \( r_{it_{min}} = \frac{1}{\sqrt{N}} \) which represents the amount of variance that is contributed by chance alone.

A good example of what can be revealed using these techniques involves the concept role performance. Role performance in the Mulford, et al., (1972) study was defined conceptually as actual behavior judged relevant to the pursuance of the job. A scale of seven items was then developed covering seven task areas designated by national officials as comprising the responsibilities of a local civil preparedness coordinator. The following tables show the results of examining the consistency of these measures.
Table 9.3: Intercorrelations \((r_{ij})^a\) of original role performance scale

<table>
<thead>
<tr>
<th>Variable</th>
<th>Var131</th>
<th>Var132</th>
<th>Var133</th>
<th>Var135</th>
<th>Var136</th>
<th>Var137</th>
<th>Var138</th>
</tr>
</thead>
<tbody>
<tr>
<td>Var131</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var132</td>
<td>0.0695</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var133</td>
<td>-0.0636</td>
<td>0.2443</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var135</td>
<td>-0.0606</td>
<td>0.3887</td>
<td>0.4446</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var136</td>
<td>0.1375</td>
<td>0.4958</td>
<td>0.2568</td>
<td>0.4692</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Var137</td>
<td>0.0260</td>
<td>0.7484</td>
<td>0.3265</td>
<td>0.4808</td>
<td>0.4488</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Var138</td>
<td>-0.0716</td>
<td>0.4055</td>
<td>0.3087</td>
<td>0.5296</td>
<td>0.4205</td>
<td>0.4391</td>
<td>1.000</td>
</tr>
</tbody>
</table>

It takes an \(r=0.138\) at .05 level with 240 d.f.

Table 9.4: Corrected item-total correlations for original scale with a minimum \(r_{it} = 0.378\)

| VAR131   | 0.0423 |
| VAR132   | 0.5553 |
| VAR133   | 0.3304 |
| VAR135   | 0.435  |
| VAR136   | 0.8442 |
| VAR137   | 0.5544 |
| VAR138   | 0.4719 |

Table 9.5: Corrected item-total correlations for revised scale with a minimum \(r_{it} = 0.408\)

| VAR132   | 0.6066 |
| VAR133   | 0.4152 |
| VAR135   | 0.5912 |
| VAR136   | 0.5662 |
| VAR137   | 0.6340 |
| VAR138   | 0.5796 |

A quick glance at Table 9.3 shows that one item appears to be significantly different from the other items. The largest intercorrelation of VAR131 (Licensing and Stocking of fallout shelters) with any of the other tasks is only .1375 and most of the others are no more than .07. The discrepancy is also evident when one looks at the item-total correlations.
in Table 9.4. VAR131 has an item-total ($r_{it}$) correlation of only .04 while most of the others are over .4. The minimum $r_{it}$ is equal to .378 which says that VAR131 contributes considerably less than one would expect by pure chance.

While it seems evident that in this sample, VAR131 (or Licensing and Stocking of fall-out shelters) does not seem to be consistent with the other measures, this, in itself, would not justify the dropping of this item from the total indicator. One should re-examine the theoretical and concrete reasons for its original inclusion. Since the conceptual definition refers to "actual behavior" rather than nationally prescribed tasks, there would seem to be some theoretical ground for omitting a task which was not relevant to the local coordinator. Further evidence of the lack of relevance, occurred some years later when the national level officials formally recognized this fact by omitting it in their description of the local coordinators responsibilities.

Having established both empirical and theoretical reasons for dropping VAR131 from the composite, Table 9.5 represents the results of the item-total correlations using a revised six-item scale. The minimum $r_{it}$ for six items was .408. All six items fall above the minimum for the six-item scale. The average intercorrelation ($\bar{r}_{ij}$) for the revised scale (omitting VAR131) was .4272 as compared to only .3255 for the seven-item scale. Thus it was felt that the
six-item, revised scale was more reflective of the actual behavior than the original seven-item scale. What this example hopefully demonstrates is that one need not be "locked into" using all the original items for an indicator, if theoretical and empirical reasons can be found to justify their being discarded. However, this does not mean that the researcher should drop items just to improve the homogeneity of his indicator. That would be allowing the data, rather than the theory, to determine his measure. And if the measure or indicator is not grounded in theory, little can be gained by its manipulation.

Warren, et al., (1969) suggest examining the range of 60% or more of the intercorrelations as a means of assessing the homogeneity of the intercorrelations. The smallest range of 60% of the intercorrelations using the seven-item scale was between 0.4958 and .2443 which amounted to a difference of 0.2515. This compares to the smallest range on the six-item scale between 0.4958 and 0.3887 with a difference of 0.1071. It seems clear from the above comparison that dropping VAR131 has increased the homogeneity or consistency of the role performance scale as well as making it more relevant to the actual (as opposed to prescribed) behavior of local coordinators. Since it is the actual (new measure) and not the prescribed (original measure) which is being examined conceptually, it was felt that this change represented a
legitimate improvement in the measure, and not merely manipulation for its own sake.

Reliability

Reliability is concerned with the accuracy or precision of a measuring instrument (indicator). In other words, if we or someone else measure the same set of objects (or a comparable set) over and over with the same (or comparable) instrument, would we get the same results? It was mentioned earlier that the criterion of reliability was originally based on the notion of homogeneous concepts. Most of the early work on estimating reliability comes from psychology where they were usually dealing with single traits having fairly stable characteristics. Since many sociological concepts are multidimensional rather than unidimensional, the adoption of this criterion by sociology has presented some problems. One approach has been to assume that the concept is unidimensional and concentrate on measuring only one aspect with a single item [this is what Brown and Taylor (1972) call mapping data into a simple order and Coombs (1953), the insuring of unidimensionality by definition]. There are of course, other approaches one can use with multiple item scales or indexes some of which will be described briefly here.

A review of the reliability literature (Bohrnstedt, 1970) reveals that reliability can be approached from three
different perspectives, stability (if we measure the same set of objects again and again with the same or comparable measuring instruments will we get the same results?), equivalence (internal consistency or homogeneity) and accuracy (the relative absence of measurement error or the proportion of true variance to total variance).

**The stability approach to reliability** A test-retest situation is used in this approach such that responses are obtained from the same individuals using the same instrument across at least two points in time. Some of the assumptions which should be met with this procedure are as follows:

\[
\begin{align*}
\rho_{X_i T_1 T_2} &= 1 \quad \text{(The correlation of an item at time 1 with the same item at time 2 should equal one.)} \\
\bar{X}_{i T_1} &= \bar{X}_{i T_2} \quad \text{(The mean of item } X \text{ at time 1 should equal the mean of the same item at time 2.)} \\
\sigma_{i T_1} &= \sigma_{i T_2} \quad \text{(The standard deviation of item } X \text{ at time 1 should equal the standard deviation of the same item at time 2.)} \\
\rho_{X_i T_1 T_2} &= \rho_{X_j T_1 T_2} \quad \text{(The intercorrelations of items from the two times should be equal.)}
\end{align*}
\]

The major problem with this type of reliability is its longitudinal nature, i.e., the effect of time. The length of time between administration of an instrument influences the reliability score. The longer the time, the lower the
coefficient. However, if an insufficient time has passed, actual recall of the questions and the individual's previous response will contaminate the "true" reliability coefficient. A third problem associated with this approach is called the reactivity problem. This refers to the fact that the respondent's sensitivity or interest in a particular subject matter may have been enhanced by the act of responding to the instrument during time one. Thus change may occur at time two because the individual's knowledge about the subject has increased as a result of heightened interest at the first administration (Campbell and Stanley, 1966). This brings up the fourth problem which simply states that the individual's true scores have a greater probability of actually changing as the time interval gets larger. This may result from environmental events (such as the energy crisis) or simply the maturation process which "molds and remolds behavioral patterns and thought processes" (Lin, 1976:177).

Some researchers have turned to preparing parallel forms of tests as a means of eliminating the time factor while avoiding recall problems. However, truly "parallel" forms are extremely difficult to create. Heise (1969) has approached the problem from another avenue. He has proposed a formula for distinguishing real change from unreliability. His method requires observations from three points in time, that time intervals be equal, and that errors in measurement
are uncorrelated.

### Reliability

\[
\rho_{xx'} = \frac{\rho_{12} \rho_{23}}{\rho_{13}}
\]

Where \( \rho_{xx'} \) is the overall reliability of the indicator and \( \rho_{12} \) is the reliability using time 1 and time 2.

### Estimate of the amount of real change over time

\[
S_{12} = \frac{\rho_{13}}{\rho_{23}}
\]
\[
S_{13} = \frac{\rho_{13}}{(\rho_{12} \rho_{23})}
\]
\[
S_{23} = \frac{\rho_{13}}{\rho_{12}}
\]

Where \( S_{12} \) is the change that has taken place between time 1 and time 2.

Because of the difficulties involved with the stability approach (not the least of which is the requirement of a longitudinal design), many sociologists have turned to the equivalence or internal consistency approach. One of the first such methods was the split-half approach. In this case, the measures or items of the indicator were randomly assigned to one of two forms or halves. The reliability coefficient is again the correlation between the two halves. However, unlike the parallel forms in the stability approach, the two halves are summed to form a single composite in any further analysis. The major problem with the split-half approach is that each split one makes could yield different reliability estimates, since the randomly assigned split halves could be far from equivalent. For this reason this approach has generally fallen into disuse.
The most popular type of equivalence approach has been the Kuder and Richardson (1937) measure of internal consistency. It examines the covariance among all of the items simultaneously rather than in an arbitrary split. It assumes that all items are parallel, that the difficulty of all items are the same, and that the condition of tau equivalence exists. [Tau equivalence says \( T_a = T_b + c \) where "a" and "b" are items and "c" is a constant. Also tau equivalence assumes that the true score variances are equal but makes no such assumption about the means or the error variances.] Further assumptions under this procedure include that the covariances and correlations should be consistent. To the extent that these assumptions are not met, particularly those associated with tau equivalence, the formulas for this type of reliability will usually tend to underestimate the reliability coefficient.

An example of this type of approach was developed by Cronbach (1951) and is commonly referred to as coefficient alpha (\( \alpha \)).

\[
\alpha = \left( \frac{n}{n-1} \right) \left( 1 - \frac{\sum \sigma_i^2}{\sigma_X^2} \right)
\]
where \( \sum \sigma_i^2 \) = sum of the individual variances
\[ \sigma_X^2 = \text{variance of the total scale} \]
\[ = \sum \sigma_i^2 + 2 \sum \sigma_{ij} \]
\[ = \Sigma \text{diagonal} + 2 \Sigma \text{of the off-diagonals} \]

These are only a few of the many possible alternatives available. In his classical article on reliability measures, Guttman listed at least six others. The variety should neither overwhelm the researcher (which one to choose?), nor make him/her complacent (Any one will do!). Results from the various methods can vary by as much as 10 points or more. It is important to realize then that true reliability can only be estimated at best. Failure to meet the assumptions corresponding to a particular formula can reduce the efficiency of the resulting estimate. Therefore the formula which can be considered as most appropriate is the one to which the researcher can come closest to meeting the assumptions. Other criteria include the state of the theory---if you have a good model, split-half may be good enough---and the intended use of the indicator.

Before leaving the notion of reliability we should perhaps examine some of the things which can influence the size of the coefficient. The classical answer has included three things: (1) the number of items, (2) the content of items, and (3) the interrelationship of items. Generally
speaking, increasing the number of items usually tends to increase the size of the reliability coefficient. However, this is not always true, and a good example is the measure for role performance discussed earlier. The seven item scale had a lower reliability coefficient (0.430) than the six item scale (0.708), primarily because the content of variable 131 set it apart from the rest so that the interrelationship of this variable with the other six in the scale was negligible.

There are also other factors which will have an effect on the size of the coefficient. The presence of a moderator variable is one. In a later chapter we will examine the effect of state on the analysis of relations between variables. The question will be raised as to whether the measures and results would have been more reliable if each state had been analyzed separately. Considerable differences in the reliability coefficient can also be obtained by using different scoring techniques (three point - agree, don't know, disagree - versus the certainty method) as well as different calculation formulas.

Finally there are a number of factors which introduce variation (or error) that may inhibit replication of the research results, and thereby affect reliability. These include situational factors (the room may be uncomfortably hot or noisy on the day of the administration); transient factors (respondent might have been tired or ill); there may
have been variations in the way the instrument was administered (one group might have been given more complete instructions or more time); and even analysis factors (coding, tabulation, rounding, and computer formatting errors) can all enter into the determination of the reliability coefficient.

It should be obvious from what has just been said that there is no way to determine the exact reliability of an indicator. We can only obtain estimates which will be adequate to the degree that (1) adequate samples are used and (2) underlying assumptions are met (Bohrnstedt, 1969).

An issue which is closely related to reliability is validity. Both are concerned with assessing the quality of an indicator, however, they approach that assessment differently. It is not sufficient to examine only reliability. An indicator can yield the same results time and time again and yet still not be valid. Validation must be established independent of reliability procedures.

Validity

Validity concerns the degree to which indicators measure what they purport to measure. Unfortunately, most procedures for establishing validity are not as straight-forward as the reliability formulas. They tend to rely more on intuitive evaluations.
In a recent book on social research, Nan Lin (1976) has proposed a framework for discussing the various approaches for confirming the validity of measures. Although the framework itself seems beneficial as a means of clarifying the distinctions between the various types of validity, the fact that Lin chose not to use conventional terminology, tends to detract from its utility. Therefore, Figure 9.3 below incorporates both Lin’s framework and Bohrnstedt’s (1969) terminology (in parentheses).

---

Figure 9.3: Approaches for establishing the validity of an indicator

Validity then can be established by either empirical or conceptual confirmation. In empirical confirmation, information from one's measure is compared to facts and outcomes found in reality. This type of validity is commonly referred to as criterion related validity. According to
Bohrnstedt there are two types of criterion related validity, (concurrent and predictive). In concurrent validity, a measure or indicator is compared to an outside criterion or direct measure of the property under investigation. For instance, an attitude scale on religious orthodoxy could be correlated with church attendance records. Predictive validity is similar except that the outside criterion is one which may occur in the future, (e.g., measure of initial marital adjustment might be correlated with divorce rates of say those who had been married for 10 years.

For most measures empirical confirmation is not possible. In those cases the researcher must turn to the more indirect conceptual confirmation approaches. In conceptual confirmation (or construct validity) validity must be inferred from conceptual evidence. Lin distinguishes between internal and external validity. Internal validity is concerned with the study (or data) situation, external validation involves the generalizability of relationships to the theoretical structure (as in the case of content validity) or to larger populations (as with cross validation).

Internal validity involves both inference across items (convergent validation) and inference across variables (discriminant validation). In convergent validity, multiple items measuring the same concept are correlated. The higher
the correlation among these items, particularly if different data gathering methods (survey, experiment, etc.) are used, the greater the convergent validity. Whereas convergent validity examines items within a single concept, discriminant validity examines items between theoretically related but distinct concepts. Suppose X and Y are two concepts that are thought to be related. The correlations in the triangles represent correlations between items of a single concept. The correlations within the rectangle represent correlation of items across concepts, i.e., an X with a Y. In discriminant validity then, the correlations in the triangles should be greater than the correlations in the rectangle. If the difference is small, it means that the items may not be tapping different dimensions. Thus either the concepts are not distinct or the measures are not valid.
Turning now to the notion of external validity, attention centers on the generalizability of the measures and findings rather than on the study (or data) situation. Thus Lin states that external validity depends on (1) the operationalization process and (2) the sampling process. In the first case, Lin seems to be referring to what has generally been described as content validity. Content validity concerns the degree to which an indicator represents the concept about which generalizations will be made.

"One's measuring instrument shows content validity to the degree that sampling from the domain of content is representative of all strata and to the degree that items tap subleties of meaning within all strata" (Bohrnstedt, 1973).

This is precisely what was suggested earlier on page 184, with respect to the comparison between the substantive meaning spheres of concepts and indicators.

The second type of external validation requires a number of studies and has been labeled cross-validation.

"The greater the variety of populations involved and the greater the variety of theoretical structures brought to bear, the more likely it is to ascertain the external validity of a measure" (Lin, 1976:175).

In other words, the more useful (in terms of being relevant for different populations and/or studies) an indicator proves to be, the more external validity will be associated with it. The North-Hatt Occupational Prestige Scale is an example of an indicator which has been applied to a great many different
populations and studies and as a consequence has gained a certain degree of external validity.

The purpose of most social research is to test the relationships between variables. Reliability influences the results obtained in data analysis and thus the inferences that reflect those results. Validity, on the other hand, influences theory testing (since it is concerned with whether the desired content was measured) and inferences. Such efforts will therefore be futile, if the indicators or measures of those variables are not valid and reliable. It was mentioned earlier that measurement error affects the reliability and validity of an indicator. A number of adjustment procedures do exist for measurement error but due to their complexity will not be discussed here.\(^1\) However, there is also a rather simple and quick method for assessing whether the variance of an indicator is due entirely to measurement error. Since this relates to the quality of an indicator, this technique will be discussed next.

---

\(^1\)Several measurement error adjustment techniques are discussed and illustrated in an unpublished paper by Warren et al. (1975).
Measurement error

According to measurement theory, the observed value of an item is equal to the true value plus measurement error.

\[ X = x + e \]
\[ \text{Obs'd} = \text{true} + \text{error} \]

The variance of the observed value becomes then:

\[ \sigma_X^2 = \sigma_x^2 + \sigma_e^2 + 2\sigma_x\sigma_e \]

the sum of the variance of the true value, the variance of the measurement error and the covariance of the true value and error value. However, if one makes the assumption that the measurement error is independent of the true score (a common assumption for most measurement and analysis techniques) the covariance term drops out leaving the simple sum

\[ \sigma_X^2 = \sigma_x^2 + \sigma_e^2 \]

\[ \text{Obs'd} = \text{true} + \text{error} \]
which can be rewritten as follows:

\[ \sigma_e^2 = \sigma_X^2 - \sigma_X^2 \]  
(4)

Reliability (\( \rho \)) can be defined as the true variance divided by the observed variance.

\[ \rho = \frac{\sigma_X^2}{\sigma_X^2} \]  
(5)

This too can be rewritten so that the true variance is a product of the reliability coefficient and the observed variance.

\[ \sigma_X^2 = [\rho (\sigma_X^2)] \]  
(5)

Substituting equation 6 for the true variance in equation 4, we get the following

\[ \sigma_e^2 = \sigma_X^2 - \sigma_X^2 \]  
(4)

\[ \sigma_e^2 = \sigma_X^2 - [\rho (\sigma_X^2)] \]  
(7)

Error = Obs'd - True

var var var

Having calculated the variance of the measurement error
in an indicator, it is now possible to calculate an F-test for the presence of true variance. The Ho hypothesis states that the variance of the true score will be equal to zero.

\[ \text{Ho: } S_{X_i}^2 = 0 \]
\[ \text{Ha: } S_{X_i}^2 \neq 0 \]

The test is then calculated as follows:

\[ F = \frac{\text{observed variance}}{\text{error variance}} = \frac{S_{X_i}^2}{S_e^2} \]

Table 9.6 below contains the calculations for this test using socialization as an example.

Since the F-value was significant the null hypothesis can be rejected and we can conclude that at least some of the variation in the socialization indicator is due to true variation. If the F-value had not been significant, we would have had to conclude that most of the variation was due to measurement error and the indicator would have to be discarded or at the very least re-examined in terms of adding or deleting items.

Examining the variance for measurement error is just one of the ways that descriptive statistics can aid in assessing
Table 9.6: Test for amount of true variance

<table>
<thead>
<tr>
<th>Variable: Socialization</th>
<th>Reliability: 0.619</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variance: 26.5607</td>
</tr>
</tbody>
</table>

\[ \text{Error} = S^2 = S_{X}^2 - \left[ \rho \left( S_{X}^2 \right) \right] \]
\[ S_{e}^2 = 26.5607 - \left[ (0.619)(26.5607) \right] \]
\[ S_{e}^2 = 10.1196 \]
\[ F = \frac{S_{X_i}^2}{S_{e}^2} = \frac{26.5607}{10.1196} = 2.62^* \]

*It takes an \( F \) of 1.26 at the .05 level of significance.

the quality of an indicator. Generally speaking, little attention is usually paid to descriptive statistics although in the next section we will discuss some of the insights that can be gained from such an examination. Table 9.7 represents an attempt to place the discussion of descriptive statistics in perspective in terms of the single variable statistics that will be discussed here (Roman numeral I) and those dealing with relationships between two concepts that will be discussed in the next chapter.
Table 9.7: Summary of types of statistics to be examined

I. Single Variable Statistics
A. Descriptive
   1. Central tendency
      - mean
      - median
      - mode
   2. Dispersion
      - Variance
      - Range (Min, Max)
      - Average Deviation
   3. Overall Distribution

B. Inferential
   1. Standard errors of estimates
   2. Sampling distributions of our statistics
      (Important for construction of confidence intervals)

II. Bivariate or Two Variable
A. Descriptive
   1. Correlation
      - Not concerned with making inferences
   2. Regression
      - beyond the sample.

B. Inferential
   1. Standard Errors
   2. Sampling distributions of our statistics
      (e.g. correlation and slope)
   3. Continuous Variable Statistics
      a. Correlation
      b. Regression
      c. Analysis of variances (Comparison of groups)
         (1) Causal - Apply different level of treatment
            (independent variable) to get different degrees of the dependent variable.
         (2) Profile - Groups of the dependent variable are examined for degree of independent variable.
Descriptive statistics

Descriptive statistics contain a wealth of information about the quality of indicators, particularly with respect to their reliabilities. In the discussion that follows, a brief explanation of some of these insights will be given. This will be followed by a table containing descriptive statistics for the total sample (n=240) and for each of the three subsamples (n=36). For the purpose of demonstrating certain points about descriptive statistics, an assumption will occasionally be made that the total sample represents the population, to which the subsamples can be compared. Finally, many of these implications will be demonstrated by examining the descriptive statistics for the socialization indicator.

Means: The sum of the scores of a variable divided by the total number of scores.

a. The mean is a useful tool in many analysis procedures, however, it is highly affected by extreme values.

b. The mean may be a more meaningful measure sociologically than the total. Suppose you have a 10 item scale with a total score of 90. What does that mean? If you report a mean value of 9, it can be compared across studies unlike the total sum which is contingent on the number of items in the scale or indicator.

c. Serves as a measure of location which can be compared to a known criterion. Useful in (1) retest situations and (2) when one wants to examine groups with a specified range or value on a given criterion variable, for instance, those with high job satisfaction.
Relative Standard: are the means coming out the way they should, i.e., if we expected a high socialization mean, did we get it?

Fixed Standard: are our people reporting to us the same as they were reporting to the census data?

d. Can reveal halo or response biases, particularly on items of satisfaction or perceptions. Some people tend to use limited ranges such as 1's and 2's or 4's and 5's (response bias). In cases where bias is revealed, means can be adjusted by correcting a respondents scores by his standard deviation.

The mean should not be examined in isolation of other descriptive measures. It is useful to compare it to both the median and the mode. The halo and response biases discussed above can be revealed by inspecting the mean with respect to the mode and the median. Furthermore, the assumption of normality can be examined using the mean, median and mode. If the distribution is normal, they will all be the same.

Median: The value that neither exceeds nor is exceeded by more than half the observations in an ordered set - the middle value.

As such it represents a measure which gets around the problem of extreme values. It also serves as a criterion for whether the mean has been unduly distorted by extreme values.
**Mode:** The most frequently occurring value. In addition to indicating the distortion of the mean by outliers, the mode can have substantive importance also. A bimodal distribution may indicate that more than one meaning sphere is being addressed.

**Variance:** Attempts to describe the extent to which scores for the individuals on a given variable are spread out or dispersed.

This is important to understand and examine because if respondents answer all questions the same or if all people acted the same way then there would be no need for sociologists because there would be no differences to explain.

The mean is taken as a criterion and each individual's score is related to the mean to determine how much different it is from the average (or criterion).

\[
\frac{\sum (x_i - \bar{x})^2}{N-1} = S^2
\]

The variance then is the average of the squared deviations from the mean. Serves as an important element in figuring coefficients of reliability.

**Standard Deviation:** Is merely the square root of the variance. Brings statistic back to that of the original unit of measurement.

Suppose the variable budget is measured in dollars:

- Mean is in dollars
- Variance is in \((\text{dollars})^2\)
- Standard deviation in dollars

If variable is socialization

- Mean = 23 socialization units
Variance = 26.561 (socialization units)²
Standard deviation = 5.154 socialization units

The greater the spread about the mean the larger will be the standard deviation. Thus if we assume a normal distribution

Two-thirds of the cases should fall within ± one standard deviation.

95% of the cases should fall within ± two standard deviations.

Range: Measure of dispersion. Represents the difference between the maximum score and the minimum scores.

In any one study there is a possible range and an observed range which may or may not coincide. Perhaps if you are getting little variance you may find, as in the case of subsample (107) on the tension variable that the observed range is limited - in this case it was half that of the possible range and of that exhibited in the other subsamples.

Perhaps the items terms are not broad enough. One must question the validity of an indicator (how well one is sampling the entire substantive domain) when the observed range is less than the possible range.
Sampling distributions of statistics

Sample 1 \( \bar{X} = 22.671 \)
Sample 2 \( \bar{X} = 23.452 \)
Sample 3 \( \bar{X} = 23.731 \)

POPULATION mean for Socialization \( \mu = 23 \)

By drawing samples of size \( n \) repeatedly from a population, a statistic, such as the mean, can be computed for each sample. The distribution of this statistic can then be examined to see how nearly it approaches a normal distribution.

Standard Error: (e.g., standard error of the mean)
How much variation would there be in \( \bar{X} \) if we took a large number of random samples?

\[
S_{\bar{X}} = \frac{S_X}{\sqrt{N}}
\]

This means as \( N \) increases the error or difference between the true population mean and the sample mean will decrease.

This gives us an estimate of how accurately we have measured our mean. Some statistics books see this as examining the reliability of the mean.

Used in setting up confidence interval. We want a confidence interval (C.I.) with a narrow range. Variability and sample size affect the size of C.I. The larger the range of variability on a scale or indicator, the larger the sample size which is necessary to determine the accuracy of that scale.

Confidence Interval can be calculated for any statistic. For instance, if we translate to Z scores, correlations can be put into a C.I. Some people say we should calculate C.I. for our correlations, then make up 2 correlation matrices, one representing the low end of the C.I. for the correlations and the other matrix for the high end. Regressions run on the original correlation matrix should then be compared to results of regressions from the C.I. matrices.
Skewness: A distribution is said to be skewed when one tail is longer than the other.

CASE I

Large positive value for skewness

CASE II

CASE III

Large negative value for skewness

If the distribution is skewed to the right as in Case I (large positive scores) the mean will be greater than the median and the skewness score will be positive.

\[
\frac{3(\bar{x} - \text{Md})}{\frac{S}{\sqrt{N}}}
\]

If the distribution is skewed to the left, the mean will be smaller than the median and the skewness score will be negative. If you wind up with highly skewed distribution need to examine whether outliers really should be considered to be within your designated population.
Kurtosis: Measure of the general peakedness of the distribution of a set of scores.

![Graphs showing Leptokurtosis, Normal Curve, and Platykurtosis](image)

- **Leptokurtosis**: Positive value
- **Normal Curve**: Kurtosis = 0
- **Platykurtosis**: Negative value

CASE I  | CASE II  | CASE III
---|---|---
![Diagrams](image)

\[
\text{MEAN} \quad \frac{S_x}{n} \quad \frac{S_r}{n} \quad \frac{S_x}{n} \quad \frac{S_x}{n} \quad \frac{S_x}{n}
\]

Where \(S_x\) = Standard deviation, Case I

Standard deviation for Case I will be much smaller than that of Case II or Case III. Standard deviation (spread about the mean) for case III will be much larger than for the other two cases. Scores which are clustered in a smaller region about the mean, may mean that the categorical distinctions of the items are too broad (respondents are unable to make specific choices).

In order to demonstrate more clearly what has just been discussed, a table (Table 9.8) has been included on the next page which contains descriptive statistics for the eight
<table>
<thead>
<tr>
<th>Variable Label</th>
<th>Sample</th>
<th>Mean</th>
<th>Variance</th>
<th>St. Dev.</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>Actual Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR037 (Selectivity)</td>
<td>104</td>
<td>12.407</td>
<td>19.238</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>VAR038 (Communication)</td>
<td>104</td>
<td>12.187</td>
<td>19.230</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>VAR058 (Salience)</td>
<td>104</td>
<td>11.587</td>
<td>19.230</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>VAR083 (Pervasiveness)</td>
<td>104</td>
<td>11.587</td>
<td>19.230</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>VAR102 (Scope)</td>
<td>104</td>
<td>11.587</td>
<td>19.230</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>VAR117 (Socialization)</td>
<td>104</td>
<td>11.587</td>
<td>19.230</td>
<td>4.386</td>
<td>-0.619</td>
<td>-2.124</td>
<td>0</td>
<td>16.64</td>
</tr>
<tr>
<td>ROLE7 (New Role Performance)</td>
<td>104</td>
<td>962.404</td>
<td>207126.813</td>
<td>455.112</td>
<td>-0.934</td>
<td>-0.264</td>
<td>0</td>
<td>1838</td>
</tr>
</tbody>
</table>

Table 9.8: Descriptive statistics for subsamples and total sample
variables in the sample data set introduced earlier. [Note: These statistics will be examined in terms of both the total sample and three subsamples.]

Two questions are particularly important with respect to the information contained in that table.

1. If we take the Total Sample as the Universe, how well do the three samples reflect the universe?

2. What information can the various descriptive statistics add to our knowledge on the quality of our measures?

If we examine the descriptive statistics in Table 9.8 for the socialization indicator we get the following results:

Table 9.9: Descriptive statistics for socialization

<table>
<thead>
<tr>
<th>SAMPLE</th>
<th>108</th>
<th>107</th>
<th>104</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>23.667</td>
<td>22.167</td>
<td>23.583</td>
<td>23.000</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>19.943</td>
<td>22.771</td>
<td>35.050</td>
<td>26.561</td>
</tr>
<tr>
<td>ST. DEV.</td>
<td>4.466</td>
<td>4.772</td>
<td>5.920</td>
<td>5.154</td>
</tr>
<tr>
<td>KURTOSIS</td>
<td>0.456</td>
<td>-0.097</td>
<td>-0.031</td>
<td>-0.036</td>
</tr>
<tr>
<td>SKEWNESS</td>
<td>-0.372</td>
<td>0.625</td>
<td>-0.960</td>
<td>-0.038</td>
</tr>
<tr>
<td>MIN.</td>
<td>15</td>
<td>9</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>MAX.</td>
<td>36</td>
<td>33</td>
<td>33</td>
<td>36</td>
</tr>
</tbody>
</table>

If we assume for the moment that the total sample (N=240) represents the entire population and that the population mean (µ) is known to be 23, then the mean for the sample (N=36) (subsample #108) of 23.667 would contain the
population mean, within its critical range. In other words, it represents a "good" or reliable estimate of the population mean - given the above assumption, and the confidence interval of 23.667 ±t(.744).

It was mentioned earlier, that the mean tends to convey more substantive information than the total score, particularly since it can be compared across samples and studies. If we compare the mean across the three subsamples (#108: \( \bar{X} = 23.667 \); #104: \( \bar{X} = 23.583 \); and #107: \( \bar{X} = 22.167 \)) we can see that the mean remains relatively stable across subsamples. The information in Table 9.8 does not contain the median or the mode, however, skewness does give one a feel for the amount of distortion that is present in the mean by comparing it to the median. In this case, the skewness values for socialization in each of the subsamples (except 108) as well as the total, are all close to zero. Even in the case of subsample 108, the value is less than one. This would seem to indicate that the mean of socialization in any of the samples is not being unduly distorted by outliers.

The variance represents an attempt to describe the amount of spread or dispersion that the socialization scores have about the mean. For subsample 108, the variance was 19.943, which is considerably less than that of the total sample variance (26.561). Of course a possible explanation of this difference can be seen by examining the observed
ranges for the two samples. In the case of subsample 108, the observed range goes from 15 to 36, whereas in the total sample the observed range is much larger (from 9 to 36). On the basis of the subsample results alone we might conclude that there was a possible "false bottom" for the socialization scores, that the items were so broad that respondents couldn't make many distinctions in their positions. However, the fact that the total sample covers the entire theoretical range and that the total sample mean is fairly close to the potential population mean, (23 for the total sample and 22.5 for the potential), signifies that the problem in subsample 108 probably lies with the sample, not the socialization indicator.

The distortion of outlyers is usually examined with reference to their effect on the mean. However, the variance is often more distorted by outlyers than the mean. The kurtosis measure offers a method for examining the normality of the curve or distribution of scores. Since the kurtosis value for socialization is not large (.372 in subsample 108 and -0.038 in the total sample) we can conclude, as we did with the mean, that the variance is not being unduly distorted by outlyers.

The interrelatedness of descriptive statistics for evaluating the quality of measures or indicators should be evident. The relation between mean, variance, skewness and
kurtosis can be shown to be the first through the fourth statistical moments. Thus the examination of these four descriptive statistics, combined with the range, median and mode, can shed considerable insight on the quality of one's indicators.

Summary

It was stated in the introduction to this chapter that measurement considerations serve as a bridge between the theoretical formulation activities and the statistical techniques utilized in the analysis phase. This connection was explicitly demonstrated with each topic (design of study, sampling, data collection technique and types of measures) discussed under the subheading, "Considerations Prior to Data Collection."

The linking aspect of measurement activities in the Preliminary Assessment Techniques section of the chapter was handled more implicitly than in the first section. Here the emphasis was upon assessing the quality of the measure. Nevertheless, these techniques also serve a linking function. According to the integrated approach of theory construction that is being described, operational or empirical measures are defined within the theoretical formulation phase. But the act of constructing items and administering them occurs in the measurement phase. Once this has been done it is necessary to see if the theorist has indeed measured the
conceptual meaning he/she had intended (validity) and whether the instrument meets the assumptions (functional unity, reliability and descriptive statistics) that are required by the analysis techniques that are desired. Additional assumptions for various analysis techniques will be discussed in the forthcoming analysis chapters many of which will be based upon the very assumptions discussed here, i.e., that one's measures are reliable, valid and free from measurement error.
The analysis phase of theory construction is intended to put the theoretical hypotheses developed in the formulation phase to test. It assumes that the activities in the other phases were carried out in a meaningful fashion so that results from the statistical tests can be interpreted. This means that the concepts and linkages in the theoretical propositions are assumed to correspond to reality. Likewise the operational or empirical measures and linkages are assumed to correspond to their counterparts in both the theoretical propositions and the real world. Finally it is assumed that the data gathering was done in such a way as to provide a set of variables which are reliable, valid, and free from measurement error.

Assuming that these statements are more or less accurate, the theorist must decide which statistical test, if any, to use on the basis of his/her problem statement, hypotheses, level at which the variables were measured, and the degree to which assumptions can be met. The next three chapters will be concerned with examining some common techniques which can be used with the kind of data (interval or interval-like) and theoretical background which has been developed throughout earlier sections of this text.
"Some social scientists tend to see theory and methodology as separable, and each as something quite apart from substantive research. The general point that needs to be stressed, however, is that theoretical assumptions underlie everything we do in statistical analyses" (Namboodiri, et al., 1975).

The statement above expressed by Namboodiri, et al., in their introduction, illustrates the growing awareness and concern for the integration of theory, measurement, statistics and inference. This concern was first elucidated in Part I and has been continued throughout the remainder of this dissertation. Thus the discussion of analytic techniques in this section will take place within the substantive framework that was first introduced in Chapter Six rather than as isolated statistical approaches. In keeping with this emphasis, the first topic to be discussed will be the stating of hypotheses, particularly as they relate to statistical analyses techniques.

Hypotheses and Statistical Techniques

The writing of general hypotheses or propositions involves more than just merely making a statement about two or more concepts. A proposition should state the nature of the relationship being proposed (hypothesized) between the
concepts. The form for such statements was introduced in Chapter Eight. The terms which are used to express that nature were called linkage_words. Generally speaking, little attention has been paid to the selection of appropriate linkage words. Common linkage words consist of the following: "are related," "are positively (or negatively) related," or some form of "the higher, the higher (or lower)." But having made such statements, what do we really know about the relationship being hypothesized? Furthermore, how does one test the proposed relationship? Will any statistic do the job? That is, does the selection of the statistic depend simply on whim? Or do certain statistics test specific types of relationships?

Let us consider the following example:

**General_Hypothesis_#1:** Socialization is related to role performance.

What does this proposition tell us about the two concepts? What kind of a theory could be developed from such a statement? Actually this proposition has very little meaning, since any of the relationships depicted in Figure 10.1 could conceivably be covered by such a statement. Suppose the actual relationship was curvilinear, as depicted in Figure 10.1(b). This implies that role performance will increase with more socialization, but only to a point, then the more socialization, the less will be the role
performance. If we applied a linear correlational technique to test the hypothesis, it is unlikely that it would be upheld - not because the hypothesis is untrue (there is a relationship) but rather because the lack of specificity allowed an inappropriate test to be conducted. The same might also be true of the threshold relationship in 10.1(c). At any rate, the actual relationship would not be revealed by linear correlation even if the hypothesis found some support. Instead, a little support for the hypothesis would probably be interpreted as support for Figure 10.1(a), the linear relationship.

This example should have indicated the importance of using linkage terms that clearly specify the relationship between the variables so that an appropriate statistical test
can be selected. An examination of the types of hypotheses and how they are related will demonstrate this need even further. I will discuss them in terms of the types of hypotheses, described earlier in Chapter Eight. These consist of

1. General Hypothesis or Proposition
2. Empirical Hypothesis
3. Theoretical Statistical Hypotheses
   - Null
   - Alternative
4. Empirical Statistical Tests
   - Null
   - Alternative

In a study, then, for each general hypothesis or proposition, there should also be an empirical hypothesis, which substitutes indicators of the concepts in the general hypothesis, and two statistical hypotheses. In the past researchers have worked in terms of the statistical test rather than their hypotheses. However, when inferences were made back to the population, they were in fact being made by means of a theoretical statistical hypothesis. The relationship between a test and the hypotheses can be seen in Figure 10.2. Furthermore, the statement by Namboodiri, et al., (1975) quoted at the beginning of this chapter relates directly to the integration of statistics, theory, and measurement as depicted earlier in Chapter Eight and reproduced in Figure 10.2 on the next page. A thorough
1. General Hypothesis:
   Socialization is related to Role Performance

3a. Theoretical Statistical Hypotheses:
   Null: \( \rho_{soc.\text{FP}} = 0 \)
   Alt: \( \rho_{soc.\text{RP}} \neq 0 \)

2. Empirical Hypothesis:
   Socialization Scores will be related to Role Performance scores

3b. Empirical Statistical Test:
   Null: \( r_{soc.\text{RP}} = 0 \)
   Alt: \( r_{soc.\text{RP}} \neq 0 \)

(Decision Rule):
Accept \( H_0 \) if \( r \) is not significant at \( \alpha \) level

Figure 10.2: Hypotheses and statistical tests
explanation of the interrelationships between these hypotheses was given in Chapter Eight and will not be repeated here, except to re-emphasize that concepts, measures and linkages must be consistent with each other and with the statistical technique that is selected.

Let us examine a number of alternative statements concerning the relationship between socialization and role performance. Note that while the concepts remain the same, the linkage terms change, and in so doing there is a subsequent change in the kinds of questions that are being addressed.

General Hypothesis #2: As socialization of an individual increases his/her role performance increases.

General Hypothesis #3: The more the socialization of an individual, the higher his role performance will be.

What are the linking terms?

What questions are these two hypotheses addressing?

a. Are socialization and role performance related?
b. Do socialization and role performance vary together?

What questions are they not specifically addressing?

a. Does socialization cause or produce role performance?
b. To what extent will a change in socialization affect a change in role performance?

What type of inferential statistical test do these hypotheses imply?
General Hypothesis #4: Role performance is a linear function of socialization, the slope of which is significant.

General Hypothesis #5: The mean value of role performance will change in terms of changes in socialization.

General Hypothesis #6: A unit change in socialization will produce a corresponding change in role performance with an average magnitude of "".

What are the linking terms?

What questions are these hypotheses addressing?

a. Can we predict changes in role performance from changes in socialization?
b. To what extent will a change in socialization affect a change in role performance?

What type of inferential statistical test do these hypotheses imply?

General Hypothesis #7: The role performance of individuals who have received a high level of socialization will vary significantly from those who have received little socialization.

General Hypothesis #8: The amount of socialization will vary significantly for individuals with a high level of role performance from those with a low level of role performance.

What are the linking terms?

What questions are these hypotheses addressing?

a. Hypothesis #7 is examining the affect of different levels of socialization on Role Performance. Do groups with different levels of socialization vary on degree of role performance?

b. Hypothesis #8 is examining various levels
of role performance to see if they differ significantly in the average amount of socialization which each level has received.

Profile or Quartile Analysis

descriptive framework — what are the characteristics (in terms of mean levels of different independent variables) of high versus low role performance?

How, if any, do these two hypotheses differ from one another in the questions they are investigating?

In what ways do these two hypotheses differ from those in the other two sets, that is, to what extent are they answering different questions?

What type of inferential statistical test do these two hypotheses imply?

These hypotheses represent only a few of the ways in which we could have stated the relationship between socialization and role performance. The decision on how a researcher will phrase his hypothesis, (i.e., the linkage terms he uses), should be based on the type of theoretical questions he can legitimately expect to answer, given the nature of his (or her) theory, sample and measures.

What kinds of questions do each of these hypotheses address then? Essentially there is little difference between Hypothesis 2 and 3. Both seem to be stating that there is a positive linear relationship between the two concepts. Without additional information, however, it is not clear from the hypotheses whether one variable occurs prior to the other or whether one variable "causes" or effects the other.
this limited information then, the most appropriate statistic would seem to be one that tests no more than covariation. Assuming the measures meet the assumptions, these two hypotheses could be tested with linear correlation coefficients.

The next set of hypotheses (numbers 4, 5 and 6, are considerably different from the set above). The relationships described by the linking terms are much more specific. In Hypothesis 4, Role Performance is said to be a function of changes in socialization, i.e.

\[
RP = f(soc) \\
RP = \beta_0 + \beta_X + \beta_i soc soc + \varepsilon_i
\]

which translates into the following theoretical statistical hypothesis:

---

\[
\text{Ho: } \beta_{RP(soc)} \leq 0 \\
\text{Ha: } \beta_{RP(soc)} > 0
\]

---

Figure 10.3: Test for Hypothesis 4
Hypothesis 5 is equivalent to hypothesis 4 although it approaches the test somewhat differently. In this case we are examining the predicted line instead of the slope.

Figure 10.4: Illustration of the test for Hypothesis 5.

If we examine the terminology in Hypothesis 5 we see again that it says that "the mean value of role performance will change in terms of unit changes in socialization." The term unit change is taken to mean that we assume fixed X's or fixed socialization scores. In other words, we will be examining the distribution of role performance scores at each socialization value. The mean value \( \mu_{y|x} \) of role performance, given socialization, for each distribution will fall on a line representing the predicted values. Thus for \( X_{soc} = 15 \) we will get one \( \hat{y} \) or predicted value but,
assuming we have 6 respondents who had a socialization score of 15, we will get 6 residuals ($\xi_i$) where $Y_i - \hat{Y} = \xi_i$. The test for this hypothesis then concerns the correlation between the predicted line $\hat{Y}$ and the actual line $Y$. In terms of the theoretical statistical hypothesis, we have

$$Ho: \ R = 0 \quad \text{Where} \ R \text{ is the multiple correlation coefficient and represents the correlation}$$

$$Ha: \ R \neq 0 \quad \text{between} \ Y \text{ and} \ \hat{Y}.$$ 

Although still in a regression framework, Hypothesis 6 is somewhat different from 4 and 5. It is attempting to say that the slope of the Role Performance line remains the same throughout the entire range of socialization scores. Most of

---

**Figure 10.5:** Illustration of Hypothesis 6.
the time we just assume that it remains the same. However, this may be a false assumption. One would probably turn to the introduction of dummy variables in order to test this type of proposition.

The last set of hypotheses, numbers 7 and 8, involve both continuous and category variables. In the six hypotheses above, both variables were assumed to be continuous. As we shall see, this has an impact on the interpretation and testing of a hypothesis. In Hypothesis 7, role performance is the continuous variable and it is said to vary according to the level (category) of socialization. In other words, the mean value of role performance in the low socialization group would be significantly different from the mean value of role performance in a medium or high socialization group. This has been called an experimental or causal framework of analysis of variance (ANOVA), since it implies that the level of socialization one receives will effect the degree or amount of role performance. Assuming three categories of socialization, the theoretical statistical hypotheses would state:

\[ H_0: \mu_1 = \mu_2 = \mu_3 \quad \text{(R.P.)} \]
\[ \text{low med hi} \quad \text{(Soc)} \]

\[ H_a: \text{at least some of } \mu \text{'s are different} \]
Hypothesis 8 differs from 7 in that there is a reversal of variables. The continuous (but not dependent) variable is now socialization, rather than role performance. In Hypothesis 7, the dependent (continuous) variable was Role Performance. Each level of socialization (the categorical independent variable) was expected to produce a different mean value of role performance. The notion of production is missing however in the profile framework. The emphasis here is more on description. The respondents are divided into groups according to their levels of role performance and an attempt is made to distinguish between these categories according to the amount of a certain property (socialization) that each role performance category possesses. This leads to the following type of theoretical statistical hypothesis:

\[ H_0: \mu_1 = \mu_2 = \mu_3 \quad \text{(Socialization)} \]
\[ \text{low med hi} \quad \text{(Role Performance)} \]

\[ H_a: \text{At least some of } \mu \text{'s are different} \]

[NOTE: 'What you write on the null hypothesis is the continuous variable.]

This approach is often referred to as a profile framework and is said to communicate well to applied audiences.

Table 10.1 attempts to summarize some of the important aspects of the four types of statistical approaches (correlation, regression and the two ANOVA procedures) just
discussed, as they relate to the various types of hypotheses. In the next few sections, I will examine bivariate (2 variable) forms of these techniques in greater depth, emphasizing the interpretations that can result from their use.

Correlation

Up to this point the discussion concerning these four statistical approaches has been on a general level. The remainder of the discussion will be of a more specific nature, centering around the application of analysis techniques to the example data set. Correlation will be discussed first. Correlation coefficients examine the amount of covariation between two concepts as demonstrated in the formula below:

$$r = \frac{\text{covariance}_{XY}}{\text{std. dev. of } x \cdot \text{std. dev. of } y}$$

As such it is a relatively simple technique to understand and manipulate, which partially explains why it is so commonly utilized. However, it suffers from a number of limitations, not the least of which is the type of conclusion one can draw from its results. Correlation examines a two-way relationship but cannot distinguish between antecedent and effect. Furthermore, one must be careful that
Table 10.1: Hypotheses and statistical techniques

<table>
<thead>
<tr>
<th>CORRELATION</th>
<th>SINGLE REGRESSION</th>
<th>CAUSAL</th>
<th>ANOVA</th>
<th>PROFILE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. POPULATION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypotheses 2 &amp; 3</td>
<td>Hypothesis 4</td>
<td>Hypothesis 7</td>
<td>Hypothesis 8</td>
<td></td>
</tr>
<tr>
<td>Hypotheses:</td>
<td>Hypotheses:</td>
<td>Hypotheses:</td>
<td>Hypotheses:</td>
<td></td>
</tr>
<tr>
<td>Ho: p ≤ 0</td>
<td>Ho: p ≤ 0</td>
<td>Ho: ( \mu_1 = \mu_2 = \mu_3 )</td>
<td>Ho: ( \mu_1 \neq \mu_2 )</td>
<td></td>
</tr>
<tr>
<td>Ha: p &gt; 0</td>
<td>Ha: p &gt; 0</td>
<td>low med hi (soc)</td>
<td>low med hi</td>
<td></td>
</tr>
<tr>
<td>Model: Bivariate</td>
<td>Model: ( Y = \beta_0 + \beta_1 X + \epsilon )</td>
<td>Ha: At least some ( \mu_i )'s are not equal.</td>
<td>Ha: At least some ( \mu_i )'s are not equal.</td>
<td></td>
</tr>
<tr>
<td>normal where both</td>
<td>where ( Y ) is normally dist for each ( X ), but ( X )'s need not be normally distributed, assumes equal variance at each ( X ).</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( X ) and ( Y ) are assumed to be normal.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **B. SAMPLE**        |                             |                               |                            |                                     |
| Substitute measures or indicators for concepts above | Substitute measures or indicators for concepts above. | Substitute measures for concepts above. | Substitute measures for concepts above. |                                     |
| Ho: \( r = 0 \)      | Ho: \( b \leq 0 \)          | Ho: \( \bar{Y} = \bar{Y} = \bar{Y} \) (RP) | Ho: \( \bar{Y} = \bar{Y} \) (Soc) |                                     |
| Ha: \( r \neq 0 \)   | Ha: \( b > 0 \)             | low med hi                    | low med hi (RP)             |                                     |
| Use t test for       | Use t or F test on          | Ha: At least some \( \bar{Y} \)'s are different. | Ha: At least some of the \( \bar{Y} \)'s are different. |                                     |
| significance         | SS or \( R^2 \)             |                               |                            |                                     |
the relationship which is found is a true rather than a spurious one caused by mutual variation with a third variable. For example:

\[ X_1 \rightarrow X_2 \]

we may find this when it really is this

In making inferences on the basis of correlations, one must also consider the distinction between statistical significance and substantive significance. If we look at the correlations in Table 10.2, on the next page, we see that all of the variables listed are "significantly" related to role performance, at least in a statistical sense. However, statistical significance, in and of itself, is not enough information on which to base one's conclusions. There are a number of other criteria which should also be considered. These are:

2. Sample size
3. Degree to which assumptions have been met
4. Sociological significance
5. The state of theory

Let us consider the relationship of socialization and Role Performance. The correlation is 0.477. How do we evaluate its significance in terms of this relationship. The
Table 10.2: Correlation Table

<table>
<thead>
<tr>
<th></th>
<th>VAR034</th>
<th>VAR037</th>
<th>VAR058</th>
<th>VAR083</th>
<th>VAR102</th>
<th>VAR110</th>
<th>VAR117</th>
<th>POLPT^a</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR034</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR037</td>
<td>0.029</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR058</td>
<td>0.205</td>
<td>0.259</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR083</td>
<td>0.063</td>
<td>0.156</td>
<td>0.220</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR102</td>
<td>0.268</td>
<td>0.173</td>
<td>0.207</td>
<td>0.277</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR110</td>
<td>0.364</td>
<td>0.302</td>
<td>0.557</td>
<td>0.368</td>
<td>0.347</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAR117</td>
<td>0.198</td>
<td>0.280</td>
<td>0.393</td>
<td>0.230</td>
<td>0.136</td>
<td>0.529</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>POLPT</td>
<td>0.236</td>
<td>0.332</td>
<td>0.510</td>
<td>0.452</td>
<td>0.330</td>
<td>0.637</td>
<td>0.477</td>
<td>1.000</td>
</tr>
</tbody>
</table>

^a Takes an r of .116 at .05 level of significance for 200 d.f.
statistical significance of a correlation is affected by both its magnitude and sample size. The sample size (N) is 240 which means that both the sample and magnitude are large enough to lend support to statistical significance.

Nevertheless, it is important to evaluate the other criteria as well. One such criterion concerns assumptions. Correlation tends to be particularistic to the sample under investigation. The representativeness of the sample is an important assumption. Homoscedasticity (variation in each variable is the same throughout the population) is also important. Large differences in standard deviations between two variables will give biased coefficients. Similarly correlations across groups will be affected if the variances for the groups are not equal on those variables.

An assumption which is frequently violated concerns the presence of measurement error. The difference between two correlation matrices may be due entirely to measurement error rather than true differences or the result of unequal variances. At the present time there is no way to test correlation matrices for measurement error. Furthermore, unlike regression coefficients which are affected by measurement error in the independent variables, correlation coefficients are affected by measurement error in both the
dependent and independent variables. The effect of measurement error on correlation consists of a reduction (attenuation) of the magnitude of the coefficient. Outliers also make the coefficient smaller. Finally, correlation coefficients assume linear relationships, attenuating the coefficients where this assumption is violated. Kish (1959) has proposed a common sense approach to the question of assumption violation, these include:

1. Know what the assumptions are.
2. Attempt to show the degree to which they have not been met.
3. Know the effect of not meeting them and where possible make some correction for the effects. (The probable effects will be discussed in more detail at the end of this chapter.)
4. Keep to a minimum the number of assumptions not met.
5. Use caution when making inferences where one or more assumptions have been violated.

As far as the socialization measure is concerned, the assessment in Chapter Nine indicates that most of the assumptions for correlation were not violated to any great extent. To fully consider the criteria of assumptions with respect to assessing the relationship between socialization and Role-Performance, one would have to examine the role

---

1Actually, in correlation it is not entirely accurate to talk about "dependent" and "independent" variables, since it examines covariation, not cause. These terms were used here to facilitate the comparison of correlation with regression in terms of measurement error problems.
performance measure (indicator) in a similar fashion to that outlined in Chapter Nine.

The fourth and fifth criteria (sociological significance and the state of theory) can be discussed together. It was stated that the correlation between Role Performance and Socialization is 0.477. If this correlation is considerably below that obtained in other studies on role performance, it might indicate that socialization is not as pertinent as other variables with respect to understanding or explaining Role Performance. Only two of the variables (communication and scope) in Table 10.3 have an explained variance, \( R^2 \), of more than 25%. In other words, more than 75% of the role performance variance remains unaccounted for with the other variables, including socialization. Substantively, then, what does this say for the relationship between Socialization and Role Performance?

In a correlational framework, \( R^2 \), (or explained variance), involves the notion of commonality (examination of a two-way relationship). Knowledge of one variable (say socialization) will tell you about the other variable, (role performance), but only to the extent of their common variation. In this case the \( R^2 \) is only 0.227. Given the fact that socialization has held up well on most of the other criteria, one might conclude that socialization might add to an explanation of Role Performance when combined with other
pertinent variables (as in the case of multiple regression).
Before taking up the subject of multiple regression, however,
it would perhaps be beneficial to examine the implications of
single or simple regression. These ideas can then be applied
later to the more complicated aspects found in multiple
regression.

Simple or Single Regression

The correlation coefficient, \( r_{xy} \), is used to describe
the covariation between two variables without stating which
of the variables "causes" the other. The regression
coefficient \( \beta_{yx} \), on the other hand, is used to predict the
dependent variable \( Y \), on the basis of values of the
independent variable \( X \). Table 10.3 contains the results of a
series of single regressions on the dependent variable role
performance. One of the regressions in Table 10.3, (the
regression of role performance on socialization), has been
depicted in Figure 10.6.

Suppose we wanted to estimate individual A's role
performance. If we know the overall mean for the sample, but
nothing else, it would serve as the best estimate of
individual A's performance. In this case, as noted in the
figure, the mean or \( \bar{Y} \) is 962. Suppose we learned that the
individual in question had a score of 27 on the Socialization
scale. What would the best estimate of his role performance
score be then?
<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>$B$</th>
<th>St.Err. of $B$</th>
<th>$R^2$</th>
<th>$R$</th>
<th>St.Err. Regr</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension VAR034</td>
<td>24.525</td>
<td>6.535</td>
<td>0.236</td>
<td>14.083</td>
<td>0.236</td>
<td>0.056</td>
</tr>
<tr>
<td>Selectivity VAR037</td>
<td>121.284</td>
<td>22.344</td>
<td>0.332</td>
<td>29.465</td>
<td>0.332</td>
<td>0.110</td>
</tr>
<tr>
<td>Communication VAR058</td>
<td>8.584</td>
<td>0.939</td>
<td>0.510</td>
<td>83.484</td>
<td>0.510</td>
<td>0.260</td>
</tr>
<tr>
<td>Pervasiveness VAR083</td>
<td>55.542</td>
<td>7.105</td>
<td>0.452</td>
<td>61.107</td>
<td>0.452</td>
<td>0.204</td>
</tr>
<tr>
<td>Salience VAR102</td>
<td>5.461</td>
<td>0.985</td>
<td>0.338</td>
<td>30.742</td>
<td>0.338</td>
<td>0.114</td>
</tr>
<tr>
<td>Scope VAR110</td>
<td>61.992</td>
<td>4.860</td>
<td>0.637</td>
<td>162.704</td>
<td>0.637</td>
<td>0.406</td>
</tr>
<tr>
<td>Socialization VAR117</td>
<td>42.091</td>
<td>5.032</td>
<td>0.477</td>
<td>69.967</td>
<td>0.477</td>
<td>0.227</td>
</tr>
<tr>
<td>Job Sat VAR130</td>
<td>5.448</td>
<td>0.968</td>
<td>0.343</td>
<td>31.681</td>
<td>0.343</td>
<td>0.117</td>
</tr>
</tbody>
</table>

Requires an $F$ of 3.89 at .05 level of significance
Figure 10.6: Regression of Role Performance on Socialization

Socialization Scores

Role Performance Scores

- $y = 983$
- $y = 1154$
- $y = 962$

Amount contributed by Regr

Amount explanation due to mean

$y = 1509$
$y = 1382$
$y = 1130$
$y = 752$
$y = 625$
$y = 499$
$y = 373$

0 9 12 15 18 21 24 27 30 33 36
The overall mean ($\bar{Y}=962$) role performance score for the sample does not take account of the effect of any other variable. On the other hand, the $\mu_{Y\mid X}$ (or the mean of $Y$ for a given fixed $X$ value) estimates $Y$, after considering a given $X$ value. Thus $Y$ can be viewed as a new and improved constructed role performance mean, ($\hat{Y} = 1130$), for the individual who had a score of 27 on socialization. [NOTE: This was illustrated earlier in Figure 10.4.]

We can calculate the new mean or estimated score for individual A's role performance, then, by using the prediction equation, substituting in the values for $b_0$ and $b_1$, given in Table 10.3, as well as the value of individual A's socialization score. This results in a $\hat{Y}$ or a new constructed role performance score of 1130 for individual A, (as well as for any other individuals who had a socialization score of 27).

\[
\begin{align*}
\hat{Y} &= b_0 + b_1 X \\
&= -5.697 + 42.09(27) \\
&= 1130
\end{align*}
\]

**Prediction Equation**

\[
\hat{Y} = b_0 + b_1 X
\]

where $X =$ Socialization Score

for individual A = 27

\[
\hat{Y} = -5.697 + 42.09(27)
\]

\[
\hat{Y} = 1130
\]
Suppose, however, that individual A's actual role performance score \((Y_i)\) was 1154. We can use the known score to evaluate the regression as follows:

\[
1130 - 962 \quad \text{What we gained by knowing the individual's socialization score}
\]

\[
1154 - 1130 \quad \text{What we missed by adding the known socialization scores due to a lack of residual fit (wrong or insufficient variables) and/or pure error.}
\]

In general then:

\[
\sum (\hat{Y} - \bar{Y})^2 = \text{Sum of Squares (SS) Regression, which gives you the amount of gain over the overall sample.}
\]

and

\[
\sum (Y - \hat{Y})^2 = \text{Sum of Squares (SS) Residual, the total correction factor for the sample.}
\]

We can see the relationship between these two figures in the Analysis of Variance table (Table 10.4) on the next page.

The variance \((S_{y.x}^2)\) represents a pooled average which assumes equal variance of \(Y\) at a given \(X\). If we take the square root, we get \(S_{y.x}\) which is the standard error of the regression. According to Table 10.3, this gives us a value of 400.927. This value ought to be compared back to the standard deviation of \(Y\), \((\text{i.e., } \sqrt{S_{y}^2})\), given in Table 9.8. When we use \(\bar{Y}\) to estimate \(Y_i\) the standard deviation, \((S_y)\), is
Table 10.4: Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>$S\cdot S\cdot$</th>
<th>$M\cdot S\cdot$</th>
<th>D.F.</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regr</td>
<td>P</td>
<td>$(Y - \bar{Y})^2$</td>
<td>$MS_{Regr}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>($(# \text{ of indep var})$</td>
<td>(remains the same or increases)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resid</td>
<td>$n - p - 1$</td>
<td>$(Y - \bar{Y})^2$</td>
<td>$S_{y.x}^2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$i$</td>
<td>$y.x$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>($(\text{remains the same or decreases})$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>$n - 1$</td>
<td>$(Y - \bar{Y})^2$</td>
<td>$i$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$i$</td>
<td>$y.x$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>($(\text{always remains a constant})$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
455.112; but when we use the regression of role performance on socialization, we find the standard deviation has been reduced to 400.927.

\[
S = 455.112 \quad \text{(see Table 9.8)}
\]
\[
S = 400.927 \quad \text{(see Table 10.3)}
\]

If we look back at Table 10.4, we see that the SS Regression will remain the same or increase as one adds independent variables to the regression. Likewise so will the \( R^2 \) or explained variance since:

\[
R^2 = \frac{\sum (y - \bar{y})^2}{\sum (y - \bar{y})^2} = \frac{SS_{Regr}}{SS_{Resid}}
\]

This is why some people are very critical of using \( R^2 \) to evaluate the worth of a regression model, since all one needs to do is continue to throw in variables (no matter how insignificant to the relationship) in order to increase the \( R^2 \).
Figure 10.7 Significance of the standard deviation of regression for evaluating models

However, the same thing is not true of the standard error

1. Magnitude of the statistic of the regression. As one adds variables, the $S_{y.x}$ will decrease but only to a point. When you begin adding insignificant variables, you begin fitting error and thus the $S_{y.x}$ will increase rather than decrease. Therefore, evaluation of a model should include an examination of the standard error of regression $S_{y.x}$ in addition to $R^2$. Furthermore, in deciding upon the "best" model, one should choose the model with the minimum standard deviation rather than the largest or maximum $R^2$ in order to get the most efficient model.
Analysis of Variance

Some people are arguing that we should be using category type variables, that our measures are not truly interval and should not be treated as such.¹ What happens then when we examine the relationship between role performance and three categories of socialization? Now, instead of 10 bases

---

![ANOVA of Role Performance on three categories of Socialization]

---

¹ Thomas Wilson, 1971.
(i.e., 9, 12, 15...36 as shown in Figure 10.6) for examining the $Y$, we have only three. [NOTE: The ANOVA (or analysis of variance) of $\bar{Y}$ gives the $Y$ of regression (using the dummy variable regression technique).] The category framework gives us fewer residuals to examine. Suppose, for example, that an individual has an actual role performance score of $Y = 943$ and a socialization score of 15, giving an estimated value of $\hat{Y} = 983$. However, this same individual would now have an estimated value of $\hat{Y} = 647$ using socialization as a category type variable. The residual has been greatly enlarged, illustrating that we have lost information by going to a category type variable. This can be further illustrated by examining standard deviations as we did above.

Table 10.5: Comparison of Residuals using continuous vs category variables

<table>
<thead>
<tr>
<th>Socialization as $A$</th>
<th>Continuous Variable</th>
<th>Category Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score = 15</td>
<td>983</td>
<td>647</td>
</tr>
<tr>
<td>Predicted $\hat{Y}$</td>
<td>-40</td>
<td>+394</td>
</tr>
<tr>
<td>Residual $Y - \hat{Y}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$i$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 10.6: Comparison of variances and standard deviation

<table>
<thead>
<tr>
<th>Type</th>
<th>Variance</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S^2$</td>
<td>207,126</td>
<td>455.112</td>
</tr>
<tr>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S^2$</td>
<td>160,742</td>
<td>400.930</td>
</tr>
<tr>
<td>$y.x$ (10 value scale)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S^2$</td>
<td>169,883</td>
<td>412.17</td>
</tr>
<tr>
<td>$y.x$ (3 value category)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As indicated in the table above, the average deviation about the mean is larger in the 3-value category case than in the 10-value continuous scale case. Thus the continuous type independent variable provides a more efficient model than the category type variable, at least in the present case.

The above example should not, however, be taken as an indictment of the technique of Analysis of Variance (ANOVA). As we attempted to demonstrate earlier in the discussion on hypotheses, the selection of a statistical technique should rest upon the kind of sociological question you are attempting to answer. I have previously discussed ANOVA in terms of two frameworks, causal and profile. I shall now attempt to illustrate these two types with data. Five variables, Selectivity, Communication, Salience, Scope and Socialization were divided into three categories. A sixth variable, tension was divided into two categories. Natural
breaking points in the distribution were used so that categories were fairly similar in terms number of respondents.

Table 10.7 contains the results from an ANOVA with role performance as the criterion variable (the dependent variable) as it was broken down according to the categories of the other six variables. What the table seems to indicate from the F-test results is that there is a significant difference between some of the levels on all the independent variables examined. We have designated this as a causal framework because we are attempting to show by this analysis that the various levels of the independent variable in each case has had an effect on the degree of role performance.

In the profile framework a criterion variable is broken down by various levels of role performance. The reference to profile describes the intent of the analyses in Table 10.8. For example, how do individuals with different levels of role performance differ as to the degree of tension they experience? Do coordinators with low role performance receive significantly less communication than those with higher role performance? In this case, then, role performance serves as the category variable and an attempt is made to describe coordinators with various levels of role performance according to the amount of the criterion or
Table 10.7: Experimental or causal framework - ANOVA

Criterion: Role Performance

<table>
<thead>
<tr>
<th>Broken Down by</th>
<th>Mean</th>
<th>N</th>
<th>St. Dev.</th>
<th>F a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-14</td>
<td>886.047</td>
<td>128</td>
<td>465.505</td>
<td>7.9451</td>
</tr>
<tr>
<td>15-16</td>
<td>1049.669</td>
<td>112</td>
<td>428.522</td>
<td></td>
</tr>
<tr>
<td><strong>Selectivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2</td>
<td>821.480</td>
<td>75</td>
<td>470.544</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>902.609</td>
<td>87</td>
<td>436.782</td>
<td>13.2823</td>
</tr>
<tr>
<td>4-8</td>
<td>1164.603</td>
<td>78</td>
<td>390.167</td>
<td></td>
</tr>
<tr>
<td><strong>Communication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-80</td>
<td>715.918</td>
<td>73</td>
<td>464.956</td>
<td></td>
</tr>
<tr>
<td>81-110</td>
<td>943.253</td>
<td>91</td>
<td>399.371</td>
<td>28.4862</td>
</tr>
<tr>
<td>111-134</td>
<td>1222.092</td>
<td>76</td>
<td>364.944</td>
<td></td>
</tr>
<tr>
<td><strong>Saliency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-200</td>
<td>792.125</td>
<td>72</td>
<td>414.104</td>
<td></td>
</tr>
<tr>
<td>201-230</td>
<td>931.232</td>
<td>86</td>
<td>447.287</td>
<td>12.9987</td>
</tr>
<tr>
<td>231-260</td>
<td>1144.510</td>
<td>82</td>
<td>436.432</td>
<td></td>
</tr>
<tr>
<td><strong>Scope</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-12</td>
<td>599.875</td>
<td>80</td>
<td>390.453</td>
<td></td>
</tr>
<tr>
<td>13-16</td>
<td>1037.984</td>
<td>64</td>
<td>344.577</td>
<td>61.7716</td>
</tr>
<tr>
<td>17-24</td>
<td>1214.125</td>
<td>96</td>
<td>370.149</td>
<td></td>
</tr>
<tr>
<td><strong>Socialization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-20</td>
<td>647.217</td>
<td>60</td>
<td>450.809</td>
<td></td>
</tr>
<tr>
<td>21-25</td>
<td>1000.648</td>
<td>111</td>
<td>429.331</td>
<td>27.1954</td>
</tr>
<tr>
<td>26-26</td>
<td>1174.956</td>
<td>69</td>
<td>342.915</td>
<td></td>
</tr>
</tbody>
</table>

a Requires an $F^2$ of at least 3.04 at .05 level of significance.
Table 10.8: Profile framework - ANOVA

<table>
<thead>
<tr>
<th>Broken Down by</th>
<th>Mean</th>
<th>N</th>
<th>St. Dev.</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Role Performance Criterion Variable: Tension</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>11.112</td>
<td>80</td>
<td>4.976</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>13.050</td>
<td>80</td>
<td>3.891</td>
<td>6.2219</td>
</tr>
<tr>
<td>1251-1850</td>
<td>13.300</td>
<td>80</td>
<td>3.925</td>
<td></td>
</tr>
<tr>
<td><strong>Role Performance Criterion Variable: Communication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>74.550</td>
<td>80</td>
<td>29.522</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>96.400</td>
<td>80</td>
<td>20.855</td>
<td>28.7012</td>
</tr>
<tr>
<td>1251-1850</td>
<td>102.200</td>
<td>80</td>
<td>21.709</td>
<td></td>
</tr>
<tr>
<td><strong>Role Performance Criterion Variable: Pervasiveness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>4.000</td>
<td>80</td>
<td>3.812</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>5.725</td>
<td>80</td>
<td>3.586</td>
<td>23.4144</td>
</tr>
<tr>
<td>1251-1850</td>
<td>7.675</td>
<td>80</td>
<td>2.694</td>
<td></td>
</tr>
<tr>
<td><strong>Role Performance Criterion Variable: Salience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>205.462</td>
<td>80</td>
<td>23.610</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>214.212</td>
<td>80</td>
<td>27.091</td>
<td>12.5419</td>
</tr>
<tr>
<td>1251-1850</td>
<td>226.675</td>
<td>80</td>
<td>24.918</td>
<td></td>
</tr>
<tr>
<td><strong>Role Performance Criterion Variable: Scope</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>11.250</td>
<td>80</td>
<td>4.379</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>14.887</td>
<td>80</td>
<td>3.835</td>
<td>53.3750</td>
</tr>
<tr>
<td>1251-1850</td>
<td>17.600</td>
<td>80</td>
<td>3.430</td>
<td></td>
</tr>
<tr>
<td><strong>Role Performance Criterion Variable: Socialization</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-750</td>
<td>20.025</td>
<td>80</td>
<td>5.136</td>
<td></td>
</tr>
<tr>
<td>751-1250</td>
<td>23.662</td>
<td>80</td>
<td>4.749</td>
<td>26.803</td>
</tr>
<tr>
<td>1251-1850</td>
<td>25.313</td>
<td>80</td>
<td>4.074</td>
<td></td>
</tr>
</tbody>
</table>

\[^{a}\text{Requires an F}^2 \text{ of at least 3.04 at .05 level of significance}\]
Table 10.9: Explanation as related to standard procedures

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>Regression</th>
<th>Causal Framework</th>
<th>Profile Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explained</td>
<td>$R^2=0.227$</td>
<td>$R^2=0.227$</td>
<td>$R^2=0.1866$</td>
<td>$Y = \hat{Y}$</td>
</tr>
<tr>
<td>Variation</td>
<td>$\text{Variance (Commonality)}$</td>
<td>$\text{(Based on SS)}$</td>
<td>$\text{(Note: Residual variance)}$</td>
<td>$R^2=0.1866$</td>
</tr>
<tr>
<td>Examining two-way relation</td>
<td>Coefficient of Determination</td>
<td>Lost $R^2$ by joining one-way relationship to categories</td>
<td>In this case we had a good regression since we had a linear regression.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Y = Y = 20.025$</td>
<td>$Y = Y = 23.66$</td>
<td>$Y = Y = 25.31$</td>
<td></td>
</tr>
<tr>
<td>Residuals</td>
<td>$(1-R^2)$</td>
<td>$(1-R^2)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>alienation residual</td>
<td>$\text{(based on SS)}$</td>
<td>$\text{linear. If have nonlinear relationship it may gain by this procedure}$</td>
<td>$\text{since it looks at both linear and nonlinear aspects; but we could add continuous variable and take care of the curvilinear feature within the regression situation.}$</td>
<td></td>
</tr>
<tr>
<td>that which is coefficient of nondetermination</td>
<td>$\text{examines deviations}$</td>
<td>$\text{since it looks at both linear and nonlinear aspects; but we could add continuous variable and take care of the curvilinear feature within the regression situation.}$</td>
<td>$\text{since it looks at both linear and nonlinear aspects; but we could add continuous variable and take care of the curvilinear feature within the regression situation.}$</td>
<td></td>
</tr>
<tr>
<td>not common between the two variables</td>
<td>$\text{examine deviations}$</td>
<td>$\text{since it looks at both linear and nonlinear aspects; but we could add continuous variable and take care of the curvilinear feature within the regression situation.}$</td>
<td>$\text{since it looks at both linear and nonlinear aspects; but we could add continuous variable and take care of the curvilinear feature within the regression situation.}$</td>
<td></td>
</tr>
</tbody>
</table>
continuous variable that is possessed.¹

The discussion to this point has been concerned with examining the relationship between two variables, the dependent and a single independent variable. Table 10.9 on the next page summarizes some of the major points dealing with types of hypotheses and the various models and assumptions, which have been discussed in this chapter.

Summary

An attempt has been made throughout this chapter to demonstrate the importance of selecting one's statistical analysis technique in a context of an integrated approach to theory construction. It has not been unknown to find published articles which have utilized a statistical analysis technique that had no apparent relation to what was developed in the theoretical portion of the paper. When this occurs, what relevance can the analysis results have for the theory that is supposedly being examined?

¹I have avoided the use of the term dependent and independent in the profile discussion because the intent here is to describe characteristics (mean values of criterion variables) of the various levels of the variable of interest. The mean value of the criterion variable is not said to be "caused" by the group variable, as was the case in the causal framework. Thus the terminology of dependent vs. independent would be misleading in this framework.
To illustrate this point further, a number of general hypotheses were discussed. Although, these hypotheses used the same concepts, the linkages were changed to reflect different types of relationships. Thus different types of statistical techniques were shown to be relevant for the direct test of different types of linkage terms.

Bivariate forms of these statistical techniques (correlation, simple regression, and analysis of variance) were then examined in greater depth using the data set to illustrate the kinds of interpretations that can be made when the assumptions for each technique have been met.

Most relationships in the real world tend to be more complicated than those examined by standard two-variable procedures, that is, most social phenomena have multiple sources of variation rather than simple one-to-one causation. In the next two chapters, I shall discuss procedures which are perhaps more appropriate for examining the multiple "causes" of variation in the real world.
CHAPTER ELEVEN:
MULTIPLE REGRESSION

Multiple regression is a method of analysing the collective and separate contributions of two or more independent variables, $X$, to the variation of a dependent variable, $Y$ (Kerlinger and Pedhazur, 1973:3).

In discussing simple regression, we used the example of role performance on socialization. I shall now move to a multiple regression situation by adding two independent variables (communication and scope) to the model. If we compare the new model to the one with socialization alone, or to the mean role performance score, we see from table 11.1 below that the new model shows a substantial increase in $R^2$ (or explained variation) and a substantial decrease in the standard deviation of the regression, $(S_{Y.X})$.

Table 11.1 Comparison of Models

<table>
<thead>
<tr>
<th>Model</th>
<th>St. Dev.</th>
<th>$R^2$</th>
<th>$R$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{Y}$</td>
<td>455.112</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$Y$ on Soc.</td>
<td>400.927</td>
<td>0.227</td>
<td>0.477</td>
</tr>
<tr>
<td>$Y$ on Soc, Com and Scope</td>
<td>33.486</td>
<td>0.460</td>
<td>0.678</td>
</tr>
</tbody>
</table>
Thus we are still within that optimum range described by Figure 10.7, where the standard deviation is going down while the $R^2$ is going up. The present model can be said then to represent a more efficient estimate of $Y$ than either the mean of $Y$ or the regression of $Y$ on Socialization alone.

If we examine the individual sources of variation as shown in the ANOVA table for the new model, we see that all the partial $B$'s are significant. Furthermore, we see that the slope ($B$) for socialization, (when communication and scope are held constant), has been reduced to 14.702 as contrasted with a slope of 42.091 when socialization was the only variable entered. What does this mean sociologically? It means that these "independent" variables are actually intercorrelated, otherwise the slope for socialization would not have changed when the two additional variables were included.
Table 11.2: ANOV for Role Performance on Socialization, Communication, and Scope. N=240.

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>S.S.</th>
<th>M.S.</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regr</td>
<td>3</td>
<td>22782669.283</td>
<td>7594223.094</td>
<td>67.073*</td>
</tr>
<tr>
<td>Error</td>
<td>236</td>
<td>26720644.513</td>
<td>113223.069</td>
<td></td>
</tr>
<tr>
<td>Corr Total</td>
<td>239</td>
<td>49503313.796</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source</th>
<th>B</th>
<th>T for Ho: B=0</th>
<th>St. Err B</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-303.558</td>
<td>2.846*</td>
<td>106.671</td>
<td>0.460</td>
</tr>
<tr>
<td>Socialization</td>
<td>14.702</td>
<td>2.925*</td>
<td>5.026</td>
<td>0.166</td>
</tr>
<tr>
<td>Communication</td>
<td>3.376</td>
<td>3.446*</td>
<td>0.979</td>
<td>0.200</td>
</tr>
<tr>
<td>Scope</td>
<td>42.554</td>
<td>6.939*</td>
<td>6.133</td>
<td>0.437</td>
</tr>
</tbody>
</table>

*Requires an F = 3.88 and t = 2.617 for .01 level of significance.

Table 11.2 includes another interesting feature, i.e., the negative intercept. A negative intercept implies a negative role performance. In other words, the predicted line crosses the Y (Role Performance) axis at some negative value. The impact of a negative intercept on interpretation is an inability to accurately predict role performance at either the upper or lower ends.
In moving from a simple regression to a multiple regression with three independent variables, the nature of the curve under examination changes from a straight line to a surface. $\hat{Y}$ is on the surface whereas $Y_i$ deviates from that surface, just as it deviated from the line in single regression. Another method of evaluating a model consists of examining the residuals ($Y_i - \hat{Y}$) or the deviation of $Y_i$ from the $\hat{Y}$. Since the example above has an $N=240$, there are 240 residuals to examine. This represents quite a few residuals for the purpose of this discussion. Therefore, I have rerun the regression of role performance on socialization, communication, and scope using a subsample with an $N=36$. Table 11.3 contains the ANOVA information on the reduced sample, while the 36 residuals ($Y_i - \hat{Y}$) can be found in Table 11.4.
Table 11.3: ANOV Role Performance on Socialization, Communication and Scope N=36

<table>
<thead>
<tr>
<th>Source</th>
<th>D.F.</th>
<th>S.S.</th>
<th>M.S.</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reg</td>
<td>3</td>
<td>1849759.065</td>
<td>616586.355</td>
<td>4.058*</td>
</tr>
<tr>
<td>Error</td>
<td>32</td>
<td>4861656.157</td>
<td>151926.755</td>
<td>616586.355</td>
</tr>
<tr>
<td>Corr Total</td>
<td>35</td>
<td>6711415.222</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

St. Dev. Regr. = 389.778

<table>
<thead>
<tr>
<th>Source</th>
<th>B</th>
<th>t for H₀: B=0</th>
<th>St. Err B</th>
<th>B*</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-89.848</td>
<td>-0.239</td>
<td>375.684</td>
<td>0.0</td>
<td>0.276</td>
</tr>
<tr>
<td>Scot</td>
<td>13.501</td>
<td>0.834</td>
<td>16.184</td>
<td>0.138</td>
<td></td>
</tr>
<tr>
<td>Cont</td>
<td>3.573</td>
<td>1.114</td>
<td>3.207</td>
<td>0.242</td>
<td></td>
</tr>
<tr>
<td>Scot</td>
<td>25.848</td>
<td>1.166</td>
<td>22.160</td>
<td>0.249</td>
<td></td>
</tr>
</tbody>
</table>

3
*Requires F = 2/90 and t = 20.036 at .05 level

Examining Residuals

Upon examining the residuals, we find that three-fourths of the residuals have an absolute value of more than 100. Less than one-half (17 of 36) have a residual of less than 260. In addition there are six whose absolute value is over 500. The rather consistently high residuals might indicate that there is considerable "lack of fit" between the model and the actual role performance scores, at least in this particular sample. However, if we compare the entire regression of the subsample with that of the total (Table 11.2), we find that one of the partial regression (B)
<table>
<thead>
<tr>
<th>No.</th>
<th>SCCT</th>
<th>COMT</th>
<th>SCCTOBSERVED</th>
<th>PREDICTED</th>
<th>RESIDUAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
<td>111</td>
<td>17 331.00000000</td>
<td>1329.65729083</td>
<td>-1018.65729083</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>117</td>
<td>20 927.00000000</td>
<td>1321.15429484</td>
<td>-404.15429484</td>
</tr>
<tr>
<td>3</td>
<td>24</td>
<td>80</td>
<td>13 514.00000000</td>
<td>956.01615643</td>
<td>-442.01615643</td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>0</td>
<td>3   182.00000000</td>
<td>190.91717431</td>
<td>-8.91717431</td>
</tr>
<tr>
<td>5</td>
<td>1/4</td>
<td>120</td>
<td>19 1030.00000000</td>
<td>1072.99763436</td>
<td>-42.99763436</td>
</tr>
<tr>
<td>6</td>
<td>21</td>
<td>52</td>
<td>14 477.00000000</td>
<td>944.71578797</td>
<td>-467.71578797</td>
</tr>
<tr>
<td>7</td>
<td>27</td>
<td>105</td>
<td>15 706.00000000</td>
<td>1351.82171513</td>
<td>-645.82171513</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>46</td>
<td>11 913.00000000</td>
<td>601.84305639</td>
<td>311.15694361</td>
</tr>
<tr>
<td>9</td>
<td>24</td>
<td>93</td>
<td>10 509.00000000</td>
<td>924.91717431</td>
<td>-415.91717431</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>89</td>
<td>8   220.00000000</td>
<td>716.42659214</td>
<td>-496.42659214</td>
</tr>
<tr>
<td>11</td>
<td>24</td>
<td>25</td>
<td>0   284.00000000</td>
<td>530.28417926</td>
<td>-246.28417926</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>55</td>
<td>10 1362.00000000</td>
<td>729.66235121</td>
<td>632.66235121</td>
</tr>
<tr>
<td>13</td>
<td>27</td>
<td>93</td>
<td>17 723.00000000</td>
<td>1046.35520251</td>
<td>-323.35520251</td>
</tr>
<tr>
<td>14</td>
<td>27</td>
<td>107</td>
<td>20 1302.00000000</td>
<td>1173.91473707</td>
<td>129.91473707</td>
</tr>
<tr>
<td>15</td>
<td>21</td>
<td>87</td>
<td>19 1516.00000000</td>
<td>995.66587222</td>
<td>520.66587222</td>
</tr>
<tr>
<td>16</td>
<td>27</td>
<td>123</td>
<td>14 1644.00000000</td>
<td>1083.13567744</td>
<td>560.13567744</td>
</tr>
<tr>
<td>17</td>
<td>18</td>
<td>82</td>
<td>17 861.00000000</td>
<td>965.54303350</td>
<td>-104.54303350</td>
</tr>
<tr>
<td>18</td>
<td>24</td>
<td>127</td>
<td>21 655.00000000</td>
<td>1230.71017667</td>
<td>-575.71017667</td>
</tr>
<tr>
<td>19</td>
<td>33</td>
<td>132</td>
<td>19 1303.00000000</td>
<td>1318.39136880</td>
<td>-15.39136880</td>
</tr>
<tr>
<td>20</td>
<td>27</td>
<td>114</td>
<td>21 1476.00000000</td>
<td>1224.77067951</td>
<td>251.77067951</td>
</tr>
<tr>
<td>21</td>
<td>21</td>
<td>93</td>
<td>18 734.00000000</td>
<td>951.19362034</td>
<td>-217.19362034</td>
</tr>
<tr>
<td>22</td>
<td>13</td>
<td>109</td>
<td>13 622.00000000</td>
<td>879.61299049</td>
<td>-257.61299049</td>
</tr>
<tr>
<td>23</td>
<td>18</td>
<td>121</td>
<td>16 1007.00000000</td>
<td>999.02798556</td>
<td>115.02798556</td>
</tr>
<tr>
<td>24</td>
<td>24</td>
<td>81</td>
<td>14 972.00000000</td>
<td>985.43641937</td>
<td>13.43641937</td>
</tr>
<tr>
<td>25</td>
<td>30</td>
<td>113</td>
<td>16 1465.00000000</td>
<td>1132.46429291</td>
<td>332.46429291</td>
</tr>
<tr>
<td>26</td>
<td>24</td>
<td>121</td>
<td>16 1216.00000000</td>
<td>1206.36287333</td>
<td>9.36287333</td>
</tr>
<tr>
<td>27</td>
<td>21</td>
<td>93</td>
<td>13 1323.00000000</td>
<td>987.00333032</td>
<td>336.00333032</td>
</tr>
<tr>
<td>28</td>
<td>24</td>
<td>112</td>
<td>17 1466.00000000</td>
<td>1373.73637917</td>
<td>89.73637917</td>
</tr>
<tr>
<td>29</td>
<td>21</td>
<td>72</td>
<td>13 716.00000000</td>
<td>786.93077534</td>
<td>-70.93077534</td>
</tr>
<tr>
<td>30</td>
<td>27</td>
<td>111</td>
<td>14 763.00000000</td>
<td>1033.11929194</td>
<td>-270.11929194</td>
</tr>
<tr>
<td>31</td>
<td>27</td>
<td>91</td>
<td>10 692.00000000</td>
<td>1325.08765709</td>
<td>-633.08765709</td>
</tr>
<tr>
<td>32</td>
<td>27</td>
<td>114</td>
<td>21 1306.00000000</td>
<td>1224.77067951</td>
<td>81.77067951</td>
</tr>
<tr>
<td>33</td>
<td>18</td>
<td>111</td>
<td>17 677.00000000</td>
<td>989.41881574</td>
<td>-312.41881574</td>
</tr>
<tr>
<td>34</td>
<td>24</td>
<td>126</td>
<td>16 1744.00000000</td>
<td>1145.59641305</td>
<td>598.59641305</td>
</tr>
<tr>
<td>35</td>
<td>24</td>
<td>80</td>
<td>9   782.00000000</td>
<td>752.26557657</td>
<td>29.26557657</td>
</tr>
<tr>
<td>36</td>
<td>21</td>
<td>111</td>
<td>16 921.00000000</td>
<td>1044.31011623</td>
<td>-123.31011623</td>
</tr>
</tbody>
</table>

**SUM OF RESIDUALS =** 0.00000000  
**SUM OF SQUARES OF RESIDUALS =** 4861656.15739711  
**SUM OF SQUARES OF RESIDUALS - ERROR SS** = 4861656.15739711  
**FIRST ORDER AUTOCORRELATION OF RESIDUALS =** 0.04676497  
**DURBIN-WATSON D =** 1.473214046
coefficients is significant in the subsample whereas they all are in the total sample. Perhaps we should contribute the high residuals to sampling error (at least in the subsample) as much as to the possible specification error cited above.

Observations 5 and 23 illustrate the situation just described. Note the similarity of values on the three independent variables. This results in predicted $\hat{Y}$

<table>
<thead>
<tr>
<th>OBS</th>
<th>SOCT</th>
<th>COMT</th>
<th>SCOT</th>
<th>OBSY</th>
<th>PREDY</th>
<th>RESID</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>18</td>
<td>120</td>
<td>19</td>
<td>1838</td>
<td>1072</td>
<td>765</td>
</tr>
<tr>
<td>23</td>
<td>18</td>
<td>121</td>
<td>16</td>
<td>1007</td>
<td>999</td>
<td>7.97</td>
</tr>
</tbody>
</table>

values which are also fairly close (1072 vs. 999). However, the residuals ($Y_i - \hat{Y}$) are considerably different ($\xi_5 - \xi_{23}$) = $(765 - 7.97 = 757.03)$. Obviously something (either sampling error, measurement error, or specification error, i.e., not the correct set of variables) is causing this additional variation. [NOTE A technique for breaking down the variance into true variation, measurement error and specification error was described in Warren et al., (1974) and Warren et al., (1975).]

There are a number of plotting techniques which can be used to examine residuals. Figure 11.1 depicts the plot of $Y$
Figure 11.1: Plot of $Y$ vs. $\hat{Y}$, ($N = 36$)
Figure 11.2: Plot of $Y$ vs. $\hat{Y}$, (N = 240)
versus \( \hat{Y} \) values for \( N=36 \) as shown in Table 11.4. What would the plot have looked like if there had been a "perfect fit"? Certainly the problems discussed above, with reference to the size of the residuals, is obvious from the plot. Compare this to the plot of \( Y \) versus \( \hat{Y} \) for the total sample \( (N=240) \) in Figure 11.2. While it is true that the correspondence between \( Y \) and \( \hat{Y} \) is still far from perfect, one can get at least a feel for a linear (straight line) relationship in the total sample plot. Anyway there tends to be more of a band of points with the total sample than with the subsample plot.

"Degree of fit" raises the important question in multiple regression of how many variables should you add. I mentioned earlier that many people concentrate on trying to maximize the \( R^2 \). However, this, by itself, is not a sufficient criteria because even irrelevant variables can increase the \( R^2 \). To illustrate this further, Table 11.5 examines some alternative models.

Comparison of Alternative Models

All of the regressions listed in Table 11.5 are significant with the exception of model eight. When tension was added to the variables of model seven, it failed to reach significance. But note that even so, there was still a slight increase in the \( R^2 \) (from .5363 to .5367). This point is even more obvious, however, when we examine the amount of gain that was made by adding variables to the four variable
Table 11.5: Examining alternative models

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>$R^2$</th>
<th>$s^2$</th>
<th>$s$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>One Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCOT</td>
<td>.4060</td>
<td>.6372</td>
<td>17,959,802.246</td>
</tr>
<tr>
<td>SOCT</td>
<td>.227</td>
<td>.477</td>
<td>25,390,233.117</td>
</tr>
<tr>
<td><strong>Two Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCOT, PERVT</td>
<td>.4608</td>
<td>.6786</td>
<td>15,900,464.293</td>
</tr>
<tr>
<td><strong>Three Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONT, SCCT, PERVT</td>
<td>.4938</td>
<td>.7027</td>
<td>14,718,335.193</td>
</tr>
<tr>
<td>SOCT, SCCT, CONT</td>
<td>.4600</td>
<td>.6782</td>
<td>15,930,166.381</td>
</tr>
<tr>
<td><strong>Four Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCT, COMT, SCOT, PERVT</td>
<td>.5107</td>
<td>.7106</td>
<td>14,326,259.014</td>
</tr>
<tr>
<td><strong>Five Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCCT, COMT, SELT, SCOT, PERVT</td>
<td>.5203</td>
<td>.7213</td>
<td>13,806,673.556</td>
</tr>
<tr>
<td><strong>Six Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCT, COMT, SELT, SCOT, PERVT, JEST</td>
<td>.5289</td>
<td>.7273</td>
<td>13,509,653.673</td>
</tr>
<tr>
<td><strong>Seven Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCT, COMT, SELT, SCOT, PERVT, SALT, JBST</td>
<td>.5363</td>
<td>.7323</td>
<td>13,253,137.104</td>
</tr>
<tr>
<td><strong>Eight Variable Models</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Ten not significant)</td>
<td>.5367</td>
<td>.7326</td>
<td>13,238,186.110</td>
</tr>
</tbody>
</table>
model. The gain in $R^2$ by adding socialization to the three variable model, (communication, scope and pervasiveness), was .0169, about twice the effect of adding a variable to each of the subsequent models:

$$R^2_{5 \text{ var model}} - R^2_{4 \text{ var model}} = .5203 - .5107 = .0096$$

$$R^2_{6 \text{ var model}} - R^2_{5 \text{ var model}} = .5289 - .5203 = .0086$$

$$R^2_{7 \text{ var model}} - R^2_{6 \text{ var model}} = .5363 - .5289 = .0074$$

$$R^2_{8 \text{ var model}} - R^2_{7 \text{ var model}} = .5367 - .5363 = .0004$$

There is an obvious drop in the amount of gain made by adding a variable to the 3 variable model versus adding variables to the models thereafter. Furthermore, only 2.6% gain was made by adding four variables to the 4 variable model. One must consider whether the 2% gain was worth the addition of that many variables, i.e., do we still have an efficient model? (On an applied level this question is especially important in terms of what are the additional costs involved in concentrating on 7 or 8 variables instead of four?) Certainly there was an increase in $R^2$ with the addition of these extra four variables, but the question remains to what extent does the addition of these variables add to our knowledge and understanding of role performance?
If we examine the models in terms of variance or standard deviation, we note by the figures in Table 11.5 that the deviation begins to level off quite a bit after the four variable model. On this basis then, the four, or maybe five variable, model looks like the most parsimonious, i.e., the model which provides the most knowledge and understanding with the least number of variables.

In examining the three variable model:

\[ \hat{Y} = -89.848 + 13.50X_{soc} + 3.573X_{comm} + 25.848X_{scope} \]

for the subsample \( (N=36) \), three types of errors were discussed. These consisted of the following:

1. Sampling error
2. Measurement error
3. Specification error

The question of selecting the appropriate model, as discussed in terms of Table 11.5, is also concerned with the specification problem. In general, researchers find themselves in one of three situations with respect to the specification problem.

**Specification and the research situation**

It is possible to view the research situation in terms of a continuum ranging from a situation C where little or no theoretical or empirical evidence exists to situation A where the theoretical model has already been completely specified.
**Situation A:**
We know the model, i.e., the model has been completely specified or established by theory. This situation is therefore not interested in theory building or testing, rather in pure estimation, (what is the size of not its significance).

**Situation B:**
Partially known model, i.e., theory has established that certain variables are important elements in a model. However, other variables remain in question. We must turn to the empirical situation to help us establish what these other variables are. Both theory building and testing are often involved in this situation.

**Situation C:**
Have guidelines but unspecified model, i.e., theoretical development is at a minimum, model specification must depend entirely on the empirical situation. This situation then primarily involves theory building techniques.

<table>
<thead>
<tr>
<th>Situation C</th>
<th>Situation B</th>
<th>Situation A</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Information</td>
<td>Some Information</td>
<td>Model Completely Specified</td>
</tr>
</tbody>
</table>

Figure 11.3: Continuum of research situations

This chapter has assumed that many people know how to manipulate the various analytical techniques, but have difficulty interpreting their results. One reason we have

---

1Note: I am indebted to Dr. Richard Warren, Iowa State University, for the initial conceptualization of the analytical approach in terms of this framework.
trouble with interpretation may well be because we fail to specify in the beginning what we want to do. Thus before writing our hypotheses, we must first decide with which of the three situations we are dealing. For instance, regression involves the following types of tests:

1. \( H_0: \) All \( \beta_i = 0 \)
   
   The overall regression is not significant or the population Multiple Correlation = 0.

2. \( H_0: \) A particular \( \beta_i = 0 \)

3. Estimate values for coefficients

Note that the first two tests are concerned with questioning the model whereas the third one is only concerned with the magnitude or size of the coefficient. Therefore, if we first make clear the nature of our situation and thus what we want to prove, then develop the appropriate type of hypotheses, interpretation of analysis results will be greatly facilitated.

**Situation A**

Let's examine the three types of situations more closely. Suppose we are in Situation A where we know the model. According to theory then, Role Performance would be considered a function of socialization, communication, and scope, i.e.,
For the moment, let us assume further that the regression in Table 11.5,

\[ \hat{Y} = -303.558 + 14.702X + 3.376X + 42.554X \]

on the total sample represents the relationship of these three variables to role performance in the population. If we then examine the regressions on the three subsamples:

\[ \hat{Y} = -89.848 + 13.50X + 3.573X + 25.848X \]
\[ \hat{Y} = -62.749 + 2.984X + 8.865X + 11.834X \]
\[ \hat{Y} = -368.979 + 13.673X + 1.439X + 65.963X \]

which of the tests mentioned above would be appropriate? If we are truly under Situation A, where we know the model, there would be no reason for conducting tests 1 and 2, since these are concerned primarily with model modification, unless you were testing a sampling question. In this case the hypothesis would not be, "Role Performance is related to Socialization, Communication and Scope." Instead, you would be testing the question - "Is my sample representative?" Usually, however, the interest under this situation revolves
around the attempt to estimate the appropriate coefficients. If these estimates prove unsatisfactory, we must consider one or more of the following to be the cause:

A. Bad sample
B. Measurement Problems
C. We didn't really know the model, a Situation B approach would have been better.

Situation B:
In Situation B, past research and theory have established that certain variables should be included in a model but that these are not sufficient for a completely specified model. In other words,

\[ Y = \beta_0 + \beta_1 X + \beta_2 X + \beta_3 X + \beta_4 X + \xi_i \]

Know these have to be in the model
There is some question as to whether these go into the model also

Suppose you know that socialization, communication and scope are in the model and you suspect that tension should also be important. You might then hypothesize that these four variables are meaningful in predicting role performance. Test #1 (all variables in combination are significant) would then be appropriate. You might also propose four subhypotheses that each variable will make a unique
contribution, in which case test #2 (each $\beta_i \neq 0$) would be used. If you stopped with these questions, you would remain entirely within a testing framework. Did you explain enough of the variance? Should some variables be dropped and/or others added?

In the theory building framework, you allow theory to determine the original model but then drop and/or add variables (modify the model) according to the dictates of the empirical data. In this context, you could use a specific equation and compare it to an alternative equation.

$$F = \frac{R^2 - Y_{\hat{4}21}^2}{(1 - R^2)}/(n-k - 1)$$

Thus where there are only a few variables in question, the backward solution type of regression technique might be beneficial. Perhaps instead of using the term model building or developing, we should use model refining. We could then test a fifth subhypothesis that a refined model would do as well as say the original model.

**Situation C:**

Pure model development or building, where there are few guidelines from theory, takes place within Situation C. Generally, this involves the examination of a large number of variables for the purpose of selecting out those few which
meet certain specifications, depending upon the regression
technique that is utilized. Three of the most widely used
techniques are compared in Table 11.6. However, before
discussing the model selection techniques, a final word about
how testing should be made, especially as it applies to low B
and C situations is in order.

A single sample should not be used to do the three types
of tests discussed earlier. In other words, one should not
use the same sample to both determine (build) the model
(tests 1 and 2) and estimate the parameters. Suppose you
determine the wrong model. If you are using a single sample
you will get no indication of whether it is right or wrong.
If it is the latter, you will instead get only worthless
estimates. If one is interested in both theory building and
estimating coefficients, there are at least three alternative
approaches that can be used.

1. Use two samples. Use the first sample
to build or determine the model and the
second sample to estimate the size of the
parameters.

2. Use one large initial sample but split
the total in half and use as in approach
number one above.

3. Use multiple measures of the concepts -
splitting them and using one set to
determine the model and the other set to
estimate values. [Note: To use the
measurement approach, one must assume little
measurement error, high reliability, validity
and intercorrelation of the measures.]
In Situation C, theoretical development generally consists of little more than a list of suggested variables. Not only that, the list of potential variables is often quite large. A number of regression techniques have been developed to help reduce that number to a minimum set, while still attempting to maximize the variance that is explained. These include:

1. All possible regressions
2. Background procedure
3. Forward procedure
4. Stepwise

*Regression_Model_Selection_Procedures* In the "all possible" regression technique, every possible combination of variables is run. Usually the model yielding the largest $R^2$ (within certain ranges of significance) is the one which is selected. In order to understand the benefits or problems of each of these selection approaches, one should consider them within the context of the A, B and C frameworks discussed above. Thus a backward approach, which gives the researcher a more wholistic picture, might be extremely beneficial in a B framework where most of the model is known and the list of potential variables is not large. On the other hand, in a C framework with an extensive list of potential variables, this approach would be too costly and unwieldy in the sense that it would yield so many equations, the theoretically important ones would be lost.
Table 11.6: Regression model selection procedures

--------------------------
Forward, Stepwise, Backward Solutions

1. Deal with one variable at a time
2. Turns everything over to the machine they do not allow theory to have an impact.
3. These will not always yield the same answer with the same set of variables, that is, they do not necessarily yield unique solutions.
4. May result in equations that are hard to justify substantively.

--------------------------
Forward Solution

Begins by selecting, from the set of variables, the independent variable with the highest zero-order correlation with the dependent variable. Succeeding variables are chosen on the basis of the highest squared semipartial correlation with the dependent variable. In other words, it is the variable with the highest correlation to the dependent variable after the effects of earlier entered variables have been partialled out.

ADVAN: More economical of computer facilities than the all possible regression procedure or the Backward Procedure, particularly in low B and C frameworks. Avoids working with more X's than are necessary while "improving" the equation at each stage.

DISAD: Fails to explore the effect that the introduction of a new variable may have on the role played by a variable which entered at an earlier stage. The ultimate effect is that one may wind up with a model which contains insignificant variables.

--------------------------
Stepwise Solution

The process of selection is the same as in the forward solution. However, the stepwise solution contains an additional operation. After each new variable is selected, the contribution of each variable in the equation is re-examined. Thus the implications that new variables have on those already in the equation are considered. In other words, due to the combined contribution of variables, a variable which may have been the best single predictor at an earlier stage may become superfluous when other variable(s) are added at later stage(s).

ADVAN: 1. End up with all significant values
2. Given that set of variables, this is probably the largest $R^2$ value you can get, although not necessarily, since you do exclude certain considerations by the partialing.

DISAD: 1. May give solutions that are awfully hard to justify substantively.
2. Because of the nature of the intercorrelations of your variables you may eliminate theoretically important variables while retaining other less relevant variables.

Backward Solution:

Begins with the squared multiple correlation ($R^2$) of all independent variables with the dependent variable. Then each variable is examined as if it were the last variable to have entered the equation. The variable adding the least increment is eliminated if the loss in $R^2$ is insignificant.

ADVAN: Enables the researcher to examine all his variables at once so as not to miss anything. Really saying you are in Situation B, where at least part of the variables are known.

DISAD: More costly than forward and stepwise, especially when one has a large number of variables, as found in a C type situation.
As a means of illustrating these methods, we have included a number of sample runs. Some of these, Tables 11.7 through 11.9, are based on the total sample (N=240). The last three, Tables 11.10 through 11.11 are based on a subsample, (N=36). While an N=36 is generally considered quite small for multiple regression, the purpose here is illustration not inference, and this purpose is often better served by a smaller sample.

If we examine the regressions based on the total sample (Tables 11.7 through 11.9) we find that the backward, forward, and stepwise procedures all yielded a 7-variable model. (Each procedure began with an eight variable model.) This means that of the eight variables included in the sample data set, only Tension failed to reach a significance level of .5 when role performance was regressed on the eight-variable set. Note that regardless of the technique utilized, the information in the ANOVA table remains the same.

\[ R^2 = 0.536 \quad F = 38.332 \]

significant at the .0001 level
Table 11.7: Backward elimination procedure, \((N = 240)\)

<table>
<thead>
<tr>
<th>NUMBER IN</th>
<th>R-SQUARE</th>
<th>VARIABLES IN MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.53673297</td>
<td>scot, cont, sel, scot, per, salt, ten, just</td>
</tr>
<tr>
<td>7</td>
<td>0.53630314</td>
<td>scot, cont, sel, scot, per, salt, just</td>
</tr>
</tbody>
</table>

The variables in the above model have all been deemed significant at the 0.1000 significance level.

Analysis of variance table, regression coefficients, and statistics of fit for the above model:

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>R-SQUARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>7</td>
<td>26548782.50613443</td>
<td>3792683.21516206</td>
<td>38.33241</td>
<td>0.0001</td>
<td>0.53630314</td>
</tr>
<tr>
<td>ERROR</td>
<td>232</td>
<td>22956531.28969890</td>
<td>96941.94521422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORRECTED TOTAL</td>
<td>239</td>
<td>49503313.79583333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SEQUENTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>PARTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>scot</td>
<td>1</td>
<td>2010580.16880628</td>
<td>203.15530</td>
<td>0.0001</td>
<td>1953510.53203388</td>
<td>19.74401</td>
<td>0.0001</td>
</tr>
<tr>
<td>per</td>
<td>1</td>
<td>2772095.70426924</td>
<td>27.41098</td>
<td>0.0001</td>
<td>2040720.47432656</td>
<td>20.70629</td>
<td>0.0001</td>
</tr>
<tr>
<td>cont</td>
<td>1</td>
<td>1632101.39623672</td>
<td>16.49635</td>
<td>0.0002</td>
<td>1196362.25582189</td>
<td>12.09156</td>
<td>0.0009</td>
</tr>
<tr>
<td>scot</td>
<td>1</td>
<td>838377.53623305</td>
<td>8.47343</td>
<td>0.0043</td>
<td>451070.1165564</td>
<td>4.55896</td>
<td>0.0317</td>
</tr>
<tr>
<td>sel</td>
<td>1</td>
<td>475509.40184769</td>
<td>4.00594</td>
<td>0.0275</td>
<td>414705.0185391</td>
<td>4.19140</td>
<td>0.0392</td>
</tr>
<tr>
<td>just</td>
<td>1</td>
<td>426400.21562385</td>
<td>4.28170</td>
<td>0.0372</td>
<td>426174.56792318</td>
<td>4.30737</td>
<td>0.0366</td>
</tr>
<tr>
<td>sal</td>
<td>1</td>
<td>366398.79311780</td>
<td>3.70316</td>
<td>0.0524</td>
<td>366398.08311780</td>
<td>3.70316</td>
<td>0.0724</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>B VALUES</th>
<th>T FOR H0:B=0</th>
<th>PROB &gt;</th>
<th>STD ERR B</th>
<th>STD B VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>-733.11384067</td>
<td>6.273465690</td>
<td>0.0001</td>
<td>6.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>scot</td>
<td>27.67562253</td>
<td>4.44342</td>
<td>0.0001</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>per</td>
<td>27.34124336</td>
<td>4.45642</td>
<td>0.0001</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>cont</td>
<td>3.20540118</td>
<td>3.47729</td>
<td>0.0009</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>scot</td>
<td>10.49773328</td>
<td>2.13517</td>
<td>0.0317</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>sel</td>
<td>35.79073238</td>
<td>2.04729</td>
<td>0.0392</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>just</td>
<td>1.63895515</td>
<td>2.07541</td>
<td>0.0366</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
<tr>
<td>salt</td>
<td>1.51183553</td>
<td>1.92436</td>
<td>0.0524</td>
<td>4.000017417</td>
<td>0.02853579</td>
</tr>
</tbody>
</table>
Table 11.8: Forward selection procedure, \((N = 240)\)

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>R-SQUARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>7</td>
<td>26548782.50613443</td>
<td>3792683.21516206</td>
<td>38.33241</td>
<td>0.0001</td>
<td>0.53630314</td>
</tr>
<tr>
<td>ERROR</td>
<td>232</td>
<td>22954531.28969890</td>
<td>98941.94521422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORRECTED TOTAL</td>
<td>239</td>
<td>49503333.79583333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SEQUENTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>PARTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOT</td>
<td>1</td>
<td>20100580.16880628</td>
<td>203.15530</td>
<td>0.0001</td>
<td>1953510.53203388</td>
<td>19.74401</td>
<td>0.0001</td>
</tr>
<tr>
<td>PERN1</td>
<td>1</td>
<td>2712095.70426934</td>
<td>27.41098</td>
<td>0.0001</td>
<td>2048720.47432656</td>
<td>20.70629</td>
<td>0.0001</td>
</tr>
<tr>
<td>CM1</td>
<td>1</td>
<td>1632181.39623672</td>
<td>16.48635</td>
<td>0.0002</td>
<td>1196362.25582189</td>
<td>12.09156</td>
<td>0.0009</td>
</tr>
<tr>
<td>SD1</td>
<td>1</td>
<td>838377.53623305</td>
<td>8.47343</td>
<td>0.0043</td>
<td>451070.11665544</td>
<td>4.55894</td>
<td>0.0317</td>
</tr>
<tr>
<td>SELT</td>
<td>1</td>
<td>475509.40184769</td>
<td>4.80594</td>
<td>0.0275</td>
<td>414705.01155891</td>
<td>4.19140</td>
<td>0.0392</td>
</tr>
<tr>
<td>JBST</td>
<td>1</td>
<td>423640.21562385</td>
<td>4.28170</td>
<td>0.0372</td>
<td>426174.56792318</td>
<td>4.30732</td>
<td>0.0366</td>
</tr>
<tr>
<td>SALT</td>
<td>1</td>
<td>366398.08311780</td>
<td>3.70316</td>
<td>0.0524</td>
<td>366398.08311780</td>
<td>3.70316</td>
<td>0.0524</td>
</tr>
<tr>
<td>SOURCE</td>
<td>B VALUES</td>
<td>T FOR HO: B = 0</td>
<td>PROB &gt;</td>
<td>T</td>
<td></td>
<td>STD ERR B</td>
<td>STD B VALUES</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>----------------</td>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>MEAN</td>
<td>-733.13384067</td>
<td>4.44342</td>
<td>0.0001</td>
<td>6.27345690</td>
<td>0.28653579</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCOT</td>
<td>27.07562253</td>
<td>4.55042</td>
<td>0.0001</td>
<td>6.00851471</td>
<td>0.2225327</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERVT</td>
<td>27.34124336</td>
<td>3.47729</td>
<td>0.0009</td>
<td>0.92183268</td>
<td>0.1902932</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMT</td>
<td>10.49773328</td>
<td>2.13517</td>
<td>0.0317</td>
<td>4.91658700</td>
<td>0.1186767</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCT</td>
<td>35.79073238</td>
<td>2.04729</td>
<td>0.0392</td>
<td>17.48200201</td>
<td>0.09796565</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELT</td>
<td>1.63895515</td>
<td>2.07541</td>
<td>0.0366</td>
<td>0.78970253</td>
<td>0.10310258</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JBST</td>
<td>1.51183503</td>
<td>1.92436</td>
<td>0.0524</td>
<td>0.78562995</td>
<td>0.09363192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SALT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NUMBER IN MODEL</th>
<th>R-SQUARE</th>
<th>VARIABLES IN MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.40604514</td>
<td>SCOT</td>
</tr>
<tr>
<td>2</td>
<td>0.46083129</td>
<td>SCOT PERVT</td>
</tr>
<tr>
<td>3</td>
<td>0.49380244</td>
<td>COMT SCOT PERVT</td>
</tr>
<tr>
<td>4</td>
<td>0.51073023</td>
<td>SOCT COMT SCOT PERVT</td>
</tr>
<tr>
<td>5</td>
<td>0.52035384</td>
<td>SOCT COMT SELT SCOT PERVT</td>
</tr>
<tr>
<td>6</td>
<td>0.52890165</td>
<td>SOCT COMT SELT SCOT PERVT JBST</td>
</tr>
<tr>
<td>7</td>
<td>0.53633314</td>
<td>SOCT COMT SELT SCOT PERVT SALT JBST</td>
</tr>
</tbody>
</table>
Table 11.9: Stepwise regression procedure, (N = 240)

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>R-SQUARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>7</td>
<td>26548782.50613443</td>
<td>3792683.21516266</td>
<td>38.33241</td>
<td>0.0001</td>
<td>0.53630314</td>
</tr>
<tr>
<td>ERROR</td>
<td>232</td>
<td>22554531.28969890</td>
<td>98941.94521422</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORRECTED TOTAL</td>
<td>239</td>
<td>49503313.79583333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SEQUENTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>PARTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCOT</td>
<td>1</td>
<td>20100580.16880628</td>
<td>203.15530</td>
<td>0.0001</td>
<td>1953510.53203386</td>
<td>19.74601</td>
<td>0.0001</td>
</tr>
<tr>
<td>PERT</td>
<td>1</td>
<td>2712095.70426904</td>
<td>27.41098</td>
<td>0.0001</td>
<td>2048720.47432656</td>
<td>20.70629</td>
<td>0.0001</td>
</tr>
<tr>
<td>CMHT</td>
<td>1</td>
<td>1632181.39623672</td>
<td>16.49635</td>
<td>0.0002</td>
<td>1196362.29582139</td>
<td>12.09156</td>
<td>0.0009</td>
</tr>
<tr>
<td>SUCI</td>
<td>1</td>
<td>803877.53623305</td>
<td>8.47363</td>
<td>0.0043</td>
<td>451070.11665542</td>
<td>4.558894</td>
<td>0.0317</td>
</tr>
<tr>
<td>SELT</td>
<td>1</td>
<td>475509.40184769</td>
<td>4.380594</td>
<td>0.0275</td>
<td>414705.01153891</td>
<td>4.19140</td>
<td>0.0392</td>
</tr>
<tr>
<td>JST</td>
<td>1</td>
<td>423640.21562385</td>
<td>4.28170</td>
<td>0.0372</td>
<td>426174.56792346</td>
<td>4.30732</td>
<td>0.0366</td>
</tr>
<tr>
<td>SALT</td>
<td>1</td>
<td>366390.08311780</td>
<td>3.70316</td>
<td>0.0524</td>
<td>366390.08311780</td>
<td>3.70316</td>
<td>0.0524</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>B VALUES</th>
<th>T FOR HO = 0</th>
<th>PROB &gt;</th>
<th>STD ERR</th>
<th>STD B VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN</td>
<td>-733.11386067</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCOT</td>
<td>27.07562253</td>
<td>4.44342</td>
<td>0.0001</td>
<td>6.27365690</td>
<td>0.28653579</td>
</tr>
<tr>
<td>PERT</td>
<td>27.34124336</td>
<td>4.55042</td>
<td>0.0001</td>
<td>6.00851471</td>
<td>0.22250107</td>
</tr>
<tr>
<td>CMHT</td>
<td>3.20548118</td>
<td>3.47729</td>
<td>0.0009</td>
<td>0.92183268</td>
<td>0.19029321</td>
</tr>
<tr>
<td>SUCI</td>
<td>10.49773328</td>
<td>2.13517</td>
<td>0.0317</td>
<td>4.91658770</td>
<td>0.11883673</td>
</tr>
<tr>
<td>SELT</td>
<td>35.79073238</td>
<td>2.04729</td>
<td>0.0392</td>
<td>17.448200201</td>
<td>0.09794565</td>
</tr>
<tr>
<td>JST</td>
<td>1.63895515</td>
<td>2.07541</td>
<td>0.0366</td>
<td>0.787970253</td>
<td>0.10310258</td>
</tr>
<tr>
<td>SALT</td>
<td>1.51183503</td>
<td>1.92436</td>
<td>0.0524</td>
<td>0.78562995</td>
<td>0.09363192</td>
</tr>
</tbody>
</table>
The subsample regressions began with an initial set of six variables:

- Socialization
- Communication
- Scope
- Pervasiveness
- Job Satisfaction
- Salience

Using these six variables we ran three types of regressions: a standard regression of role performance on the entire set of six variables; and the forward and stepwise procedures which reduced the total list to three and two variable models (respectively).

The standard regression \((N=36)\) of role performance on the set of six variables above, yielded an \(R^2 = 0.471\). If we examine the partial SS, we see that only pervasiveness was significant at the .05 level when entered last. This means that only pervasiveness made a unique contribution to the regression when added last (at the .05 level). Although the models are not the same so that a one-to-one comparison is impossible, it is interesting to note that only salience, in the 7-variable models of the \(N=240\) runs, failed to reach a significance level of .05 when entered last. When we look at the sequential SS \((N=36)\) we see that only the additions of socialization, communication and pervasiveness were significant. The amount of variation added by scope, job satisfaction and salience to the variation already accounted for was minimal. For example, the operations below, adapted from Table 11.10 illustrate that the addition of scope to
Table 11.10: Analysis of variance, regression coefficients and statistics of fit, (N=36)

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SUM OF SQUARES</th>
<th>MEAN SQUARE</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>R-SQUARE</th>
<th>C.V.</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>6</td>
<td>3160955.17050393</td>
<td>526325.86175066</td>
<td>4.3039</td>
<td>0.0035</td>
<td>0.47098191</td>
<td>36.286175</td>
</tr>
<tr>
<td>ERROR</td>
<td>29</td>
<td>3550460.05171829</td>
<td>122429.65695580</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CORRECTED TOTAL</td>
<td>35</td>
<td>6711415.22222222</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>DF</th>
<th>SEQUENTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
<th>PARTIAL SS</th>
<th>F VALUE</th>
<th>PROB &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOCT</td>
<td>1</td>
<td>697534.29099013</td>
<td>5.6343</td>
<td>0.69743</td>
<td>81810.38028049</td>
<td>0.66822</td>
<td>0.5144</td>
</tr>
<tr>
<td>CONT</td>
<td>1</td>
<td>945521.46687347</td>
<td>7.72298</td>
<td>0.0092</td>
<td>284641.98443099</td>
<td>2.32494</td>
<td>0.1246</td>
</tr>
<tr>
<td>SCOT</td>
<td>1</td>
<td>206703.30969150</td>
<td>1.65834</td>
<td>0.2015</td>
<td>31851.25193329</td>
<td>0.26016</td>
<td>0.6196</td>
</tr>
<tr>
<td>JBST</td>
<td>1</td>
<td>8712.95915432</td>
<td>0.07117</td>
<td>0.7874</td>
<td>2059.38347455</td>
<td>0.01682</td>
<td>0.8930</td>
</tr>
<tr>
<td>SALT</td>
<td>1</td>
<td>9042.75591319</td>
<td>0.07386</td>
<td>0.7838</td>
<td>9042.75591319</td>
<td>0.07386</td>
<td>0.7838</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>B VALUES</th>
<th>T FOR H0:B=0</th>
<th>PROB &gt;</th>
<th>STD ERR B</th>
<th>STD B VALUES</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERCEPT</td>
<td>-512.63048845</td>
<td>-0.69982</td>
<td>0.5037</td>
<td>732.51362245</td>
<td>0.0</td>
</tr>
<tr>
<td>SOCT</td>
<td>15.02668354</td>
<td>0.81745</td>
<td>0.5744</td>
<td>18.38240164</td>
<td>0.15324411</td>
</tr>
<tr>
<td>CONT</td>
<td>4.42457399</td>
<td>1.52478</td>
<td>0.1346</td>
<td>2.90178545</td>
<td>0.3003205</td>
</tr>
<tr>
<td>SCOT</td>
<td>11.20214964</td>
<td>0.51006</td>
<td>0.0133</td>
<td>18.54441578</td>
<td>0.43410905</td>
</tr>
<tr>
<td>JBST</td>
<td>0.45315478</td>
<td>0.12970</td>
<td>0.8930</td>
<td>3.49398686</td>
<td>0.02822591</td>
</tr>
<tr>
<td>SALT</td>
<td>0.86392342</td>
<td>0.27177</td>
<td>0.7838</td>
<td>3.17883627</td>
<td>0.04784562</td>
</tr>
</tbody>
</table>
socialization and communication did not add significantly to the regression, whereas the addition of a fourth variable, pervasiveness, did. On the basis of the sequential SS, then, we might decide to reduce the variables in the model to socialization, communication and pervasiveness, (the model that was selected by the forward procedure). On the other hand, using the partial SS as a guide, we would probably select the two-variable model chosen by the stepwise procedure (pervasiveness and communication). An examination of the B's (partial regression coefficients) suggests perhaps a third possibility - a one variable model (pervasiveness) - since it was the only "B" that was significant at the .05 level. This demonstrates that it is often possible to arrive at the same model, using standard regression, as one would get using a selection procedure, although, not as quickly.
Table 11.11: Forward selection procedure, (N = 36)^

<table>
<thead>
<tr>
<th>STEP 1</th>
<th>VARIABLE</th>
<th>PENTRY</th>
<th>R SQUARE = 0.4314241 cin</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>1</td>
<td>153150.0,2441426</td>
<td>9.07</td>
</tr>
<tr>
<td>CPT</td>
<td>35</td>
<td>5.100223,0150850</td>
<td>152442.01910483</td>
</tr>
<tr>
<td>TOTAL</td>
<td>36</td>
<td>5.111452,00222222</td>
<td>4.97</td>
</tr>
<tr>
<td>B VALUE</td>
<td>STD ERROR</td>
<td>TYPE II SS</td>
<td>F</td>
</tr>
<tr>
<td>Intercept</td>
<td>143.12105633</td>
<td>16.36366286</td>
<td>1.321586.0461526</td>
</tr>
<tr>
<td>PENTRY</td>
<td>23.724279666</td>
<td>1.9321586.0461526</td>
<td>0.29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEP 2</th>
<th>VARIABLE</th>
<th>CONT</th>
<th>ENTRLD</th>
<th>R SQUARE = 0.4314190 cin</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>2</td>
<td>252105,099759</td>
<td>1467542.35493976</td>
<td>12.72</td>
</tr>
<tr>
<td>CPT</td>
<td>35</td>
<td>5.790777.1131377</td>
<td>114850.08828221</td>
<td>0.000</td>
</tr>
<tr>
<td>TOTAL</td>
<td>37</td>
<td>5.111452,00222222</td>
<td>4.97</td>
<td>0.2333</td>
</tr>
<tr>
<td>B VALUE</td>
<td>STD ERROR</td>
<td>TYPE II SS</td>
<td>F</td>
<td>PROOF</td>
</tr>
<tr>
<td>Intercept</td>
<td>153.22833434</td>
<td>1.3221056332</td>
<td>1.321586.0461526</td>
<td>12.26</td>
</tr>
<tr>
<td>PENTRY</td>
<td>51.51375001</td>
<td>14.36764186</td>
<td>1.321586.0461526</td>
<td>12.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEP 3</th>
<th>VARIABLE</th>
<th>SOCT</th>
<th>ENTERED</th>
<th>R SQUARE = 0.4314315 cin</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>3</td>
<td>5.09233,1767779</td>
<td>1.321586.0461526</td>
<td>9.06</td>
</tr>
<tr>
<td>CPT</td>
<td>32</td>
<td>5.628591.10444515</td>
<td>113393.15507764</td>
<td>0.000</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>5.111452,00222222</td>
<td>4.97</td>
<td>0.2333</td>
</tr>
<tr>
<td>B VALUE</td>
<td>STD ERROR</td>
<td>TYPE II SS</td>
<td>F</td>
<td>PROOF</td>
</tr>
<tr>
<td>Intercept</td>
<td>15.790777.1131377</td>
<td>0.0092105633</td>
<td>1.321586.0461526</td>
<td>1.43</td>
</tr>
<tr>
<td>SOCT</td>
<td>5.74589150</td>
<td>2.09012092</td>
<td>1.321586.0461526</td>
<td>7.56</td>
</tr>
<tr>
<td>PENTRY</td>
<td>51.51375001</td>
<td>14.36764186</td>
<td>1.321586.0461526</td>
<td>12.39</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEP 4</th>
<th>VARIABLE</th>
<th>SOCT</th>
<th>ENTERED</th>
<th>R SQUARE = 0.4633631 cin</th>
</tr>
</thead>
<tbody>
<tr>
<td>REGRESSION</td>
<td>4</td>
<td>5.143195,45563641</td>
<td>7.05799.03805910</td>
<td>6.83</td>
</tr>
<tr>
<td>ERROR</td>
<td>31</td>
<td>5.600215.76678582</td>
<td>1.321586.0461526</td>
<td>0.000</td>
</tr>
<tr>
<td>TOTAL</td>
<td>35</td>
<td>5.111452,00222222</td>
<td>4.97</td>
<td>0.2333</td>
</tr>
<tr>
<td>B VALUE</td>
<td>STD ERROR</td>
<td>TYPE II SS</td>
<td>F</td>
<td>PROOF</td>
</tr>
<tr>
<td>Intercept</td>
<td>15.790777.1131377</td>
<td>0.0092105633</td>
<td>1.321586.0461526</td>
<td>1.43</td>
</tr>
<tr>
<td>SOCT</td>
<td>5.74589150</td>
<td>2.09012092</td>
<td>1.321586.0461526</td>
<td>7.56</td>
</tr>
<tr>
<td>PENTRY</td>
<td>51.51375001</td>
<td>14.36764186</td>
<td>1.321586.0461526</td>
<td>12.39</td>
</tr>
</tbody>
</table>

*Other variables met the 0.0001 significance level for entry into the model*
Table 11.12: Stepwise regression procedure, \((N = 36)\)\(^a\)

<table>
<thead>
<tr>
<th>Step 1 Variable Entered</th>
<th>R Square</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residual</td>
<td></td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 2 Variable Entered</th>
<th>R Square</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
<tr>
<td>PCRT</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 3 Variable Entered</th>
<th>R Square</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
<tr>
<td>PCRT</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Step 4 Variable Entered</th>
<th>R Square</th>
<th>Sum of Squares</th>
<th>Mean Square</th>
<th>F</th>
<th>P</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
<tr>
<td>PCRT</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>1521590.3441625</td>
<td>2.97</td>
<td>0.0031</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)And other variables met the \(p \leq 0.0000\) significance level for entry into the model.
It might require several runs to arrive at a model which can be selected more directly by a selection procedure.

Running a **forward solution** (with a significance level of .5) yielded a three-variable model containing socialization, communication and pervasiveness with an \( R^2 = 0.459 \). The **stepwise solution** (with a significance level of .1) rejected the three-variable model above and selected a model containing only communication and pervasiveness (\( R^2 = 0.435 \)), since socialization was not significant at the .1 level. An examination of the Sequential SS in both procedures explains why the stepwise selected the two variable model. The amount of variation, which was added by socialization after pervasiveness and communication were already in the model, was not significant. However, according to the partial SS, in both techniques, pervasiveness and communication would have been selected, even if they had entered last.

As stated above, a predetermined significance level on regression selection procedures determines the model that is selected, (in the sense that it limits the selection or addition of variables to only those which meet the specified significance level). The levels used in the forward and stepwise procedures for \( N=36 \), were .5 and .1, respectively. These are somewhat lower than the levels which are generally used for most statistics. However, selection procedures are usually utilized in a low B or C framework where the
researcher often has little more than a list of potential variables. Furthermore, sampling and measurement errors, as well as the order of introduction of the variables, affect the outcome of a regression analysis. It's possible then that theoretically important variables might be overlooked when a conservative (.05 or .01) significance level is used. A lower significance level in the model development stage allows for more flexibility and lessens the likelihood that these outside factors will obscure the consideration of important variables, (Bancroft, 1968). This can be important from either a theoretical or an applied sense.

Theoretically speaking, a sound model may not reach conservative limits in the initial sample, but subsequently prove adequate in more representative samples. Likewise, in terms of application, the most manipulable variables are not always the most statistically significant. For instance, increasing initial socialization efforts for groups of new coordinators might prove to be the most economical approach for improving overall role performance rather than attempting to change established communication and pervasiveness patterns after new coordinators have been dispersed in the field. Using this subsample and a conservative significance level, socialization would not have been included in the model.
In the present example the difference between the forward and stepwise solutions was caused by a difference in the required significance level. Thus the more conservative significance level for the stepwise solution (.1 versus the .5 of the forward solution) eliminated socialization, yielding a two-variable model rather than the three-variable model chosen in the forward solution.

This does not mean, however, that these procedures will always yield the same models, given the same required significance level. On the contrary, with another set of variables (and/or sample) the solutions could still vary because the forward solution, unlike the stepwise, fails to reassess the significance of variables after they have once entered the equation. Thus the combined contribution of variables which enter at a later stage may reduce the usefulness (significance) of variables already in the equation. If this occurs, the final model selected by the forward procedure will contain insignificant variables which would not be contained in the model selected by the stepwise procedure.

It should perhaps be emphasized that we have been discussing the use of regression selection procedures in terms of low B or C situations. These constitute exploratory research situations where theoretical guidelines are limited or unavailable. Some would argue that selection procedures
should never be used, that they just throw a bunch of variables together which may or may not have any substantive meaning or interpretation. Certainly theory, rather than the computer, should determine the selection of variables. It is for this reason, therefore, that we have placed the discussion of these selection techniques within the framework of exploratory situations, where theoretical development is vague or nonexistent. Even within exploratory situations, however, one must be aware of the "atheoretical" nature of these techniques.

We can illustrate this aspect of the selection procedures by examining the regressions for the total sample. The two variables which consistently enter first are scope and pervasiveness. Furthermore, a study by Warren, et al., (1975), using the same total sample data set, found that scope and pervasiveness were the only significant variables affecting role performance when corrections for measurement error were taken into consideration. Let us assume for the moment that this represents the "true" situation in the population. To what extent, then will the selection procedures reflect this relationship when used on the subsamples?

At the .1 level of significance, none of the models selected for the sub-samples contained both scope and pervasiveness. Furthermore, a different model was selected
Table 11.13: Models selected for the $N=36$ subsamples

<table>
<thead>
<tr>
<th>Subsample #</th>
<th>Selection Procedure</th>
<th>107</th>
<th>108</th>
<th>104</th>
</tr>
</thead>
<tbody>
<tr>
<td>107</td>
<td>Forward (0.5 sig. level)</td>
<td>Scope</td>
<td>*Pervasiveness</td>
<td>*Communication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>*Job Satisfaction</td>
<td>*Communication</td>
<td>*Pervasiveness</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pervasiveness</td>
<td>Socialization</td>
<td>*Job Satisfaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Communication</td>
<td>Scope</td>
<td>Scope</td>
</tr>
<tr>
<td>108</td>
<td>Stepwise (0.1 sig. level)</td>
<td>Scope</td>
<td>Pervasiveness</td>
<td>Communication</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Job Satisfaction</td>
<td>Communication</td>
<td>Pervasiveness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Job Satisfaction</td>
</tr>
</tbody>
</table>

*Would have been selected if the significance level had been 0.1.
for each of the subsamples. In the #108 subsample, scope was not included in the model at the .1 level of selection when the set of six variables were examined. However, when a standard regression of role performance on scope and pervasiveness was run for this subsample, both proved significant at the .01 level, regardless of the order of entrance. Clearly, then, the regression selection techniques eliminated an important variable (scope) in terms of the "true" relationship in the population (assuming, for purposes of illustration, that the total sample represents the population) and also in terms of the subsample itself, since pervasiveness and scope were significant when these two variables were considered alone. In other words, the "true" relationship within the population, would have been revealed in the sub-sample analysis, if theoretical guidelines had specified the model, but not if the researcher had simply allowed the computer to select a model from the set of six variables.

Regression selection procedures concentrate on maximizing explained variance ($R^2$). Thus the emphasis is on prediction (i.e., how well did the model fit this set of data) since the larger the $R^2$ the better the "fit" or prediction. But this is prediction in terms of a particular set of data and not scientific prediction. (Scientific prediction would attempt to predict scores for a new set of
people from a population and would therefore be based on theory.) Like scientific prediction, explanation is concerned with relationships within the population, not just a particular sample. The utility of selection techniques, therefore, is confined to predictive or model building goals rather than explanation or model testing.

Searching for an Explanation

There are a number of criteria that can be examined when explanation is the goal, some of which we have discussed earlier in the chapter. We can think of these criteria in terms of a series of questions, the answers to which provide different types of information which together contribute to our understanding of the substantive relations among the variables. These include the following:

1. Test of $R^2$
   (Is the joint contribution of $X$ and $X$ significant)
   \[ 1 \quad 2 \]

2. $R^2$ shrinkage
   (What is the extent to which the sample $R^2$ has been overestimated?)

3. $MS_{resid}$
   (What is the most efficient model i.e., the model where $MS_{resid}$ is at a minimum?)

4. Significance of the partial regression coefficient (Is the contribution of $X$ significant after $X$ has been taken into account, and vice versa?)
5. Commonality Analysis (What is the unique contribution of each variable?)

6. Partial and Semipartial Correlations (What is the incremental contribution of each variable according to the order of inclusion?)

As a means of facilitating the discussion, we shall examine the questions as they relate to the following regression of role performance on scope and pervasiveness, (Table 11.14).

**Test of R-Square**

In order to answer the first question, i.e., "Is the joint contribution of $X_1$ and $X_2$ significant?", we must test the statistical significance of $R^2$, and thus the regression. This can be done several ways.

\[
F = \frac{R^2/k}{(1-R^2)/(N-k-1)}
\]

\[
F = \frac{46083/2}{(1-.46083)/(240-2-1)}
\]

\[
F = 101.52
\]

\[
F = 3.05 \text{ at } .05 \text{ level}
\]

\[
F = 101.283
\]
Table 11.14: Regression of Role Performance on Scope and Pervasiveness

\[
\hat{Y} = 10.50940 + 52.98685(X_1) + 30.92942(X_2)
\]

\[R^2 = .46083\]

<table>
<thead>
<tr>
<th>ANOV</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regr</td>
<td>2</td>
<td>22812675.873</td>
<td>11406337.936</td>
<td>101.283</td>
</tr>
<tr>
<td>Resi</td>
<td>237</td>
<td>26690637.923</td>
<td>112618.725</td>
<td></td>
</tr>
</tbody>
</table>
The results of the significance tests indicate that the $R^2$ is significant and thus one can conclude that the joint contribution of $X_1$ and $X_2$ is significant.

**$R^2$-Square Shrinkage**

Explained variance or $R^2$ is data bound, (i.e., specific to a particular sample) and is therefore generally discussed in a predictive framework. However, attempts have been made to estimate the $R^2$ in the population (Nie, et al., 1975) by adjusting the sample $R^2$ for the number of independent variables and the size of the sample. In calculating the sample $R^2$, the zero order correlations, used to compute the weights necessary for a maximum $R$, are treated as if they were error-free. Since they are never error-free, an element of chance is introduced, biasing the computed $R$ upwards (Kerlinger and Pedhazur, 1973). The amount of overestimation of $R$ cannot be determined precisely, however, there are several alternatives for estimating the amount of shrinkage, all of which represent a more conservative, but nevertheless biased, estimate of the percent of variance explained in the population. Both procedures are based on the following logic:

$$R^2 \text{ in the population} =$$

$$1 - \frac{\text{error variance in } Y \text{ in the population}}{\text{total variance in } Y \text{ in the population}}$$

Thus we have
\[ \hat{R}^2 = 1 - (1-R^2) \left( \frac{N-1}{N-k-1} \right) \]

and

\[ \hat{R}^2 = R^2 - \left[ \sum_{k=1}^{N-k-1} (1-R^2) \right] \]

as taken from Kerlinger and Pedhazur (1973:283) and Nie, et al., (1975:358) respectively. Applying these two formulas to the example we have:

| \( \hat{R}^2 = 1 - \left( 1.46083 \right) \left( \frac{239}{237} \right) \) | \( \hat{R}^2 = .46083 - \left[ \frac{1}{238} \right] \) |
| \( \hat{R}^2 = 1 - .54371 \) | \( \hat{R}^2 = .46083 - .00226 \) |
| \( \hat{R}^2 = .45629 \) | \( \hat{R}^2 = .45857 \) |

NOTE: The formula on the right is the one which is output by SPSS.

According to Kerlinger and Pedhazur, the overestimation of \( R \) gets larger as the ratio between the number of independent variables and sample size increases. Therefore, the larger the sample size, the less bias will be incurred.

"Some authors recommend that the ratio of independent variables to sample size be at least 30 subjects per independent variable. This is a rule of thumb that does not satisfy certain researchers who say that samples should have at least 400 subjects" (Kerlinger and Pedhazur, 1973:282).
Minimizing mean square residual

Explained variance ($R^2$) gives the researcher an estimate of the relevance of his (her) model. In general a model with a relatively large $R^2$ probably contains more of the pertinent variables than does one with a low $R^2$. The problem with depending upon $R^2$ is that the addition of any variable, relevant or not, can raise the $R^2$. A better method of obtaining the most efficient model is to select the model in which the mean square residual ($MSresid$) reaches a minimum rather than attempting to maximize the $R^2$. This was discussed at some length earlier in this chapter. In essence, the $MSresid$ continues to decrease with each additional variable until one begins adding error. At that point the $MSresid$ will increase. Table 11.15 below illustrates how the addition of a theoretically unimportant variable can nevertheless increase the $R^2$, however, instead of decreasing the $MSresidual$, as was the case in Model 2, the $MSresidual$ increased, indicating that the addition of tension to the model actually resulted in fitting error rather than true variation.
Table 11.15: Comparing mean square residuals to explained variance

<table>
<thead>
<tr>
<th>Model #</th>
<th>Independent Variables</th>
<th>MS residual</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Scope</td>
<td>123540.898</td>
<td>.40605</td>
</tr>
<tr>
<td>2</td>
<td>Scope, Pervasiveness</td>
<td>112618.725</td>
<td>.46083</td>
</tr>
<tr>
<td>3</td>
<td>Scope, Pervasiveness, Tension</td>
<td>112975.683</td>
<td>.46140</td>
</tr>
</tbody>
</table>
Minimizing the standard error of the estimate

Another check, which is similar to the above, consists of examining the standard error of the estimate \( \sqrt{\frac{MS}{\text{resid}}} \). This represents an estimate of the dispersion or variance of the predicted \( \hat{Y} \) values about the regression. As such, it can be compared to the standard deviation of the \( Y \) scores \( (S = 455.112) \). In general, the smaller the standard error is, compared to the standard deviation of \( Y \), the better. In this case the standard error is considerably smaller, indicating that the prediction is fairly successful. The standard error will approach (or even exceed) the standard deviation of \( Y \) as the \( R^2 \), or explained variance, approaches 0. Conversely, then, the greater the precision of the model in predicting \( Y \), the smaller the standard error will be. This statistic can also be useful in deciding between models, i.e., does the introduction of \( X_4 \) add to the precision of the model containing \( X_1, X_2, \) and \( X_3 \)? [Note, however, that the standard error can approach zero simply by adding enough variables, therefore, care should also be taken when interpreting this statistic.] If we examine models 1 and 2, we see from table 11.16 that the \( R^2 \) for Model 2 is considerably less than the \( R^2 \) for models 1 or 3 and therefore its standard error is closer to the standard deviation of \( Y \) \( (S = 455.112) \). In the case of the fourth model, the \( R^2 \) is practically the same (although slightly larger, as shown in Table 11.16) as that
<table>
<thead>
<tr>
<th>Model #</th>
<th>Model</th>
<th>Standard Error of Estimates</th>
<th>F²</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Y on Scope</td>
<td>351.483</td>
<td>.406</td>
<td>36.5%</td>
</tr>
<tr>
<td>2</td>
<td>Y on Pervasiveness</td>
<td>406.821</td>
<td>.204</td>
<td>42.3%</td>
</tr>
<tr>
<td>3</td>
<td>Y on Scope, Pervasiveness</td>
<td>335.587</td>
<td>.461</td>
<td>34.9%</td>
</tr>
<tr>
<td>4</td>
<td>Y on Scope, Pervasiveness, Tension</td>
<td>336.118</td>
<td>.461</td>
<td>34.9%</td>
</tr>
</tbody>
</table>
in model 3. However, the standard error of model 4 is larger than that for model 3 indicating that the 2 variable model #3 is a more efficient or precise model than the three variable one (model 4).

Another method for examining the decrease in the standard error of the estimate, according to Draper and Smith (1966), is to consider it in relation to the mean response, \( \frac{S}{\bar{Y}} \). Thus we see in Table 11.16 that the standard error of the estimate of Model 3 has been reduced down to 34.9% of the mean role performance score (\( \bar{Y} = 962.404 \)). As with the two previous criteria, there is no set standard to which this measure can be compared, rather one must depend on "prior knowledge and personal feelings" to decide whether this represents a satisfactory level of precision.

The criteria to this point have been concerned with answering questions about the whole regression. The remainder will involve various methods or techniques for analyzing the individual contribution of variables or groups of variables.
Assessing the significance of partial regression coefficients

An examination of the partial regression coefficients can answer the question, "Is the contribution of $X_i$ significant after all other $X$'s have been taken into account?" A partial regression coefficient ($B_i$) represents the amount of change that can be expected in $Y$ for a given unit increase in one of independent variables ($X_i$) while controlling for or holding all of the others constant.

Another way of explaining the partial $B_i$ according to Nie, et al., (1975) is in terms of residuals.

"The partial $B_i$ is equivalent to a simple $B$ between $Y$ and the residuals of $X_i$ from which the effects of $X_2$ is taken out. That is, if $x_i = (X - X'_i)$ where $X'_i = A + BX_2$, then the partial $B_i$ is equivalent to the simple regression coefficient $B$ from the equation $Y = A + BX_i$" (page 330).

Given this explanation of the partial regression coefficient, it should be clear, then, that the test $H_0: B_i = 0$ represents one approach to the question in the previous paragraph. Putting the discussion in the context of Table 11.17, we see from the F tests that both $B$'s in Model One are significant ($H_0: B_i = 0$ was not upheld). Note, however, that the partial regression coefficient for Tension in Model Two was not found to be significant. The betas ($\beta$) or standardized partial regression coefficients give the relative influence of each variable, given this set of variables. (I will discuss this coefficient in greater depth
Table 11.17: Unstandardized and standardized partial regression coefficients

<table>
<thead>
<tr>
<th>Model #</th>
<th>Independent Variable</th>
<th>Partial Regression Coefficient (β)</th>
<th>Beta (β)</th>
<th>Std. Error of B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Scope</td>
<td>52.98685</td>
<td>0.545</td>
<td>4.990</td>
<td>112.763*</td>
</tr>
<tr>
<td></td>
<td>Pervasiveness</td>
<td>30.92942</td>
<td>0.252</td>
<td>6.302</td>
<td>24.082*</td>
</tr>
<tr>
<td>2</td>
<td>Scope</td>
<td>51.99901</td>
<td>0.534</td>
<td>5.372</td>
<td>93.683*</td>
</tr>
<tr>
<td></td>
<td>Pervasiveness</td>
<td>31.18715</td>
<td>0.254</td>
<td>6.334</td>
<td>24.247*</td>
</tr>
<tr>
<td></td>
<td>Tension</td>
<td>2.67588</td>
<td>0.026</td>
<td>5.339</td>
<td>0.251*</td>
</tr>
</tbody>
</table>

*Significant at the .01 level, where $F_{238} = 4.69$ and $F_{237} = 4.86$ respectively.
in the next chapter). At this point, I shall only comment on the fact that the $\beta$ for tension of .026 is negligible when compared to the Betas of .534 and .254 for Scope and Pervasiveness.

The partial regression coefficients listed in Table 11.17 represent estimates of the true values in the population. In order to relate these estimates to the population, one must construct confidence intervals for each estimate. Using the variables in Model 1 of Table 11.17, we get the following confidence intervals:

<table>
<thead>
<tr>
<th></th>
<th>$b \pm t(S_i) \text{ with } 1 \text{ and } (N-k-1) \text{ d.f.}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>$52.98685 \pm 1.96 (4.990)$</td>
</tr>
<tr>
<td>Pervasiveness</td>
<td>$30.92942 \pm 1.96 (6.302)$</td>
</tr>
</tbody>
</table>

It has been pointed out by Draper and Smith (1966:65) that this formula yields a 95% confidence interval (or range) for the partial regression coefficients when examined separately, irrespective of the value of any other parameters. In other words, the estimate does not provide for the joint confidence region. The calculation of joint confidence regions is somewhat involved and will not be discussed here.
Examine partial correlations

Another approach to examining the relation of one variable after controlling for all the others, involves the use of partial correlation. Somewhat similar to the explanation of the partial regression coefficient, the partial correlation has been defined as "the correlation between the residuals of the regressions of Y on X₂ and X₁ on X₂" (Blalock, 1960:435). Partial correlation removes the effect of a control variable(s) from both Y and X₁. What is being correlated then are errors with respect to the control variable. The formula for a first order partial consists of the following:

\[
r_{ij.k} = \frac{r_{ij} - (r_{ij} \cdot r_{ik} \cdot r_{jk})}{\sqrt{1 - r_{ij}^2} \sqrt{1 - r_{ik}^2} \sqrt{1 - r_{jk}^2}}
\]

If

\[i = \text{Role Performance}\]
\[j = \text{Scope}\]
\[k = \text{Pervasiveness}\]

\[r_{ij} = 0.63722\]
\[r_{ik} = 0.45199\]
\[r_{jk} = 0.36774\]

then

\[
r_{ij.k} = \frac{(0.63722) - (0.45199)(0.36774)}{\sqrt{1 - (0.45199)^2} \sqrt{1 - (0.36774)^2}} = \frac{0.7101}{0.8295} = 0.8578
\]
and

\[
\begin{align*}
    r_{ij.k} &= \frac{(.45199) - (.36774)(.63772)}{\sqrt{(1 - .63772)^2 \sqrt{1 - (.36774)^2}}} = \frac{.3808}{.8295} \\
    r_{ik.j} &= 0.304
\end{align*}
\]

The results would seem to indicate that most of the variation between pervasiveness and role performance is shared by scope. In terms of application, then, the researcher might recommend emphasizing only the scope of the organization (i.e., provide more opportunities for coordinators to interact with one another) in order to increase the coordinators' role performance.

**Commonality analysis**

Another approach to breaking down the variance is a method called Commonality Analysis. Commonality analysis is concerned with assessing the unique contribution of each variable as well as their various common contributions. According to Kerlinger and Pedhazur (1973:298):

"The unique contribution of an independent variable is defined as the variance attributed to it when it is entered last in the regression equation."

Applied to the model of role performance on scope and pervasiveness, we get the following components:
$U(1) = R^2_{y.12} - R^2_{y.2}$

where

$y = \text{role performance}$  
$R^2_{y.12} = \text{explained variance of } Y \text{ on } 1 \text{ and } 2$

$1 = \text{scope}$  
$R^2_{y.2} = \text{explained variance of } Y \text{ on } 2$

$2 = \text{pervasiveness}$  
$U(1) = \text{unique contribution of scope}$

$U(1) = .461 - .204$

$U(1) = .257$

Similarly $U(2) = R^2_{y.12} - R^2_{y.1}$

$U(2) = .461 - .406$

$U(2) = .055$

The commonality of the two independent variables can be expressed by the formula

$C(12) = R^2_{y.12} - U(1) - U(2)$

Thus $C(12) = .461 - .257 - .055$

$C(12) = .149$

We can now express the correlation between role performance and scope in terms of the common and unique contributions.

$R^2_{y.1} = U(1) + C(12)$

$(.637)^2 = .257 + .149$

$\sqrt{.406} = .406 \sqrt{ }$
Kerlinger and Pedhazur (1973) provide a more detailed discussion of commonality analysis, including expansion formulas for cases involving more than two independent variables. In elaborating the problems of commonality analysis, they seem to stress two points:

1. The uniqueness of variables depends on the relations among the specific set of variables under study. Addition or deletion of any variables causes a subsequent change in the uniqueness attributed to the variables.

2. Commonality analysis, at present, holds more utility for a predictive than an explanatory framework since the procedure enables one to select those variables which have large unique components and small commonalities.

It would seem that this approach does have more to offer theory building than theory testing, and yet it would also seem to be useful in understanding the interrelationships between variables. (One need not place the examination in terms of a lack of specificity of the indicators as Kerlinger and Pedhazur seem to imply.) The value of commonality analysis for explanatory purposes then, rests not on its theory testing capabilities but rather on the intuitive information it can supply concerning a set of theoretically defined variables. In terms of the example, this means that we allow theoretical considerations to determine our model, (Role Performance on Scope and Pervasiveness) and then utilize commonality analysis (in conjunction with other techniques discussed in this chapter) to help describe the
nature of the relationships which are established by our tests.

Kerlinger and Pedhazur describe the unique contribution formula as a squared semipartial correlation between $Y$ and the variable of interest, after partialing all of the other independent variables from it, (the variable of interest). The semipartial correlation (sometimes referred to as the part correlation) and its square can also provide important information concerning the interpretation of multiple regression analysis.

**Semi-partial correlations**

I noted earlier that the partial correlation ($r_{y1.2}$) represented the correlation between $Y$ and variable 1, after the effects of variable 2 had been removed from both $Y$ and variable 1. In other words, it was a correlation between two residuals. In the semipartial or part correlation, however, the effect of the control variable(s) is(are) removed from only one (generally the independent variable) of the two correlated variables. Thus the semipartial represents a correlation between $Y$ and the residual of $X_1$ after the effect of one or more independent variables has been removed. The formulas below show the relationship that exists between the semipartial and the partial correlation.
Semi-partial correlation

\[
\rho_{1.2} = \frac{\rho_{y_1 y_2} - \rho_{y_1 y_2}^2}{\sqrt{1 - \rho_{y_1 y_2}^2}}
\]

\[
\rho_{1.2} = \frac{.637 - (.452)(.368)}{\sqrt{1 - (.368)^2}} = .9295
\]

\[
\rho_{2.1} = \frac{.897}{\sqrt{1 - (.368)^2}} = .9295
\]

\[
\rho_{1.2} = .507
\]

\[
\rho_{2.1} = .234
\]

Partial correlation

\[
\rho_{1.2} = \frac{\rho_{y_1 y_2} - \rho_{y_1 y_2}^2}{\sqrt{1 - \rho_{y_1 y_2}^2}}
\]

\[
\rho_{1.2} = \frac{.568}{\sqrt{1 - (.368)^2}} = .9295
\]

\[
\rho_{2.1} = .304
\]
The semipartial correlation \([r_{y(1.2)}]\) of Role Performance and the residual of scope, after removing the effect of pervasiveness, is equal to .507. Another way of expressing this semipartial correlation is in terms of its square, i.e.

\[
\frac{r^2}{y(1.2)} = \frac{R^2}{y.12} - \frac{R^2}{y.2}
\]

where

\[
R^2_{y.12} = \text{explained variance of } Y \text{ on } 1 \text{ and } 2
\]

\[
R^2_{y.2} = \text{explained variance of } Y \text{ on } 2
\]

\[
r^2_{y(1.2)} = .461 - .204 = .257
\]

\[
r_{y(1.2)} = .257 = .507
\]

It follows, then that the \(F\)-test applied to differences between \(R^2\)'s constitutes a statistical test of significance for the semipartial correlation.

\[
\text{Significance test for Semipartial correlation of Role performance and scope:}
\]

\[
F = \frac{R^2_{y.12} - R^2_{y.2}}{(1-R^2_{y.12})/(240-2-1)}
\]

\[
F = \frac{.461 - .204}{(1-.461)/237} = .257 = 113.215
\]
Significance test for \( F = \frac{(R^2 - R^2_y)}{2 - 1} \)

where \( F = 6.76 \) is significant at the .01 level.

Thus both semipartial correlations \( r_{y(1.2)} \) and \( r_{y(2.1)} \) were statistically significant.

A further use of the semipartial correlation involves the decomposition of the explained variance (\( R^2 \)). In the case where all independent variables are uncorrelated, the \( R^2 \) can be expressed as the simple sum of squared correlations of each independent variable with the dependent variable, i.e.,

\[
r^2 = r^2_{y.12} + ... + r^2_{y.k}
\]

However, few sociologically relevant variables are truly "independent". Rather, there is a certain amount of overall or duplication of variance of each independent variable with the others. Using the squared semipartial correlation, the correlated independent variables can be adjusted to equal zero so that the relative contribution (for a particular order of introduction) of each independent variable can be calculated. Thus we have the following general formula expressing the \( R^2 \) of a model with four independent variables,
\[ R^2_{.1234} = r^2_{y1} + r^2_{y(2.1)} + r^2_{y(3.12)} + r^2_{y(4.123)} \]

[Note, however, that the relative contribution for a particular variable will change with a different order of introduction. Therefore, the semipartial correlation has its greatest utility where the order of introduction is specified by theory.]

Applying this to the model of Y on scope and pervasiveness, we can get the following two breakdowns, depending on which variable is introduced first.

<table>
<thead>
<tr>
<th>Order 1</th>
<th>Order 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ R^2_{.12} = r^2_{y1} + r^2_{y(2.1)} ]</td>
<td>[ R^2_{.12} = r^2_{y2} + r^2_{y(1.2)} ]</td>
</tr>
<tr>
<td>[ .461 = (.637)^2 + (.234)^2 ]</td>
<td>[ .461 = (.452)^2 + (.507)^2 ]</td>
</tr>
<tr>
<td>[ .461 = .406 + .055 ]</td>
<td>[ .461 = .204 + .257 ]</td>
</tr>
<tr>
<td>[ .461 = .461 \checkmark ]</td>
<td>[ .461 = .461 \checkmark ]</td>
</tr>
</tbody>
</table>

What this says is that most of the explained variance can be accounted for by scope, when scope is entered first.

However, when pervasiveness is entered first the relative
contribution of scope, with pervasiveness already entered is considerably less than when scope is entered first. This illustrates the importance of theoretically specifying the order of introduction, when one wants to use the semipartial correlation. The semipartial, and its square, contribute important information towards the understanding and explanation of the interrelationship between variables. In the next few paragraphs, we shall discuss how they can be extended to provide a measure of the relative influence of groups of independent variables.

Assessing the influence of groups of variables with the multiple partial correlation

The complexity of social phenomena in general and some abstract concepts in particular have led to the consideration of groups of independent variables. Blalock (1969) utilizes groups of independent variables in his block-recursive technique. Mood (1971), Mayeske, et al. (1969), and Wisler (1969) suggest the grouping of independent variables as a means of facilitating the interpretation of commonalities. Others have suggested using groups of indicators to represent abstract concepts (called the multiple indicator approach) rather than forming composites. (Note: the multiple indicator approach was discussed briefly in Chapter 9.) The multiple partial correlation, based on a simple extension of the formulas for multiple and partial correlations discussed
above, provides a method for computing a multiple correlation between a dependent variable and a group of independent variables, while controlling for one or more independent variables (Blalock, 1960: 458-459). The general formula consists of:

\[
\begin{align*}
    r_i^2 &= \frac{R^2_{i.jk...w} - R^2_{i.tu...w}}{1 - R^2_{i.tu...w}} \\
    i(jk...n).tu...w
\end{align*}
\]

suppose we wanted \( R^2_{y(12),34} \)

then \( r^2_{y(12),34} = \frac{R^2_{1234} - R^2_{34}}{1 - R^2_{y.34}} = \text{Squared Multiple Partial Correlation} \)

The numerator consists of the proportion of variation explained by variables 1 and 2, after the variation of variables 3 and 4 have been removed, and therefore represents the squared semipartial correlation. Thus the relationship between the partial and semipartial correlations discussed above also holds between the multiple partial and semipartial correlations, i.e.,
Multiple Partial = \[
\begin{vmatrix}
1 \\
\frac{1}{1 - R^2_{y.34}}
\end{vmatrix}
\]
(Semipartial correlation)

where

\[
R^2_{y(12.34)} = R^2_{y.1234} - R^2_{y.34}
\]

= Squared semipartial correlation.

Charles Mulford, Iowa State University, has grouped the independent variables in his causal model of Effectiveness (as elaborated in the Mulford, et al., 1972, article) into what he describes as Career Phases. These consist of (A) Initiation Decision, (B) Interaction, (C) Evaluation, and (D) Achievement. Suppose we examine the effect of the first three phases on role performance, (Achievement). In order to work out the required multiple partial correlations, four regressions will be required, i.e.,

1. \(Y\) on \(X_1, X_2, X_3, X_4, X_5, X_6, X_7\) (Y on Phase A, B, C)
   \[R^2 = .53019\]

2. \(Y\) on \(X_4, X_5, X_6, X_7\) (Y on Phases B, C)
   \[R^2 = .47553\]

3. \(Y\) on \(X_1, X_2, X_3, X_6, X_7\) (Y on Phases A, C)
   \[R^2 = .47553\]

4. \(Y\) on \(X_1, X_2, X_3, X_4, X_5\) (Y on Phases A, B)
   \[R^2 = .52034\]
Figure 11.4: The effect of career phases on Role Performance (Achievement)
It follows, then that the multiple partials consist of:

\[ r^2 = \frac{R^2_B - R^2}{1 - R^2} = \frac{D_{ABC} - D_{BC}}{D_{BC}} = \frac{.53019 - .47553}{1 - .47553} \]

\[ r = .32281 \]

\[ F = 8.996 \text{ where } F = 3.88 \text{ at .01 level} \]

*where the subscripts denote phases rather than variables

\[ r^2 = \frac{R^2_B - R^2}{1 - R^2} = \frac{D_{ABC} - D_{AC}}{D_{AC}} = \frac{.53019 - .40670}{1 - .40670} \]

\[ r = .45662 \]

\[ F = 30.49073 \text{ where } F = 4.71 \text{ at .01 level} \]

\[ r^2 = \frac{R^2_B - R^2}{1 - R^2} = \frac{D_{ABC} - D_{AB}}{D_{AB}} = \frac{.53019 - .52034}{1 - .52034} \]

\[ r = .14326 \]

\[ F = 2.430 \text{ where } F = 3.04 \text{ at .05 level} \]

where the F-value is calculated by

\[ F = \frac{R^2 - \frac{1}{k} \sum_{i=1}^{k} \frac{1}{N - k - 1}}{1 - R^2} \]

The multiple partial correlations for phases A and B proved to be statistically significant, whereas the multiple partial
for phase C did not. Substantively, then a multiple partial correlation reflects the amount of influence that a variable (or group of variables) has on a dependent variable, when it is entered last, for a given set of variables. Given this interpretation for the multiple partial correlation, then, Figure 11.5 becomes

![Diagram](image)

Figure 11.5: Models with significant coefficients showing the effect of Career Phases on Role Performance

The path between the group of variables in the third phase (evaluation) and role performance has been omitted because it was not statistically significant. By examining the magnitudes of the two significant paths, one would conclude that the second phase (Interaction) contains the most important group of variables, given the model under investigation. Note that the relative contributions reflected by the multiple partial coefficients would change with the addition or deletion of one or more variables or
sets of variables. Furthermore by treating the variables included in a block as indicators this same procedure can be adapted to the use of multiple indicators.

A comment on the use of multiple indicators In the case of multiple indicators, the block (or phase as described in the previous example) becomes the variable. The multiple indicator approach seems to be most beneficial when the concepts under examination prove to be heterogeneous rather than homogeneous. The method is still in the developmental stages but those who are interested in reading further should see Curtis and Jackson (1962), Jacobson and Lalu (1974), Costner and Schoenberg (1972), Sullivan (1971 and 1974) and Burke and Schuessler (1973-1974).

This chapter has devoted a great deal of space to discussing various procedures for examining the relationship between variables. The selection of one or more of these techniques should depend upon theoretical considerations; that is, which technique(s) answers the type(s) of substantive questions the researcher is interested in addressing? Each of these techniques involves the adoption of certain assumptions. The degree to which these assumptions can be met tend to have differential effects on analytical results and their interpretation.
Examining the Assumptions of Multiple Regression

Violations concerning assumptions of homoscedasticity (equal variances for each level of X) and normality do not generally cause serious distortions (Bohrnstedt and Carter, 1971; Zeller and Levine, 1974). Furthermore, whatever distortion that does exist, "the effects are [generally] conservative; that is they lead to errors of underestimation of the size of relationships" (Borgatta and Bohrnstedt, 1972:11). There has been considerable controversy over the assumption requiring interval measures, with purists such as Mayer (1971) and Wilson (1971) arguing that the formal properties of ordinal type variables are too weak to substantiate conclusions about theoretical models. Borgatta and Carter disagree with this purist approach and describe a number of studies (Labovitz, 1967, 1970a; Baker, et al., 1966; Borgatta 1968, 1970) which have reached the opposite conclusion:

"The point is that under almost any conceivable research situation, statistical tests are robust enough to allow the researcher to use them with little fear of gross errors regardless of whether or not he has an interval or ratio scale so long as his ordinal measure is monotonically related to the underlying true scale."

Furthermore, dummy variables can be utilized with ordinal or even nominal data if one does not want to violate this assumption, providing the dependent variable is more than nominal (Boyle, 1970; Lyons and Carter, 1971).
Violation of the three types of assumptions (homoscedasticity, normality and interval measures) just discussed generally have minimal effects on estimates. Multicollinearity and small sample size can also be minimal, except in extreme cases. Regression coefficients tend to become more unstable as the intercorrelations between independent variables increase and the sample size gets smaller.

"The rule, then, is to use large samples, (over 100 and preferably 200 or more), and independent variables whose intercorrelations are as low as possible" (Kerlinger and Pedhazur, 1973:442).

It should be noted that estimates of both \( R^2 \) and the partial regression coefficients will usually be biased (\( R^2 \) is over estimated; single regression coefficients are underestimated; and partial regression coefficients are either over or underestimated). Furthermore, the bias tends to decrease as the sample size increases. On the other hand, procedures have been developed to minimize these effects when small sample sizes must be used. Labovitz (1965) discusses three such approaches.

Apparently the two assumptions which prove to be the most troublesome, especially with respect to multiple regression, concern \textit{specification error} [that we have included all of the pertinent variables as well as the correct form (linear vs. curvilinear) of their relationship]
and measurement error (Bohrnstedt and Carter 1971). Obviously, if the model is wrong, the regression has little or no meaning. Reliance on theory in the building of models and the comparison of alternative models (as discussed earlier in this chapter) represent the best approaches to minimizing the effect of specification error.

In connection with the specification problem the theorist should consider the possibility of one or more moderator variables having an effect upon results. Sometimes relationships between concepts hold for some groups of units but not others. When this happens the relationship is said to be affected by a moderator (a variable that serves as the basis for dividing a set of units into meaningful subunits). Both behavioral and socio-demographic variables have been shown to act as moderators.

The subject is raised here under specification error because failure to recognize the existence of a moderator in the real world can affect both measurement and statistical test results (Warren et al., 1973). The reader should therefore be aware of their implications and make an attempt to assess the possibility of their presence.

In the example data set, it can be demonstrated that state serves as a moderating variable. This should not be surprising since the states were purposively chosen precisely because they each represented a different part of the country.
and a different type of organizational hierarchy. In such a case, the surprise would occur if state didn't prove to be a moderator.

In terms of the above statement, what are the implications of basing our interpretations on results using the \( N=240 \) (i.e. of combining the state samples into a single sample of \( N=240 \))? Detailed examination of the moderating effect of state (which will be presented briefly in Chapter 13) reveals that it is important in this data set, but only for some variables in certain states. Procedures for examining the differential effects of moderators have been described in greater depth by Specht and Warren (1974).

The assumption which is probably violated most often and with the most serious implications is the one that assumes the variables are measured without error. Few sociological indicators have a reliability of over .8; therefore, according to Bohnstedt and Carter, we can expect parameter estimations in the two-variable case \((Y \text{ on single } X)\) to be attenuated (underestimated). Path coefficients and partial regression coefficients can be either under or over estimated. Until recently, most sociologists seemed to be content to ignore the situation or, at best, lament the lack of precision and then proceed as if it didn't exist.

"Except for a few noted exceptions, sociologists seem to be blatantly unconcerned with the problems of measurement error. The most cursory review of the major journals should
convince one of this fact" (Bohrnstedt and Carter, 1971:142).

The review by Stokes and Miller (1975) substantiates the statement above. They, too, found very little consideration of reliability and validity or measurement error (see the earlier discussion on page 5). This need not be the case, however. There are approaches available to correct our estimates for at least random measurement error. These will not be discussed here. For those who wish to know more about these techniques, see Warren et al. (1974).

Statisticians speak in terms of the "robustness" of various statistical tests. This is "the ability of a statistical test to maintain its logically deduced conclusion when one or more assumptions have been violated" (Labovitz, 1967). In general most of the parametric techniques are considered fairly robust making Kish's approach (discussed on page 335) a reasonable one.

Summary

This chapter has been concerned with a discussion of multiple regression with special emphasis on interpretation of results. Many of the issues which arise in connection with the use of multiple regression have been discussed in this chapter. Some of these include the advantages of examining residuals; comparison of alternative models; the most appropriate type of procedure for each of the three
types of situations in which theorists (and/or researchers) often find themselves with respect to the development of theory; and a number of criteria that can be examined in an attempt to arrive at explanatory (as opposed to purely predictive) inferences.

The final topic in this chapter concerned a discussion of common assumptions for multiple regression. This discussion dealt not only with a description of the assumptions but also with how to examine the degree to which those assumptions have been met as well as the possible implications of failing to meet them. Descriptions are fairly common in most works on multiple regression but the other two aspects are not.

Many of the issues and assumptions discussed in this chapter are also important for path analysis—the analysis technique which will be examined in the next chapter. Actually path analysis is a type of regression technique based on both the assumptions of multiple regression plus some additional one's concerning the causal model as a whole. These will be taken up more fully in the next chapter.
The technique of path analysis was introduced and applied to genetic research by Sewell Wright in the early 1920s. Almost forty years later social scientists began to recognize its potential utility for model testing and theory construction in their own substantive areas. In terms of its application to sociological concerns, the earliest attempts included Simon's formulations in the 1950's and Blalock's extension of Simon's techniques. Their approach examines the correlations between hypothesized relations with respect to whether they are direct, indirect or spurious. Boudon (1965) later described their approach as simply a special or weak form of path analysis. However, the work which is considered to be the classic example of path analysis as applied to sociological data is O. D. Duncan's (1966) "Path Analysis: Sociological Examples." The publication of this article marked the real beginning of the use of path analysis in sociology.

Briefly stated, path analysis represents an analytical

\footnote{Kerlinger and Pedhazur make the statement that path analysis is useful in testing theory not generating it but we will demonstrate that path analysis has application for exploratory situations as well, but only if at least a general causal ordering is known.}
technique for examining the interrelationships and relative contributions of variables in a theoretically prescribed model. As such, it spells out explicitly the basic theoretical assumptions underlying the model to a greater extent than is required in ordinary regression analysis. According to Duncan, this represents its major advantage over other possible techniques.

"The great merit of the path scheme, then, is that it makes the assumptions explicit and tends to force the discussion to be at least internally consistent so that mutually incompatible assumptions are not introduced surreptitiously into different parts of an argument extending over scores of pages" (Duncan, 1966:4)

As noted earlier in the text, path analysis has become a popular technique among current journal contributors and, even though it has sometimes been misused, it nevertheless represents an important tool with which sociologists should be acquainted. This chapter will therefore be concerned with developing an understanding of the path analytic approach. As such, it will be concerned with more than mere description of procedure, rather I shall first deal with the dependence of the technique on a theoretical framework followed by a discussion of the types, and implications, of the assumptions required for its use.

The basic principles of the approach will be described briefly and then illustrated with the example data used in the previous chapters. Different modes of interpretation
(scientific versus applied) will also be discussed as well as the implications of various criticisms that have been made with reference to its current application (or at times, misapplication) in sociology. One such criticism deals with the indiscriminant usage of the technique without providing the proper theoretical foundation [e.g. as when one erroneously uses it in a pure "C" situation (see Chapter Twelve) where practically no theoretical guidelines exist.] Because of its special importance, therefore, I shall deal with this aspect first.

The Role of Theory in Path Analysis

It was explained in the introduction that one of the primary goals of this dissertation was to give the reader a feel for the integration of theory and methods. As we shall see, the path analysis approach is particularly well suited to achieve that goal, since its legitimate use is limited to either a situation "A" (fully specified model) or a situation "B" (model at least partially specified).

Theory plays an important part in the path analysis technique whether one is interested in theory building or theory testing, precisely because it is validity of the model that is examined, and which therefore determines the degree of certainty one can place on the path analysis results. A path analysis is really quite meaningless in a pure situation "C", where nothing is known about the model. While it is
quite true that one can test for goodness of fit with the path analysis technique, it is nevertheless also true that any number of models can be drawn that would achieve an equally good fit with the data (Kerlinger and Pedhazur, 1973:307). Only theory can determine which model approximates reality the best.

This does not mean that path analysis should be restricted to model testing alone. Model building is also a proper activity for path analysis, but only where it is possible to specify at least a causal ordering of the variables. For instance, as we see in Figure 12.1, one might argue that since selectivity and socialization occur early in one's job career they should come before scope and pervasiveness, because the effect of these variables can only be felt after one had been in the organization for a time. Likewise one might place scope and pervasiveness in a temporal order preceding the dependent variable, Role Performance, because they represent interaction variables which the human relationists feel influence effectiveness or achievement. (Also they could be viewed as throughput variables in a general systems framework).

While this example does not seem to contain much theory, it should be emphasized, again, that the more theoretical reasoning which goes into such a model, the more the researcher can rely on his (her) results. Thus in model
testing the researcher should specify, within the model, the exact nature of the expected relationships. This of course, can not be done unless a fairly well-developed theory (as in a high situation "B" or situation "A") already exists. If, for example, we were interested in testing the model in Figure 12.1, then we would need to specify the nature of the relations between the variables that we expect to find, rather than just a general temporal ordering. If we are testing theory, then the relations in the model should be based on theoretical evidence. Such evidence (Mulford, et al., 1972 and Warren, et al., 1975) can easily be found to support the slightly different but more specific model found in Figure 12.2 below.
Theory, then, can be said to be essential to the path analysis technique. Even when that theory is not exact, the method depends upon theory to at least specify the order if not the specific relations. Furthermore, when relations in the path model have been completely specified on the basis of theory and then upheld by the analysis, the researcher can claim far more confidence in his data manipulation than would be possible if there had been little or no theory. This then brings up the problem of specification as well as the assumptions in path analysis which concern themselves with

1The curved arrow between Scope and Pervasiveness is meant to indicate covariation (a noncausal relation) whereas the straight arrows signify that a causal relationship is expected between the two variables in the direction of the arrow.
Specification

Specification is the problem of determining (a) which variables should be included in the model (including the the form they may take - \( X_1, X_2, X_3 \), etc.) and (b) at least a weak causal ordering of those variables. The importance of theory with respect to the specification problem cannot be overemphasized. The relevance or meaningfulness of a path model rests upon the ability of the researcher to select the "true" model. It follows then that the first assumption implies that the selected model is the correct one.

Assumption #1: The relevant variables can be theoretically identified and included in the model.

Furthermore, the model represents the "true" relationships of those variables.

Assumption #2: The causal laws governing the system are established sufficiently to specify the causal priorities among variables in a way that is undebatable (Heise, 1969:52).

It should be noted here that neither the author nor Heise are implying the necessity of a full-scale theory (Situation A). Rather, one must have at least a partial theory "which permits the ordering of variables in terms of their causal priorities."
Considerable space was devoted in Chapter Eight to a general discussion of causality and its importance for the development of causal models. Nie, et al., (1975) have offered an "operational" definition of causality that applies to the specific case of path analysis.

"X is a cause of X if and only if X can be changed by manipulating X and X alone. Note, first, that the notion of causation implies prediction, but prediction of a particular kind. It implies the notion of possible manipulation...to understand what is meant by 'alone' in the definition one must understand the notion of causal hierarchy and the notion of relevant control...note [further] that manipulation of X alone does not imply that all the other causes of X are controlled or held constant" (1975:384).

This notion of cause led Heise and others to make the additional assumptions of no causal feedback.

**Assumption #3:** The system of concern contains no reciprocal causation or feedback loops; i.e., if X causes Y, Y can not affect X, either directly or through a chain of other variables.

This assumption is necessary when the ordinary least squares regression procedure is used to determine the path coefficients. The first two models on the next page would not meet this assumption. However, other techniques are available, making this assumption necessary only when the least squares solution is used.

For instance, Figure 12.3(a) could be solved using a two-stage least squares, i.e.
For the most part, these techniques are beyond the scope of this text. Therefore, Assumption #3 will apply unless otherwise noted.

Once the model has been specified according to the first three assumptions above, the next step concerns the problem of estimating parameters for each of the hypothesized paths. This is referred to in the literature as the identification problem.

**Identification**

"Identification is the problem of estimating the unknown parameters in a model from available empirical data" (Heise, 1969:52). Identification becomes a problem when one does not have experimental or longitudinal data. This means that the solution must rest on the known empirical correlations between the variables. However, if one has three independent variables, this results in 8 unknowns,
Unless something is done, this cannot be solved. This problem therefore leads to the adoption of a fourth assumption.

**Assumption #4**: The path coefficients of a recursive model can be identified from the empirical correlations among variables if disturbances in the dependent variables are controlled with each other or with the inputs (exogenous variables).

What this means is that the solution, using empirical correlations, can be solved if one assumes that correlations containing an error variable \( r_{st}, r_{sl}, r_{tl} \) are equal to zero. This assumption then reduces the number of unknowns so that a solution can be obtained for the remaining three unknowns. Note \( P_t \) and \( P_3 \) can be solved in terms of the first three unknowns, given the previous assumptions.

The solution which is generally utilized is that of least squares. When this procedure is selected, one must make the additional assumptions that are made for linear regression.

**Assumption #5**: (a) Sample units are drawn independently
(b) Interval level of measurement
(c) Homoscedasticity is a factor
(d) Multicollinearity is not a factor

and further,

**Assumption #6:** Measurement instruments used to obtain empirical data have high reliability.

These constitute most of the assumptions that are generally made and/or required when one uses the path analysis technique. Having discussed the assumptions we shall present a brief summary of the steps involved in the method and then illustrate the procedure with an empirical example.

**The Method**

Path analysis has been described by Sewell Wright (1921) as a method of combining the knowledge of the degree of correlation among variables with theoretical knowledge as to the causal relationships that exist between them. Theory and data then are combined "to determine whether a proposed set of interpretations is consistent throughout". The following is a brief summary of the steps that are generally taken in utilizing a path analytic approach.

1. Select and/or develop a model using all available theoretical evidence. At a very minimum, temporal and/or causal ordering must be established.
2. Write a set of recursive equations which correspond to the paths depicted in the model. Exogenous variables may be set equal to zero.

\[ X = e_1 \]
\[ X = e_2 \]
\[ X = b_3 X + b_{31} X + e_3 \]
\[ X = b_4 X + e_4 \]

3. Examine the intercorrelations of the variables. Compare the results to your model. Correlations which typify hypothesized paths should have more significant values than those corresponding to nonexistent paths, e.g., we could expect that \( r_{23} > 0 \) and \( r_{24} = 0 \), or at least significantly low.

4. Run the required regressions in order to calculate the partial regression coefficients. Interpret your results. Would need to run 2 regressions:
5. Part 4 above can be repeated until the model contains only significant paths. Suppose \( b^* \) was not significant, would then run \( X \) on \( X \) only.\(^1\)

6. Calculation of residuals for each endogenous variable. Tells you how much of this particular variable was not explained by the variables in the model.

7. Examine direct and indirect effects.

8. Interpretation

This represents a brief summary of the various steps in path analysis. We shall now follow these steps by developing a theoretical model and then testing it with path analysis.

---

\(^1\)This has been the traditional method as described by O. D. Duncan (1966). Some have recently criticized the re-running of the model to eliminate insignificant paths (Step 5 above), especially where the model was only partially specified by theory (Situation - B). In the latter case, Step 5 really amounts to the building and testing of a model with the same data set (not an entirely legitimate operation). The reader should, therefore, be aware that the further one gets from a Situation A model, the less certainty he/she can place in the revised model. Consequently, there may be times when the theorist may want to omit Step 5 and proceed to 6.
Examination of an Effectiveness Model

The model on the following page was developed from four primary sources. The major part of the model was based on the following: (1) Amitai Etzioni's (1961) work on the correlates of effectiveness, (2) the Mulford, et al., (1972) paper attempting to create and test a model based partially on Etzioni's work, and (3) a paper by Warren, Faisal, Mulford, Klonglan and Fuller (1975) attempting to re-examine the Mulford, et al., model when considering measurement error. Finally the variable job satisfaction has been added to the model as another type of effectiveness, along with role performance, on the basis of the research (4) done by Vroom (1964), March and Simon (1958) and Scott (1962), all of which have demonstrated only a slight relationship between the two effectiveness variables. (Vroom found an average correlation of .14 when he examined a sizeable number of studies attempting to relate the two.) As a result no causal arrow will be hypothesized between the two effectiveness variables. The following set of recursive equations represent the relationships depicted by the model in Figure 12.4.
Figure 12.4: Model developed on the basis of theory
Scope is said to be influenced by Socialization ($X_1$) and Communication ($X_2$) on the basis of the Warren, et al., (1974) and the Mulford, et al., (1972) final models. Pervasiveness is said to be influenced by Socialization only, as per the final Warren, et al., model. Likewise Salience is said to be influenced by scope as a result of the Warren, et al., study. Tension, which was not examined in the Warren, et al. paper, is said to be influenced by both Scope ($X_3$) and Salience ($X_5$) as per the Mulford, et al., final model. Role performance ($X_8$) is said to be influenced by scope ($X_3$) and
perceived salience (X^3) and tension (X^5) on the basis of the final model in the Warren et al., study.

The basis for including job satisfaction in the model as a second effectiveness variable has been touched on briefly above. Talcott Parson's delineated four types of problems which must be solved for the system to remain effective (viable). Role Performance represents the Goal Attainment problem while job satisfaction is a solution to the latent pattern maintenance problem. The theoretical reasoning for saying that scope (X^1) pervasiveness (X^4) salience (X^3) and tension (X^5) influence job satisfaction is not as well backed by theory or research findings. However, these types of variables reflect some of the same types of dimensions as those which concerned the human relations studies on job satisfaction.

Having established a model on the basis of available theoretical evidence, the next step is to examine the correlations between the variables. The correlations representing relationships for which we have hypothesized paths should be significant. Table 12.1 on the next page shows that most of the hypothesized relationships do have fairly high correlations. The exceptions involve pervasiveness and tension (note the weak correlations in boxes for these two variables.)
Table 12.1: Zero order correlation coefficients for variables in model

<table>
<thead>
<tr>
<th>Variables</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>X Socialization</td>
<td>1.000</td>
<td>.393</td>
<td>.529</td>
<td>.230</td>
<td>.136</td>
<td>.198</td>
<td>.368</td>
<td>.477</td>
</tr>
<tr>
<td>X Communication</td>
<td></td>
<td>1.000</td>
<td>.557</td>
<td>.220</td>
<td>.207</td>
<td>.205</td>
<td>.187</td>
<td>.510</td>
</tr>
<tr>
<td>X Scope</td>
<td></td>
<td></td>
<td>1.000</td>
<td>.368</td>
<td>.347</td>
<td>.364</td>
<td>.357</td>
<td>.637</td>
</tr>
<tr>
<td>X Pervasiveness</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.276</td>
<td>.064</td>
<td>.133</td>
<td>.452</td>
</tr>
<tr>
<td>X Salience</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.268</td>
<td>.108</td>
<td>.338</td>
</tr>
<tr>
<td>X Tension</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.092</td>
<td>.236</td>
</tr>
<tr>
<td>X Job Satisfaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.343</td>
</tr>
<tr>
<td>X Role Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>
After examining the correlations, the next step in the procedure is to run the prescribed regressions as listed in the set of recursive equations. The standardized partial regression coefficients (betas) are then placed on the arrows in the model. [NOTE: curved arrows represent relationships not causal ordering. Therefore the figure on the curved arrow is a correlation rather than a beta.]

\[ B = .94 \]

\[ \text{SAL.} \]

\[ X \]

\[ .76 \]

\[ .93 \]

\[ .367 \]

\[ \text{SCOT.} \]

\[ X \]

\[ .389 \]

\[ .423 \]

\[ .347 \]

\[ R = .76 \]

\[ \text{COMM.} \]

\[ X \]

\[ .252 \]

\[ .545 \]

\[ R = .93 \]

\[ \text{PERVT.} \]

\[ X \]

\[ .302 \]

\[ R = .97 \]

\[ \text{ROLE PERF.} \]

\[ X \]

\[ .161 \]

\[ R = .73 \]

\[ \text{TEN.} \]

\[ X \]

\[ .92 \]

\[ R = .92 \]

Figure 12.5: Path Model I with Significant Path Coefficients
This model contains only the significant paths from the first regression (residuals calculated on the basis of first regression).

When the regressions are repeated using only the significant variables shown in Figure 12.5 above, we get the following model:

\[ R = .94 \]

\[ R = .76 \]

\[ R = .93 \]

\[ R = .97 \]

\[ R = .73 \]

\[ R = .92 \]

Figure 12.6: Revised model
Thus all of the hypothesized paths were significant except some of those going to job satisfaction. In the latter case, only scope had a significant path to job satisfaction. A curved arrow between job satisfaction and role performance was also added reflecting the sizable correlation which was unexpectedly found. Evidence in support of most of the hypothesized relationships is not unexpected nor can it really be considered as "evidence in support" since both the Mulford and the Warren studies used this same set of data (with minor modifications of the Role Performance concept). However, job satisfaction was not included in these studies. Several points about this effectiveness variable bear consideration in so far as it differs from the initial model.

The unexpectedly high correlation between role performance and job satisfaction might support the Porter and Lawler model Role Performance \( \rightarrow \) Job Satisfaction (1968). A regression was run (in addition to those specified in the model) with Scope, Pervasiveness, Tension, Salience and Role Performance on Job Satisfaction. Both Scope and Role Performance turned out to be significant. When job satisfaction was included with scope and pervasiveness on role performance there was still a significant path from job
satisfaction to role performance but the magnitude was much less. This tends to give some partial support then to the model by Porter and Lawler described above.

The only paths that were not significant in the original model dealt with those leading to job satisfaction. Basing results on the first set of Job satisfaction equations, the regression was re-run, regressing only scope on job satisfaction. The results were as follows:

<table>
<thead>
<tr>
<th>$X$ (Job Satisfaction)</th>
<th>$F$</th>
<th>$E^*$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>on $X$ (Scope)</td>
<td>37.023**</td>
<td>0.367</td>
<td>0.135</td>
</tr>
</tbody>
</table>

According to the second equation in the Table above we should have run both scope and role performance on job satisfaction; however, this was thrown in after the analysis was done, in view of the unexpectedly high correlation between job satisfaction and role performance. As an aside note of interest, since there was no relationship hypothesized between them, two differences of $R^2$ tests will be run to see if (1) adding role performance made a significant difference in explaining job satisfaction and (2) if adding job satisfaction made a difference in explaining role performance, since the direction has not yet been
Table 12.2: Partial Regression analysis of Path Model I

<table>
<thead>
<tr>
<th>Variables</th>
<th>F Value</th>
<th>β</th>
<th>% Explained Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Scope</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Socialization</td>
<td>.46.872**</td>
<td>.367</td>
<td>.424</td>
</tr>
<tr>
<td>X Communication</td>
<td>.59.475**</td>
<td>.413</td>
<td></td>
</tr>
<tr>
<td>X Pervasiveness</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Socialization</td>
<td>.13.314**</td>
<td>.230</td>
<td>.053</td>
</tr>
<tr>
<td>X Salience</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope</td>
<td>.32.54**</td>
<td>.347</td>
<td>.120</td>
</tr>
<tr>
<td>X Salience</td>
<td>.6.37*</td>
<td>.161</td>
<td></td>
</tr>
<tr>
<td>X Tension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope</td>
<td>.23.42**</td>
<td>.308</td>
<td>.155</td>
</tr>
<tr>
<td>X Salience</td>
<td>.6.37*</td>
<td>.161</td>
<td></td>
</tr>
<tr>
<td>X Job Satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope</td>
<td>.29.663**</td>
<td>.389</td>
<td>.137</td>
</tr>
<tr>
<td>X Pervasiveness</td>
<td>.022</td>
<td>-.003</td>
<td></td>
</tr>
<tr>
<td>X Salience</td>
<td>.044</td>
<td>-.013</td>
<td></td>
</tr>
<tr>
<td>X Tension</td>
<td>.470</td>
<td>-.045</td>
<td></td>
</tr>
</tbody>
</table>

*Takes an F=3.04 at .05 level of significance
**F=4.35 significant at .01 level
<table>
<thead>
<tr>
<th>Dependent and Independent Variables</th>
<th>F Value</th>
<th>$B^*$</th>
<th>(%) Explained Variance ($R^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Job Satisfaction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope 3</td>
<td>11.680**</td>
<td>-291</td>
<td>.160</td>
</tr>
<tr>
<td>X Pervasiveness 4</td>
<td>0.595</td>
<td>-.053</td>
<td></td>
</tr>
<tr>
<td>X Salience 5</td>
<td>0.250</td>
<td>-.033</td>
<td></td>
</tr>
<tr>
<td>X Tension 6</td>
<td>0.524</td>
<td>-.047</td>
<td></td>
</tr>
<tr>
<td>X Role Performance 3</td>
<td>6.516*</td>
<td>.210</td>
<td></td>
</tr>
<tr>
<td>X Role Performance 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope 3</td>
<td>112.763**</td>
<td>.545</td>
<td>.461</td>
</tr>
<tr>
<td>X Pervasiveness 4</td>
<td>24.082**</td>
<td>.252</td>
<td></td>
</tr>
<tr>
<td>X Role Performance 8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Scope 3</td>
<td>84.945**</td>
<td>.498</td>
<td>.475</td>
</tr>
<tr>
<td>X Pervasiveness 4</td>
<td>24.661**</td>
<td>.252</td>
<td></td>
</tr>
<tr>
<td>X Job Satisfaction 7</td>
<td>6.211*</td>
<td>.126</td>
<td></td>
</tr>
</tbody>
</table>
established.

Job Satisfaction as Dependent:

\[
R^2 - R'^2 / d - d
\]

\[
F = \frac{.160 - .137/5-4}{.023} = 6.407*\]

Pole Performance as Dependent:

\[
-.475 - .461/3-2 = .014\]

It is apparent that the Pole Performance \(\rightarrow\) Job Satisfaction adds slightly more to explaining the relationship than does the Job Satisfaction \(\rightarrow\) Role Performance although the increase is not especially large. However, if the relationship was examined (corrected) for measurement error, support might well be added to the Porter and Lawler model (the first direction above) particularly if scope was also considered.

Decomposition of Effects

One type of analysis which we did not consider in the example above concerns the decomposition of effects, both total and indirect. The model in Figure 12.6 is more complex than necessary for the purpose of illustrating this type of procedure. Therefore, we will drop Tension, Salience and Job Satisfaction, leaving a model somewhat similar to that
derived by Warren, et al., (1975), except that is a completely recursive model, and illustrate a new procedure for decomposition of effects.

In terms of equations, the model above translates itself into the following set:

\[ X = e \]
\[ X = e \]
\[ X = b_3 X + b_{31} X + b_{32} X + e \]
\[ X = b_4 X + b_{41} X + b_{42} X + X + e \]
\[ X = b_5 X + b_{51} X + b_{52} X + b_{53} X + e \]
The original method for breaking down the total effect was developed by Duncan. However, he has recently opted for acceptance of a new procedure - that of Alwin and Hauser (1975) - which will be discussed here.

Alwin and Hauser approach to decomposition of effects

We shall begin with a correlation table for the set of five variables. The table indicates that a significant correlation exists for all of the hypothesized paths. [NOTE this does not verify the model, only indicates that some relationship does exist where we have hypothesized a causal path.]

Table 12.3: Correlation table

<table>
<thead>
<tr>
<th>Variable</th>
<th>SOC.</th>
<th>COM.</th>
<th>Scope</th>
<th>PERV.</th>
<th>R.P.</th>
</tr>
</thead>
<tbody>
<tr>
<td>X 1</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X 2</td>
<td>0.393</td>
<td>1.000</td>
<td>0.557</td>
<td>0.220</td>
<td>0.510</td>
</tr>
<tr>
<td>X 3</td>
<td>0.529</td>
<td>0.557</td>
<td>1.000</td>
<td>0.368</td>
<td>0.637</td>
</tr>
<tr>
<td>X 4</td>
<td>0.230</td>
<td>0.220</td>
<td>0.368</td>
<td>1.000</td>
<td>0.452</td>
</tr>
<tr>
<td>X 5</td>
<td>0.477</td>
<td>0.510</td>
<td>0.637</td>
<td>0.452</td>
<td>1.000</td>
</tr>
</tbody>
</table>

^Takes an r of 0.116 at .05 level of significance
Returning to the model we can identify the following set of reduced-form regression equations:

(1) Regress $X$ on $X_3$, $X_1$, $X_2$

(2) Regress $X$ on $X_4$, $X_1$, $X_2$

(3) Regress $X$ on $X_4$, $X_1$, $X_2$, $X_3$

(4) Regress $X$ on $X_5$, $X_1$, $X_2$

(5) Regress $X$ on $X_5$, $X_1$, $X_2$, $X_3$

(6) Regress $X$ on $X_5$, $X_1$, $X_2$, $X_3$, $X_4$

The solution of these equations provides the following standardized partial regression coefficients:
Table 12.4: Standardized beta coefficients, reduced form

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Reduced-Form Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X on X X X X X</td>
</tr>
<tr>
<td>Soc</td>
<td>.36686  .17010  .04741  .32695  .16648  .15501</td>
</tr>
<tr>
<td>Com</td>
<td>.41325  .15295  .01475  .38120  .20044  .19687</td>
</tr>
<tr>
<td>Soc</td>
<td>.33443  .43742  .35650</td>
</tr>
<tr>
<td>Perv</td>
<td>.24195</td>
</tr>
</tbody>
</table>

Note: The reduced-form equation includes multiple X variables, indicating a regression model with multiple independent variables.
With this knowledge we can now begin to decompose the various effects. Before actually doing the calculations, however, it might be beneficial to consider the verbal definitions of these various effects as given by Alwin and Hauser.

**Total Effect**: The total effect of one variable on another is the part of their total association which is neither due to their common causes, to the correlation among their causes, nor to unanalyzed (predetermined) correlation.

A total effect tells us how much change in a consequent variable is induced by a given shift in an antecedent variable, irrespective of the mechanisms by which the change may occur.

**Indirect Effects**: Indirect effects are those parts of a variable's total effect which are transmitted or mediated by variables specified as intervening between the cause and effect of interest in a model.

That is, they tell us how much of a given effect occurs because the manipulation of the antecedent variable of interest leads to changes in other variables which in turn change the consequent variable.

**Direct Effects**: The direct effect of one variable on another is simply that part of its total effect which is not transmitted via intervening variables.

It is the effect which remains when intervening variables have been held constant.
The overall Total Effect of $X$ on $X$ can be represented by $g$ where $g = .32695$ as derived from the reduced form equation of P.P. on socialization and communication. Once we know the overall total effect of $X$ on $X$, we can calculate the total effects of $X$ which operates (1) through $X$, (2) through $X$ but not $X$, and (3) $X$ on $X$ unmediated by either $X$ or $X$.

1. The effect of $X$ on $X$ operating through $X$ is:

$$g_{15} - g^*_{51}$$

where

$$g^*_{51} = .16648 = B^* \text{ for } X \text{ on } X, X, X$$

$$g_{51} = .32695 = B^* \text{ for } X \text{ on } X, X$$

so that

$$(.32695 - .16648) = .16047$$

2. The effect of $X$ on $X$ through $X$, but not $X$ can be expressed by the following:

$$g^*_{51} - P_{51}$$

where $P$ is one of the path coefficients in the full model obtained by regressing $X$ on all the other variables in the recursive system, (i.e., $X$ on $X, X, X, X$).
so that

\[
\frac{q^* - P}{51} 51
\]

The effect of \( X \) on \( X \) through \( X \), but not \( X \),

\[
\begin{align*}
X_1 & \rightarrow X_5 \\
X_4 & \rightarrow X_3 \\
\end{align*}
\]

\[
(0.16648 - 0.15501) = 0.01147
\]

(3) The effect of \( X \) on \( X \) unmediated by either

\[
X_{15} \rightarrow X_{34}
\]

\[
\text{coefficient } P = \frac{q^* - P}{51}
\]

which results from the

\[
\text{regression of } X \text{ on } X, X, X, X. \text{ This }
\]

\[
\begin{align*}
X_5 & \rightarrow X_1, X_2, X_3, X_4 \\
\end{align*}
\]

Obviously follows from the fact that the path
coefficient represents the standardized partial
regression coefficients, controlling for the other
independent variables. Utilizing these calculated
values we can now partition the total effect into
its constituent parts, as shown in Table 12.5.
Table 12.5: Calculations for the decomposition of total effects

<table>
<thead>
<tr>
<th>Decomposition of Total Effect</th>
<th>Equation</th>
<th>% of Total Effect of total effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Effect of X on X</td>
<td>$1 - \frac{g^*}{g}$</td>
<td>$1 - \frac{.16648}{.32695} = \frac{.50919}{.32695} \approx 49.08%$</td>
</tr>
<tr>
<td>mediated by X</td>
<td>$\frac{(g - q^*)}{q}$</td>
<td>$\frac{(.16648 - .15501)}{.32695} \approx 3.50855$</td>
</tr>
<tr>
<td>Total Effect of X or X</td>
<td>$\frac{(g^* - P)}{q}$</td>
<td>$\frac{(.16648 - .15501)}{.32695} = \frac{.15501}{.32695} \approx 47.41%$</td>
</tr>
<tr>
<td>unmediated by X</td>
<td>$\frac{P}{q}$</td>
<td>$\frac{.15501}{.32695} \approx 47.41%$</td>
</tr>
<tr>
<td>mediated by X</td>
<td>$\frac{g^*}{g}$</td>
<td>$\frac{.16648}{.32695}$</td>
</tr>
<tr>
<td>unmediated by either X or X</td>
<td>$\frac{P}{q}$</td>
<td>$\frac{.15501}{.32695} \approx 47.41%$</td>
</tr>
</tbody>
</table>
Note that the total effect of $X$ on $X$ can now be broken down as follows:

\[
\text{Total effect of } X \text{ on } X = .32695 = \frac{1}{5}
\]

Mediating Effect of $X$

\[
\frac{3}{4} \quad + \quad \text{Mediating Effect of } X \text{ without } X \frac{4}{3}
\]

Unmediated Effect

OR

\[
.32695 = .16047 + .01147 + .15501
\]

\[
.32695 = .32695
\]

We can now examine the decomposition of the indirect effects, as shown in Table 12.6 on the next page.
Table 12.6: Calculations for the decomposition of indirect effects

<table>
<thead>
<tr>
<th>Indirect Effects</th>
<th>Equation</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indirect Effect of X&lt;sup&gt;1&lt;/sup&gt; on X directly Transmitted by way of X&lt;sub&gt;3&lt;/sub&gt;</td>
<td>((q - q^*) (\frac{p}{q}))</td>
<td>((0.32695 - 0.16648) \cdot 0.35650 = 0.35650)</td>
</tr>
<tr>
<td>Indirect Effect of X&lt;sup&gt;1&lt;/sup&gt; on X Transmitted by way of X&lt;sup&gt;3&lt;/sup&gt;'s Effect on X and X&lt;sup&gt;4&lt;/sup&gt;'s subsequent Effect on X</td>
<td>((q - q^*) (\frac{p}{q}))</td>
<td>((0.32695 - 0.16648) \cdot 0.35650 = 0.35650)</td>
</tr>
</tbody>
</table>
The mediating effect of $X_3$ on the relationship between $X_1$ and $X_5$ was found to be 49.08%. We can break this down into 40% [or $(.13078/\cdot.32695)$] being transmitted by $X_3$ directly and 9.08% [or $(.02968/\cdot.32695)$] being transmitted via the $X_3$ effect on $X_4$ and its subsequent effect on $X_5$.

Alwin and Hauser have suggested a method of arranging these calculations in a table that facilitates their interpretation. I shall follow their lead by presenting a similar table containing the appropriate calculation formulas (Table 12.7) and a table containing the values calculated above (Table 12.8). Before doing so, however, a few explanatory comments are in order. In the case where an exogenous variable affects an initial dependent variable (i.e., where there are no intervening variables such as $X_1$ and $X_2$ on $X_3$ or $X_4$ on $X_5$) the direct effect equals the total effect. Furthermore, where an independent variable has an effect on a dependent variable with only one variable intervening (such as $X_1$ and $X_2$ on $X_4$ with $X_3$ intervening), the indirect effect can be computed by subtracting the direct effect from the total effect (e.g., $q_{41}$ - $p_{41}$).

Definitions were given earlier for total effect, indirect and direct effects. Those definitions concentrated on the substantive meaning of the different components. The following descriptions, on the other hand, are aimed at aiding in computation and filling in the tables on the next
Total Effect of a variable is its coefficient in the first reduced form equation in which it appears as a regressor.

Noncausal Components: The sum of noncausal components of association may be found as the difference between a total effect and the corresponding zero order measure of association.

Direct Component: It represents the coefficient in the last (structural) equation in the sequence.

Indirect Components: Represent the differences between coefficients of a causal variable in two equations in the sequence, where the mediating variable (or variables) is that which appears as regressor in one equation and not in the other (Alwin and Hauser, 1975:42).

We have applied these and other directions in Table 12.7 on the following page. Utilizing the formulas in Table 12.7 and the empirical data in Table 12.4, we get the breakdown of effects shown in Table 12.8 for the variables in the model in Figure 12.7, (given this set of data).
Table 12.7: Calculation formulas for decomposing effects

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Independent Variable</th>
<th>Total Effect</th>
<th>Indirect Effect Via</th>
<th>Direct Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>3</td>
<td>1</td>
<td>31</td>
</tr>
<tr>
<td>X</td>
<td>1</td>
<td>32</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>X</td>
<td>q</td>
<td>41</td>
<td>1</td>
<td>41</td>
</tr>
</tbody>
</table>

or

\[ g - p \]

\[ 41 \]

2

3

4

5

or

\[ g - p \]

\[ 42 \]

3

4

5

Arranging the decomposition of effects in this fashion greatly facilitates their interpretation. Note that the results from Table 12.5 can now be found in Table 12.8, under
Table 12.8: Decomposition of effects with empirical data

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Independent Variable</th>
<th>Total Effect</th>
<th>Indirect Effect Via Variable 3</th>
<th>Indirect Effect Via Variable 4</th>
<th>Direct Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>0.36686</td>
<td>---</td>
<td>---</td>
<td>0.36686</td>
</tr>
<tr>
<td>X</td>
<td>1</td>
<td>0.41325</td>
<td>---</td>
<td>---</td>
<td>0.41325</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(72.13%)</td>
<td>(27.87%)</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>0.17010</td>
<td>0.12259</td>
<td>---</td>
<td>0.04741</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>(90.36%)</td>
<td>(9.64%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(49.08%)</td>
<td>(3.5%)</td>
<td>(47.41%)</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>0.32695</td>
<td>0.16047</td>
<td>0.01147</td>
<td>0.15501</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td></td>
<td>(47.4%)</td>
<td>(3.5%)</td>
<td>(47.41%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(47.4%)</td>
<td>(3.5%)</td>
<td>(47.41%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(47.4%)</td>
<td>(3.5%)</td>
<td>(47.41%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(18.5%)</td>
<td>(81.5%)</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>0.24195</td>
<td>---</td>
<td>---</td>
<td>0.24195</td>
</tr>
</tbody>
</table>
the column heading "Direct Effects." Furthermore, we can see from Table 12.8 that 72% of the effect of socialization \((X_1)\) on pervasiveness \((X_4)\) is mediated by scope \((X_3)\). On the other hand, 90% of the effect of communication \((X_2)\) on pervasiveness \((X_4)\) is mediated by scope \((X_3)\). This may help explain why the final model in the Warren, et al., 1975 paper included an arrow from socialization to pervasiveness but not from communication.

Alwin and Hauser (1975) conclude their article on a cautionary note that should perhaps be re-emphasized here. Examining the direct and indirect effects can be beneficial in sorting out inter-relationships. However, the procedure amounts to nothing more than busy work if the results are not interpretable.

"We should be disappointed if our efforts to elucidate such causal interpretations were to lead researchers to generate vast quantities of uninteresting or meaningless components. Sometimes a detailed interpretation will speak to an important research question, and other times it will not. We [meaning this method] offer no substitute for the thoughtful interpretation of social data" (p. 47).

What the authors were trying to emphasize is that their procedure, as with path analysis in general, must be tied to theory if it is to have any real meaning. Furthermore, this approach results in a decomposition of a coefficient rather than explained variance. Therefore, where one independent variable is relatively unimportant in explaining the
dependent variable, this approach really may not offer much insight after all. For example, if you find that that 90% of a coefficient (which was close to zero) has an indirect effect through another variable, you still do not know very much, since 90% of nothing (or nearly nothing) is till nothing. What I am saying, therefore, is that the reader may want to assess the importance of particular breakdown before creating an elaborate interpretation for a relatively unimportant variable. Another problem with this approach is that it requires a completely recursive model. A great many sociological models however do not meet this requirement.

Throughout this chapter, we have utilized standardized path coefficients in our discussion and models. However, there are good reasons for substituting the raw unstandardized coefficients for the standardized path coefficients. We shall discuss these reasons in the next section.

Standardized Versus Unstandardized Coefficients

There has been considerable discussion concerning whether or not one should use standardized or nonstandardized partial regression coefficients. Many of the arguments against the use of standardized coefficients are equally applicable to correlation coefficients since both represent standardized (unit free) measures.
Hammond (1973) states that correlations may give biased results in terms of analysis, aggregating and measurement theory, whereas slopes (unstandardized partial regression coefficients) are far less likely to do so.

Correlations and path coefficients can be changed by a change in standard deviations. Thus a few outliers in one sample can increase the standard deviation to the point where differences between the groups may be found where they don't really exist. Therefore, one should be certain that $\sigma_1^2 = \sigma_2^2$ (equal variance across the two groups) if the researcher wishes to compare results across populations.

Path coefficients and correlations can also be described as particularistic to the sample under investigation. This is essentially what Blalock (1971) is arguing when he discusses the use of standardized versus nonstandardized coefficients and closed populations. He states that path coefficients (standardized regression coefficients) are legitimate when one wants information concerning relative contributions of elements within a specific population (or sample).

"...we must keep clearly in mind the distinction between working with causal laws and unstandardized coefficients, and attempting to measure relative importance of variables in specific populations. We must recognize that relative importance cannot be evaluated in the abstract. The contribution of each factor to the total variation in a dependent variable is a function of how much the various independent variables happen to vary in that given
population" (p. 150).

Blalock concludes his discussion by saying that the unstandardized regression coefficients measure laws (invariant relationships from one population to the next) rather than specific changes within a single population. Therefore, he argues that the use of standardized coefficients be confined to descriptive or practical purposes, within a single sample or population. For comparisons across populations, however, one should only use the unstandardized coefficients which are not subject to the possible differences of variation between the two populations.

In the introductory chapter, some mention was made about criticisms concerning causal model analysis in general and path analysis in particular. I shall discuss a few of these criticisms in the next section.

Criticisms Concerning the Use of Path Analysis

(1) failure to use the procedure correctly

(2) whether sociological theory and methodology are as yet sophisticated enough to use the technique.

With regard to the first category, one sees such complaints as a failure to integrate theory and the path
model. It has been emphasized here that the model must be based or grounded in theory. This is not always done, however. Researchers have been known to simply draw up a model with little or no reference to theoretical justification. Others, it would seem have "slipped into their papers" what appears to be a path model simply because path models are the present fad and have otherwise ignored it throughout the remainder of their papers.

Another gross misuse of the method concerns the analysis of residuals. In the discussion and example, we utilized the conventional $\sqrt{1 - R^2}$ to measure the residual, defining it as the amount of variation of the dependent variable not explained by the variables in the model. In a survey conducted by Miller and Stokes (1975), they cited 9 out of 10 (89.6%) of the authors in their survey as not specifying what they used to calculate the residuals. If everyone used the conventional form this would not be a problem. However, some researchers report the coefficient of alienation $(1 - R^2)$ rather than $\sqrt{1 - R^2}$. Still others ignore the issue completely.

The magnitude of the residual term can be used as a method of evaluating a path model. Miller and Stokes make the charge that maximizing explained variance (thus minimizing the residual) has become first priority in research reports rather than selection of path models which
have eliminated weak paths (thus reducing the $R^2$).

The criticisms just considered have been failures of the researchers and could have been avoided if the method had been legitimately followed. The criticisms in the next category are not so easily disposed of. These concern the questions of whether or not sociological theory has advanced far enough to build models containing undebatable causal priorities, as called for by Heise (1969); and/or whether sociological methodology is sophisticated enough to measure those concepts once they have been theoretically specified.

In their survey, Miller and Stokes found that 40% of the residuals reported in their study were greater than or equal to .85. Twenty-five percent had residuals of over .90.

"One article in four explained less than 20% of the variance with an average of five independent variables" (p. 199).

Furthermore, only 20% had residuals of less than .70. One must question just how much relevance can be given to many of the models described in their study, particularly the 25% with residuals over .90. If the theoretical and methodological tools are truly lacking, we must consider whether we have gained anything once we have developed our path model, or as Sorokin (1956:133) has stated, whether we have:

"A beautiful blueprint of a marvelous building for whose construction the necessary material is unavailable."
Obviously I am not suggesting that path analysis be placed on a shelf (as some have done) to await the future development of theory and methods. However, the researcher should be aware of some of the common misuses and problems that have surrounded its use in the past so that they can hopefully avoid them in the future.

Summary

This chapter has been concerned with the technique of path analysis. Considerable emphasis has been placed on the necessity of building a theoretical model prior to (or as a first step) in utilizing this method. Unless the model is grounded in theory, path analysis is nothing more than a mathematical exercise.

The procedure itself was summarized briefly, with the major concern of the chapter being placed on interpretation and understanding rather than formulas. Most of the actual calculations are done on the computer anyway and are, therefore, not really necessary for gaining an understanding of the method. An empirical example, working through each of the steps, was also included to illustrate the technique.

Considerable space was given to the decomposition of effects as developed by Alwin and Hauser (1975). More attention was given to formulas with reference to this approach because it is fairly new and because it requires some hand calculation. As with the path analysis technique
in general, however, it was emphasized that this approach is meaningless unless tied to substantive theory.

Finally a number of common pitfalls or misuses were discussed, as was the criticism by some that sociological theory and methods are not yet ready for path analysis. These criticisms were raised because it was felt that the method could be more intelligently applied where the researcher is aware of such problems.

This concludes the discussion of analysis procedures. I have attempted to include some of the most widely used parametric techniques. Of course others could have been included, but these probably offer as useful a set of tools (in terms of applicability to most data sets and range of interpretation) as most of the readers of this dissertation will require.

The last phase, (Inference) of my integrated approach to theory construction will be dealt with in the next section. It will serve the dual purposes of (1) explicitly demonstrating the integrated nature of the inferential process and (2) comparing my approach to those that were discussed earlier in Part I.
INFEERENCE

The last major phase in the comprehensive, integrated approach to theory construction that is being discussed here is the inferential phase. Inference consists of the act of drawing conclusions on the basis of sample data and empirical tests and generalizing them to propositions that relate to the population.

As with activities throughout the other phases, inference must be done in a context of an integrated approach to theory construction if the resulting theory is to meet the criteria set forth in Chapter Four. Inferences are made from the point of data collection and continue throughout the final stages. They are influenced by decisions or choices that have been made throughout the entire theory construction process. As such the inferential phase serves the purpose of "putting it all together." In an attempt to try to "capture" or demonstrate the nature of the inferential process, (as described above), I have placed the discussion of inference within the framework of this example. This example is concerned with re-examining the model presented in the last chapter for possible moderating effects of state. The previous chapters were also illustrated at various points with real data, however those illustrations were restricted to the particular topic under discussion at that point in the process. In the next chapter, on the other hand, I will
attempt to give the reader a feeling of the process as a whole, as it relates to the generating of inferences, beginning with the theoretical development of a model (actually an extension of the framework used throughout the dissertation) and proceeding through the method of theory construction that has been the subject of Part II of this dissertation.

The generating of inferences represents the final phase of the method under discussion. As was true with the previous phases, inference should not be isolated from activities in the other phases. I have attempted to demonstrate that sound sociological theory construction should take place within a comprehensive, integrated procedure that consists of interdependent activities covering formulation measurement, analysis and inferential processes (phases). The last section of this chapter will consequently be concerned with assessing whether this text did, in fact, meet the objectives laid out in Chapter One as necessary for developing such a procedure.
CHAPTER THIRTEEN:  
DISCUSSION

This chapter will attempt to illustrate that inference is not just something that one slips into the end of a discussion. Inferential conclusions can be drawn from the sample data during measurement assessment as well as analysis. Furthermore the inferences that one can draw from the sample data are tempered by choices and decisions that were made earlier in the other phases. Thus the statement of the problem determines the types of hypotheses which, in turn, determine the kind of measures that are used to collect the data, as well as the type of statistical tests which are most appropriate. All of these then help to determine the inferences that can be drawn from the sample data analysis and then generalized to the population.

It is hoped that by illustrating the integrated, comprehensive approach to theory construction with a more complete example than those used in the previous chapters, the role of inference can more readily be understood in terms of the way it "puts everything together" and the way other activities impinge upon the kinds of inferences that can be subsequently generated. Furthermore, the example (summarizing as it does, many of the topics discussed throughout the earlier parts of the text) should serve as a focal point around which an assessment of the procedure as a
whole can be performed.

In order to avoid excessive repetition, in a text which is already quite long, certain activities within the comprehensive approach will not be explicitly included in the discussion. Where replication is necessary for continuity, it will be done so in as brief a manner as possible and the reader will be directed to earlier parts of the text for more elaboration.

Putting It All Together

The example which will be discussed in this topic consists of a re-examination of the causal model developed in Chapter Twelve on Path Analysis. It was based primarily on the work of Amitai Etzioni (1961, 1975) who uses a systems perspective within the "Social Factist" theoretical orientation as outlined in Chapter Five. The application of Etzioni's theory will be on a basically middle-range level and will utilize an explanatory analytical approach.

The problem for this re-examination involves an investigation of the possible moderating effects of state on the results obtained by the previous analysis. This example has been selected because it represents an alternative elaboration of the theoretical framework and data set that was utilized in the previous chapters of this text. The reader should note that the emphasis in discussing this example will be placed upon the inferential process and not
upon a detailed description of moderating analysis procedures. In keeping with the notion of presenting an overall procedure, a brief summary of the method will be included but the method remains secondary to the goal of illustrating the inferential process from the beginning to the end of the theory construction activity.

**Problem statement**

Most of the theoretical framework for the proposed model was presented in Chapter Twelve, and will not be repeated here. Instead the problem statement will confine itself to a delineation of the theoretical justifications for considering state as a possible moderator, however, the reader should be aware that the problem statement would require theoretical development of the model if it had not been done in the previous chapter.

The term, moderator, originally came from psychology. Its use suffers from the fact that it is never quite clear whether the moderating variable(s) is a predetermined variable or an 'intervening one.

---

^1 For a more extensive elaboration of moderators and the analysis techniques which will be utilized here, the reader is directed to see Specht and Warren (1975).
In sociology, the term has been expanded to include many of the things done in statistics to obtain control. Generally speaking then, if the theorist gets better prediction by examining subgroups based on different levels of a variable, that variable is considered to be a moderator.

**State as a Moderator**

The moderator variable to be examined here will be state. State may be thought to moderate (have a differential influence on) the results, considering the fact that each of the three states were originally chosen because they reflect (1) a different internal structure, (2) different areas of the country—north, south and midwest, and (3) different types of expected disasters. In Georgia, the organization is highly hierarchical and intimately related to the military; in Massachusetts, the organization is based on townships; whereas in Minnesota, it is on a regional basis. These different types of organizational structure could reflect different orientations to variables such as selectivity, socialization, communication, scope, pervasiveness, etc.
Each of the three states could also be expected to reflect different orientations strictly because they represent different areas of the country, each with somewhat varying historical and cultural backgrounds. Finally, each state reflects different concerns about particular types of disasters. Minnesota and Georgia are subjected to tornadoes, Massachusetts is not. Minnesota and Massachusetts have problems with heavy snows, blizzards, etc. All states to a greater or lesser degree are confronted by flooding. However, hurricanes are only characteristic of Georgia, not the other two states. In addition, homes in Massachusetts and Minnesota generally are more prone to have basements than homes in Georgia, thus in Georgia the problem of stocking and securing shelters is probably more relevant than in the other two states. In conclusion, the fact that these three states were originally chosen because they were expected to reflect some of these differences justifies the use of state in examining its possible moderating effect.

The causal model The following causal model was previously developed and tested in Chapter Twelve. I shall consider the initial model rather than the revised one because some of the relationships which were found to be insignificant in the earlier examination may prove to be significant when the possible moderating effect of state is considered.
In keeping with current practice, as discussed in Chapter Eight, the causal model above will take the place of the usual sets of general, empirical and statistical hypotheses. Note that the linkages are assumed to be causal in the Abell and Rhoads sense of contingent causality. One such condition discussed earlier concerned the fact that this model assumes a normative type organization. This re-examination will be concerned with establishing if and when state (i.e. a particular state) affects the relationships differentially. This would require a specification of those instances as part of the conditional statement.

Although I will not go into the adequacy of the concepts and measures included in the model above, it should be re-emphasized that inferences based on analysis results will
only be as viable as the concepts and measures are adequate. To the extent that the measures are invalid (do not correspond to their respective concepts and/or reality), unreliable (do not elicit the same responses from similar individuals) lack functional unity (do not approach homogeneity) and consist of a high degree of measurement error, the inferences reflecting data results will be invalid. Likewise the inferences will also be invalid to the extent that the concepts and linkages fail to correspond to their real world counterparts.

Examining the data

A cross-sectional design was originally used so that some of the measures depend upon recall data (e.g. socialization refers to initial socialization which for many director/coordinators involved a considerable time span). Care must be taken when interpreting cross-sectional data in terms of a logical time sequence because of the recall problem.

Preliminary assessment of the measures was discussed extensively in Chapter Nine. The results of that assessment will not be repeated here, except to reiterate that the measures met most of the criteria within acceptable boundaries and should not exert undue influence on the conclusions that are derived from the analysis.
Examining for the effect of a moderator

The following steps represent a summary of procedures which have been advocated by Specht and Warren (1975) for the examination of a moderating effect on a proposed relationship.

1. Provide a logic for examining the moderator variable. Why do you think grouping on the variable will cause differential results.

2. Examine preliminary statistics for possible effects
   (a) on reliabilities
   (b) on descriptive statistics
   (c) On the assumption of equal means and variances  [See Tables 13.1, 13.2, and 13.3 respectively]

3. Comparison of correlations for the total sample to those within the various groups  [See Tables 13.4(a-d) ]

4. Examine the regressions for the total and each group in terms of
   (a) the regression as a whole
   (b) the regression coefficients for each independent variable  
   [See Tables 13.5(a-b) and 13.6(a-b) ]

5. Examine various models using
   (a) a standard regression with common intercepts and overall slopes (Model 1)
   (b) separate slopes and intercepts (Model 2)
   (c) separate intercepts and pooled slopes (Model 3)
   (d) common intercept but different slopes (Model 4)
   (e) different intercepts (slopes not considered) (Model 5)

The first four of these steps are fairly descriptive in nature (i.e., they do not test relationships). They also require little or no further explanation as to "how" to do
them, since the techniques required by these steps have all been discussed in earlier chapters. Step 5, on the other hand, involves the testing of alternative (and more complicated) models. Unfortunately, in the process of explaining and illustrating the procedures involved in step 5, most readers would probably lose the intended connection between the example and the inferential process it is supposed to be illustrating. Therefore only those results from step 5 which are important to the generation of the final inferences will be included here.

I have already discussed the possible theoretical reasons (Step 1) why state may moderate or affect the relationships depicted in the model. If state does have an effect, it should be reflected by the sample data in terms of both measurement considerations (Step 2) and analysis results (Steps 3-5).

**Reliability** The effect of state as a moderating variable on reliability can be shown by examining a few of the variables in the model above.
Table 13.1: Differential effects of state on reliability

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Sample</th>
<th>Mass.</th>
<th>Georgia</th>
<th>Minn.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socialization</td>
<td>.6190</td>
<td>.6632</td>
<td>.5466</td>
<td>.6000</td>
</tr>
<tr>
<td>Scope</td>
<td>.6895</td>
<td>.6102</td>
<td>.6954</td>
<td>.7191</td>
</tr>
<tr>
<td>Role Performance</td>
<td>.6278</td>
<td>.5565</td>
<td>.6932</td>
<td>.6280</td>
</tr>
</tbody>
</table>

*Coefficient Alpha has been used to determine the reliability scores.

The results in Table 13.1 would seem to indicate that a scale may work better in one state than it does in the others. For instance, socialization seems to work best in Massachusetts, Scope in Minnesota, and role performance in Georgia.

If we examine the descriptive statistics by state in Table 13.2, we see that in all cases across the eight variables the highest mean value is located in Georgia. All (but tension whose variance is next to the lowest) on the other hand also exhibit the lowest variance in Georgia. We might expect then to find more differential results in Georgia than in the total or the other two states. [Note: Only the two end variables will be examined for the moderating effects of state, which means that communication and socialization will be dropped from further analysis since they are not thought to have direct effects on either dependent variable.]
Table 13.2: Descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>All States (N=240)</th>
<th>Mass. (N=84)</th>
<th>Georgia (N=80)</th>
<th>Minn. (N=76)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>S</td>
<td>X</td>
<td>S</td>
</tr>
<tr>
<td>Scope</td>
<td>14.58</td>
<td>4.68</td>
<td>15.08</td>
<td>4.28</td>
</tr>
<tr>
<td>Pervasiveness</td>
<td>5.80</td>
<td>3.70</td>
<td>5.56</td>
<td>3.93</td>
</tr>
<tr>
<td>Salience</td>
<td>215.45</td>
<td>28.19</td>
<td>210.15</td>
<td>29.29</td>
</tr>
<tr>
<td>Tension</td>
<td>12.49</td>
<td>4.39</td>
<td>12.49</td>
<td>4.09</td>
</tr>
<tr>
<td>Role Perf.</td>
<td>962.4</td>
<td>455.11</td>
<td>1031.69</td>
<td>425.02</td>
</tr>
<tr>
<td>Job Sat.</td>
<td>95.412</td>
<td>28.63</td>
<td>91.88</td>
<td>31.84</td>
</tr>
<tr>
<td>Communication</td>
<td>23.00</td>
<td>5.15</td>
<td>23.50</td>
<td>5.41</td>
</tr>
<tr>
<td>Socialization</td>
<td>91.05</td>
<td>27.02</td>
<td>96.43</td>
<td>23.02</td>
</tr>
</tbody>
</table>
Table 13.3: Tests on means and variances

<table>
<thead>
<tr>
<th>Variable</th>
<th>F-test on Means</th>
<th>Probability Level</th>
<th>Bartlett-Box F</th>
<th>Probability Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>VAR 083 Pervasiveness</td>
<td>1.334</td>
<td>.264</td>
<td>.504</td>
<td>.604</td>
</tr>
<tr>
<td>VAR 034 Tension</td>
<td>1.611</td>
<td>.200</td>
<td>.953</td>
<td>.367</td>
</tr>
<tr>
<td>VAR 102 Salience</td>
<td>3.493*</td>
<td>.031</td>
<td>.334</td>
<td>.679</td>
</tr>
<tr>
<td>VAR 110 Scope</td>
<td>12.42**</td>
<td>.000</td>
<td>1.458</td>
<td>.225</td>
</tr>
<tr>
<td>VAR Role Performance</td>
<td>19.15**</td>
<td>.000</td>
<td>1.016</td>
<td>.363</td>
</tr>
<tr>
<td>VAR 130 Job Satisfaction</td>
<td>2.527</td>
<td>.080</td>
<td>2.175</td>
<td>.111</td>
</tr>
</tbody>
</table>
There is a statistically significant difference as shown in Table 13.3 between at least one state and the others on the mean values of salience, scope, and role performance. However, none of the variances proved to be significantly different on these variables so homogeneity of variance is a reasonable assumption across the three states for these variables, even though equal means is not.

If we examine the four correlation matrices [Tables 13.4(a) - (d)] we see that Massachusetts has three correlations which were significant in the total sample but not significant in Massachusetts. These are pervasiveness with salience, scope with tension, and tension with role performance. Two correlations which were not significant in the total sample are significant in Georgia. These are the correlation between pervasiveness and salience and the correlation between tension and job satisfaction.
Table 13.4(a): Correlations - total sample

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05.  
**Significant at .01.  

N = 240
Table 13.4(b): Correlations - Massachusetts

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>1.000</td>
<td>.459**</td>
<td>.329**</td>
<td>.128</td>
<td>.390**</td>
<td>.582**</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pervasiveness</td>
<td>1.000</td>
<td>.172</td>
<td>.050</td>
<td>.202*</td>
<td>.488**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salience</td>
<td>1.000</td>
<td>.258**</td>
<td>-.009</td>
<td>.240*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tension</td>
<td>1.000</td>
<td>-.142</td>
<td>-.049</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Satisfaction</td>
<td>1.000</td>
<td>.388**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Role Performance</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05.  N=84

**Significant at .01.
Table 13.4(c): Correlations - Georgia

<table>
<thead>
<tr>
<th></th>
<th>X 3</th>
<th>X 4</th>
<th>X 5</th>
<th>X 6</th>
<th>X 7</th>
<th>X 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>X Scope</td>
<td>1.000</td>
<td>.438**</td>
<td>.404**</td>
<td>.532**</td>
<td>.406**</td>
<td>.569**</td>
</tr>
<tr>
<td>X Pervasiveness</td>
<td>1.000</td>
<td>.395**</td>
<td>.217*</td>
<td>.171</td>
<td>.464**</td>
<td></td>
</tr>
<tr>
<td>X Salience</td>
<td>1.000</td>
<td>.310**</td>
<td>.113</td>
<td>.424**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Tension</td>
<td>1.000</td>
<td>.269**</td>
<td>.492**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Job Satisfaction</td>
<td>1.000</td>
<td>.330**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>X Role Performance</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Significant at .05.
**Significant at .01.

N=80
Table 13.4(d): Correlations - Minnesota

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>X Scope</th>
<th>X Pervasiveness</th>
<th>X Salience</th>
<th>X Tension</th>
<th>X Job Satisfaction</th>
<th>X Role Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1.000</td>
<td>.202*</td>
<td>.331**</td>
<td>.385**</td>
<td>.303**</td>
<td>.619**</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>1.000</td>
<td>.255*</td>
<td>-.097</td>
<td>-.044</td>
<td>.426**</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>1.000</td>
<td>.225*</td>
<td>.193*</td>
<td>.404**</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.164</td>
<td>.218*</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
<td>.298**</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>

*Significant at .05.  
**Significant at .01.  
N=76
In addition, one correlation (between pervasiveness and job satisfaction) is no longer significant. This same correlation is the only correlation in Minnesota which was different (as far as significance level is concerned) from the total sample. Tension is a more significant variable in Georgia than in any of the other states or the total sample. These seemed to be the major overall differences, between the four correlation matrices. But what do these differences mean substantively?

If we put this discussion in terms of the model in Figure 13.3, three of the hypothesized relations would seem to be in doubt, depending upon the state examined. All of the problem relationships involve job satisfaction. Generally speaking, the arrows between salience and job satisfaction and between tension and job satisfaction failed to receive support from the correlational analysis. A weak relationship was found for salience and job satisfaction, but only in Minnesota. Tension and job satisfaction also found support in only one state, Georgia, where it was of medium rather than weak strength. The third hypothesized relationship to find only partial support was that between pervasiveness and job satisfaction. No more than a weak relationship was found in the total and in one state, Massachusetts. On the basis of the correlations, one might expect job satisfaction to be only related to scope, except
perhaps in Georgia where tension might have more of an input. In the next section I will examine these relationships more specifically in terms of the effect of the independent variables on the two dependent variables rather than mere relationships.

**Multiple regression**

With the use of multiple regression, the relationships that were hypothesized in Figure 13.3 and examined in terms of strength and direction (using correlation analysis) will now be decomposed into the degree of linear dependence of the dependent variable upon one or more independent variables.

Just as in Table 13.2, descriptive statistics, we can see from Table 13.5 (a) that prediction of role performance by scope and pervasiveness using a standard regression seems to work better in Minnesota than in the other two states, ($R^2$ of .48 in Minnesota as compared to .38 in Georgia and .40 in Massachusetts). The results of both the regressions in Table 13.5(a) and the previous correlations would tend to indicate that state is having some effect with respect to the prediction of role performance. The difference in standard errors between the states could, at least in part, translate this effect into the presence of specification error.

In Table 13.5 (b) the situation is somewhat different. Here the dependent variable is job satisfaction. The regression is significant in the total sample but
Table 13.5(a): Regression of Role Performance on Scope Pervasiveness

<table>
<thead>
<tr>
<th>State</th>
<th>SS REGR</th>
<th>SS RESID</th>
<th>SS TOTAL</th>
<th>R²</th>
<th>F</th>
<th>ST. ERR.</th>
</tr>
</thead>
<tbody>
<tr>
<td>All States (N=240)</td>
<td>22,612,675.87</td>
<td>26,690,637.92</td>
<td>49,503,213.79</td>
<td>.461</td>
<td>101.28**</td>
<td>335.59</td>
</tr>
<tr>
<td>Mass. (N=84)</td>
<td>6,004,609.93</td>
<td>3,002,304.97</td>
<td>9,006,914.90</td>
<td>.40</td>
<td>27.05**</td>
<td>333.13</td>
</tr>
<tr>
<td>Georgia (n=80)</td>
<td>4,545,605.92</td>
<td>7,384,909.76</td>
<td>11,930,515.69</td>
<td>.38</td>
<td>23.70**</td>
<td>309.69</td>
</tr>
<tr>
<td>Minn. (n=76)</td>
<td>7,488,742.12</td>
<td>8,204,290.83</td>
<td>15,693,032.94</td>
<td>.48</td>
<td>33.32**</td>
<td>335.24</td>
</tr>
</tbody>
</table>

*Significant at .05. **Significant at .01.

Table 13.5(b): Regression of Job Satisfaction on Scope Pervasiveness

<table>
<thead>
<tr>
<th>State</th>
<th>SS REGR</th>
<th>SS RESID</th>
<th>SS TOTAL</th>
<th>R²</th>
<th>F</th>
<th>ST. ERR.</th>
</tr>
</thead>
<tbody>
<tr>
<td>All States (n=240)</td>
<td>26,792.28</td>
<td>169,109.88</td>
<td>195,902.16</td>
<td>.14</td>
<td>9.31*</td>
<td>26.826</td>
</tr>
<tr>
<td>Mass. (n=84)</td>
<td>16,947.87</td>
<td>67,192.94</td>
<td>84,140.81</td>
<td>.20</td>
<td>4.98*</td>
<td>29.164</td>
</tr>
<tr>
<td>Georgia (n=80)</td>
<td>8,765.49</td>
<td>42,003.31</td>
<td>50,768.80</td>
<td>.17</td>
<td>3.91</td>
<td>23.665</td>
</tr>
<tr>
<td>Minn. (n=76)</td>
<td>6,756.38</td>
<td>50,144.82</td>
<td>56,901.20</td>
<td>.12</td>
<td>2.39</td>
<td>26.58</td>
</tr>
</tbody>
</table>

*Significant at .05. **Significant at .01.
insignificant in all but one state, Massachusetts (although even here it is close since it takes an F of 4.33 to be significant). Since it was shown in Table 13.3 that the variances were not significantly different we can surmise that the total sample might be a better predictor of job satisfaction using this set of variables than any one state.

The discussion above was in terms of the regressions as a whole. I shall now turn to the individual regression coefficients to see what insight can be brought to bear in terms of state on the individual variables. The regression coefficient and standardized regression coefficient [shown in Table 13.6(a)] for scope is more nearly approximated by the Minnesota subsample than by the others, although, the same is not true for pervasiveness. The fact that scope contributes more to the prediction of role performance than pervasiveness may explain why Minnesota comes off better in the total equation than the other two states even though its approximation of pervasiveness is less than that in the other two states. It might also indicate, as did the standard errors above, that some equations may need fewer variables than others (i.e., the presence of some specification error).

As explained earlier, the results of Step 5 will not be reported here, except to say that the most efficient role performance was found to be Model 3 (separate intercepts and pooled slopes). This model allows for the greatest increase
Table 13.6(a): Regression of Role Performance on Scope and Pervasiveness

<table>
<thead>
<tr>
<th>State</th>
<th>Variables</th>
<th>B</th>
<th>St. Err. B</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>All States</td>
<td>Scope</td>
<td>52.99</td>
<td>.545</td>
<td>4.99</td>
</tr>
<tr>
<td>(N=240)</td>
<td>Perv.</td>
<td>30.93</td>
<td>.252</td>
<td>6.30</td>
</tr>
<tr>
<td>(N=84)</td>
<td>Perv.</td>
<td>30.35</td>
<td>.280</td>
<td>10.48</td>
</tr>
<tr>
<td>Georgia</td>
<td>Scope</td>
<td>42.37</td>
<td>.452</td>
<td>9.34</td>
</tr>
<tr>
<td>(N=80)</td>
<td>Perv.</td>
<td>29.06</td>
<td>.266</td>
<td>10.88</td>
</tr>
<tr>
<td>Minn.</td>
<td>Scope</td>
<td>51.08</td>
<td>.555</td>
<td>7.95</td>
</tr>
<tr>
<td>(N=76)</td>
<td>Perv.</td>
<td>40.12</td>
<td>.314</td>
<td>11.05</td>
</tr>
</tbody>
</table>

*Significant at .05.
**Significant at .01.
Table 13.6(b): Regression of Job Satisfaction on Scope

<table>
<thead>
<tr>
<th>State</th>
<th>Variable</th>
<th>B</th>
<th>St. ERR. B</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>All States (N=240)</td>
<td>Scope</td>
<td>2.384</td>
<td>0.389</td>
<td>.438</td>
<td>29.683**</td>
</tr>
<tr>
<td></td>
<td>Perv.</td>
<td>-0.025</td>
<td>-0.003</td>
<td>.514</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Sal.</td>
<td>-0.014</td>
<td>-0.014</td>
<td>.068</td>
<td>.044</td>
</tr>
<tr>
<td></td>
<td>Ten.</td>
<td>-0.297</td>
<td>-0.045</td>
<td>.433</td>
<td>.470</td>
</tr>
<tr>
<td>Mass. (N=84)</td>
<td>Scope</td>
<td>3.241</td>
<td>.436</td>
<td>.880</td>
<td>13.57**</td>
</tr>
<tr>
<td></td>
<td>Perv.</td>
<td>.246</td>
<td>.030</td>
<td>.918</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>Sal.</td>
<td>-0.123</td>
<td>-0.113</td>
<td>.119</td>
<td>1.075</td>
</tr>
<tr>
<td></td>
<td>Ten.</td>
<td>-1.323</td>
<td>-0.170</td>
<td>.810</td>
<td>2.665</td>
</tr>
<tr>
<td>Georgia (N=80)</td>
<td>Scope</td>
<td>2.353</td>
<td>.385</td>
<td>.838</td>
<td>7.886**</td>
</tr>
<tr>
<td></td>
<td>Perv.</td>
<td>0.094</td>
<td>.013</td>
<td>.864</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>Sal.</td>
<td>-0.069</td>
<td>-0.074</td>
<td>.113</td>
<td>0.376</td>
</tr>
<tr>
<td></td>
<td>Ten.</td>
<td>0.498</td>
<td>0.084</td>
<td>.742</td>
<td>0.452</td>
</tr>
<tr>
<td>Minn. (N=76)</td>
<td>Scope</td>
<td>1.562</td>
<td>0.282</td>
<td>.709</td>
<td>4.856*</td>
</tr>
<tr>
<td></td>
<td>Perv.</td>
<td>-1.024</td>
<td>-0.133</td>
<td>.918</td>
<td>1.243</td>
</tr>
<tr>
<td></td>
<td>Sal.</td>
<td>0.132</td>
<td>0.131</td>
<td>.123</td>
<td>1.145</td>
</tr>
<tr>
<td></td>
<td>Ten.</td>
<td>0.075</td>
<td>0.013</td>
<td>.721</td>
<td>0.011</td>
</tr>
</tbody>
</table>

*Significant at .05.
**Significant at .01.
in $R^2$ using the fewest variables (given this data set and sample). In the case of job satisfaction, a modified Model 4 would seem to be the most efficient, since some (but not all) of the individual variable slopes are different but intercepts do not seem to be different.

**Summarizing the results of the moderator analysis**

The state in which a director/coordinator works was shown to have a moderating effect upon the prediction of his/her role performance and job satisfaction. The organizational set up within the three states included in the sample are thought to represent the three types of organizations which characterize agencies within the civil preparedness organization. However, the three states also have other differentiating factors which could be operating either alone or in combination. For instance, is it the type of organization or the type of disasters that makes the difference? At this point no definitive conclusions can be drawn with respect to these questions. However, this analysis has demonstrated that state does have an impact upon the prediction of role performance and job satisfaction, albeit a different impact upon the former than the latter. It was further shown that this impact could be translated into specifying separate intercepts (along with pooled slopes) for role performance; while it meant specifying separate slopes for certain (not all) variables and common
This example has attempted to bring together and to summarize most of the important topics which were discussed throughout the text. Emphasis was placed here, as it was in the body of the text on the necessity for integration and the interdependent nature of the various activities. A number of activities in the theoretical formulation and measurement phases have an effect on the kinds of inferences that can be made after data analysis. Some of these have been touched upon in the discussion above; all have been discussed previously in the text. I have listed some of the more prominent activities below.

1. Theoretical orientation, values and etc.
2. Faulty concepts and/or measures
3. Choosing the wrong design
4. Not meeting statistical assumptions
5. Moderator variables
6. Specification error
7. Sampling error
8. Measurement error

If one takes an integrated approach to theory construction, as this method suggests, the implications of prior activities will more readily be made apparent.

This concludes the discussion on the integrated, comprehensive approach to theory construction that was proposed in Chapter Three. The approach was developed as a means of closing the gap between theory and methods. Before bringing closure on the text, however, the method should be assessed to see whether the objectives for this method, as
laid out in the introduction, have been met.

Assessment of the Comprehensive Approach to Theory Construction

Assessment of the proposed technique will be made in terms of the degree to which the objectives discussed in the introduction were met. These will also include the criteria in Chapter Two that were used to assess a number of current theory construction texts. The criteria can be summarized as follows:

1. Review current status of theory construction to develop a *working definition* of comprehensive theory construction
2. Scope of material should cover all four phases of theory construction
3. Stress should be placed on *interdependence and integration*
4. Discussion supplemented with empirical example in terms of the process as a whole as well as individual activities.

**Development of a working definition.**

There is a vast amount of material that has been written on theoretical formulation, measurement and analysis activities. It was necessary to discover some method for reducing that material down to a manageable number of texts that could be reviewed. The survey offered such a method, since it gave an indication, at least, of the texts that were currently being used in teaching theory construction. This reduced the total number to seven of the most widely cited texts of which one (Kuhn) was eliminated since it dealt with
theoretical orientations or paradigms rather than theory construction, per se. In its place, two other books, (Abell, 1971 and Hage 1972) were examined for the scope of the material they covered.

These texts were examined first in terms of the authors’ stated goals for their texts, since all had intended their texts as guides to the construction of theory. It was found that each of the authors had affirmed either explicitly or implicitly, the need for integration of theory and methods. A synthesis of their variously stated goals led to the following working definition of sociological theory construction.

Comprehensive Sociological Theory

construction is a process whereby both theoretical and methodological activities interact to produce a generalized conception of reality, commonly referred to as a theory or model. As such the end product (theory) is a resultant of constant interaction between theory formulation, measurement, analysis and reformulation activities.

Scope of material

Having arrived at a working definition which was derived from stated goals by the authors of theory construction texts, the next step consisted of a content analysis of those books to see whether any had reached their goals of
developing such a comprehensive approach.

This examination was dealt with in terms of both scope and integration (criteria numbers 2 and 3 above). In order to compare the present approach to the others, Table 2.3 from Chapter Two has been reproduced on the next page as Table 13.11 with the addition of this text. The reader should note that the present text has covered all the items in the list, although the coverage of measurement error and moderators was not extensive.

Interdependence and integration of material

The numbers in parentheses in Table 13.7 refer to the chapters where these activities, or their implications, are discussed. These were included as a means of graphically illustrating that activities in one phase have implications for activities which come after them and that this has been built into the discussion of the method both implicitly (in the general discussion of an activity) and explicitly (many topics are followed by an implication section).

A recently published text by Nan Lin, The Foundations of Social Research (1976), approaches the stated goal of this text in that it incorporates theoretical formulation, measurement and analysis procedures. The major difference between it and this text resides in the emphasis that the present text has placed upon the interdependence and integration of the various activities. Another difference is
Table 13.7: Material covered by current theory construction
texts

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Theoretical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chapt. 5, 9, 13</td>
</tr>
<tr>
<td>Orientation</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theoretical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chapt. 6, 7</td>
</tr>
<tr>
<td>Background</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>8, 9, 10, 13</td>
</tr>
<tr>
<td>Unit of Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chapt. 6, 7</td>
</tr>
<tr>
<td>Population and</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6, 9, 13</td>
</tr>
<tr>
<td>Concepts</td>
<td>***</td>
<td>*</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>chapt. 7, 9, 10, 13</td>
</tr>
<tr>
<td>Linkages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chapt. 8, 10, 13</td>
</tr>
<tr>
<td>Propositions</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>chapt. 6, 10</td>
</tr>
<tr>
<td>(arrangement of)</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>chapt. 6, 13</td>
</tr>
<tr>
<td>Propositions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Empirical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>chapt. 7, 9, 13</td>
</tr>
<tr>
<td>Indicators</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic</td>
<td>*</td>
<td>**</td>
<td>***</td>
<td>****</td>
<td>Chapt.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>---</td>
<td>----</td>
<td>-----</td>
<td>------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8, 9, 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Design</td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td>9, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Gathering Technique</td>
<td></td>
<td>***</td>
<td></td>
<td></td>
<td>9, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preliminary Analysis</td>
<td></td>
<td>**</td>
<td></td>
<td></td>
<td>9, 10, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis Procedures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10, 11, 12, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement Error</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9, 10, 11, 12, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8, 9, 10, 11, 12, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inference</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>****</td>
<td>9, 10, 11, 12, 13</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the use of a single data set as the basis for most of the illustrations, which should enable the reader to gain a feeling of continuity (i.e., how activities in one chapter relate to those in other chapters).

**Empirical application**

The discussion of the various activities has been in terms of choices or alternatives, each narrowing the scope of future choices. The use of the single data set graphically demonstrates this aspect of the method. For example, a variety of theoretical orientations, study designs, data collection techniques and etc. are discussed in general terms. Then a specific choice is illustrated with the empirical example that was dictated by previous choices.

If the method can effectively be used in the integrated fashion as discussed in the text then one should be able to demonstrate this with an empirical example. In Kaplan's terminology this means that the discussion of the method itself is an example of "reconstructed logic", whereas the discussion of the example used throughout the text to illustrate the various activities within the method, can be viewed more as an example of "logic-in-use."
Utility of the comprehensive approach

This text has been written with several target audiences (students and researchers) in mind. The first, and most obviously in need of such a text, are beginning sociologists. The method begins, as Romans suggested, by defining explicitly what is meant by the term "theory". (After, all how can we expect students to construct theory if they don't know what it is?) Discussion of the method proceeded in a step-by-step fashion, generally in terms of alternatives that the theorist must choose between, followed by the implications his/her choice may have on future choices. This format should be useful to both beginning students (introducing them not only to the activities but their implications as well) and to the researcher (offers a step-by-step guide which can be followed in developing and carrying out a research project).

Most of the discussion is directed at the beginning student and takes place on a fairly elementary level. However, as one moves through the analysis chapters the depth of the topics tend to increase. Thus for example, Chapter Ten introduces the student to the bivariate forms of correlation, regression and analysis of variance, tying them first into types of hypotheses and inferences and then into the statistical procedures and assumptions. Even the chapter on multiple regression begins on a fairly introductory level,
one that is primarily concerned with prediction. Some of the issues dealing with explanation, at the end of the chapter, require more statistical background. Even so, these need not prevent the beginning student from using a multiple regression procedure as outlined in the earlier sections.

On the other hand as the student's knowledge matures, the text will continue to be useful, both in terms of future learning and as a guide in the conduct of research and theory construction. Evaluation of other research should also be facilitated since the text offers a quick guide (Table 2.3) to the topics that should receive consideration. Furthermore, the reader should find that this text is compatible with disciplines outside sociology such as statistics, philosophy, and computer analysis programs since each of these has entered explicitly into either the general discussion or the empirical example. In fact, references to sources from these disciplines are made throughout the text.

Turner (1974) has suggested that a theory should be assessed in terms of what the author states he/she is going to do rather than what others think he/she should do. This has been the reasoning that the present assessment has followed with respect to the comprehensive approach to theory construction that was outlined in the text.

The current gap between theoretical and methodological activities has been defined by many as a barrier to the
construction of sound sociological theory. This has been explicitly recognized by most authors of prevailing texts on sociological theory construction, even though they have, for the most part, failed to carry through with that goal in their texts. This text proposed to fill that gap by outlining comprehensive approach to theory construction that integrates theoretical, measurement, analysis, and inferential activities. Furthermore, it specified a number of objectives which were to be met by that method. This section has used those objectives as criteria for assessing the approach. On the basis of the discussion above, it would seem fair to conclude that the criteria have been met.

Summary

The final phase of the comprehensive approach to theory construction is the inferential phase. As with the other phases however, it is not something that is merely tacked on to the end of the analysis phase. The theorist must begin to make inferences throughout the measurement and analysis phases. Some would even argue that inference is involved in selecting a measure for a concept (i.e. when a measure is substituted for a concept the inference is made that the measure does in fact represent the concept and its real world counterpart).

Furthermore, the act of making inferences cannot be fully understood unless it is placed within the context of an
integrated approach to theory construction. Many of the decisions which are made in the theoretical formulation and measurement phases have an effect on the kinds of inferences that can be made. Some of these include values, orientations (a theorist from one orientation may not interpret the meaning of "facts" in the same way as one from another orientation), faulty concepts and/or measures, choice of design, not meeting assumptions, presence of moderators, sampling, specification and measurement errors. To the extent that these enter into the construction process, the inferences that can be drawn regarding the applicability of the proposed theory to that population must be modified, or at least tempered. An attempt was made to demonstrate the impact that some of these activities have on inference by working through the process with an empirical example.

Having completed the discussion and illustration of the method, the second part of this chapter turned to an assessment of the method, using criteria suggested in the first and second chapters of this dissertation. After examining the working definition of theory construction, the scope of material, the integration of that material, and the proposed utility, the conclusion was drawn that the method as discussed in this text did indeed meet the specified criteria.
CHAPTER FOURTEEN:
SUMMARY AND CONCLUSIONS

The problem for this dissertation revolved around the need to develop an integrated, comprehensive approach to theory construction. The dissertation itself was divided into two main parts. The first part was concerned with establishing the need for a more integrated and comprehensive approach than is currently available. The second part then was devoted to a detailed description of a proposed alternative. A survey of theory construction courses revealed a wide diversity in terms of whether theory construction was viewed as a theoretical activity, a methodological one, or something that should be studied in its own right. Taken as a whole, the material that was required in the courses that dealt with theory construction was massive. However, a number of textbooks were repeatedly listed as required texts. Since it was impossible to do a content analysis on all the material that was listed, it was decided that these frequently cited texts would serve as the basis for establishing a working definition of theory construction. Consequently, by examining the prefatory statements of the authors of six frequently cited texts, a definition calling for an integrated and comprehensive (in terms of involving formulation, measurement and statistical activities) approach to theory construction was developed.
Having arrived at a definition of theory construction by reviewing statements of these authors, the next step was to assess the degree to which they were able to produce such a technique. This assessment took two forms: the scope of the material covered, and the amount of integration of activities that was emphasized.

In order to assess the scope of each text a list of prominent theoretical formulation, measurement, and analysis activities was drawn up using a variety of sources external to those being examined. The list contained seventeen items: 6 to 9 of which dealt with formulation activities; 4-6 with measurement, and 2 to 4 with analysis, (depending on who is doing the classification). Most of the texts that were reviewed concentrated on theoretical formulation activities. None were found to include more than ten of the items. At the conclusion of the above examination, it was evident therefore, that the problem gap between theory and methods still remained.

The problem for this dissertation, then, as laid out in part one, was to develop an integrated and comprehensive approach to theory construction. This approach was to include activities from each of the four phases (theoretical formulation, measurement, analysis, and inference) in keeping with the working definition of a comprehensive approach to theory construction that was developed in Chapter One. In
addition to the scope of material, the interdependence of the various activities was to be demonstrated in order to emphasize the necessity for applying an integrated approach to theory construction. A final goal was to include material at two levels of understanding, an elementary one for the beginning student and more complicated techniques which could be used as the student gains expertise.

An approach was proposed in Chapter Three which purported to meet these criteria. A brief overview of the approach was presented, along with some of the thought processes that were involved in the selection of activities and terminology for the approach. This fulfilled the first criterion (i.e. developing an approach consistent with the working definition). Demonstration of the scope, integration and depth required by the other criteria were undertaken in Part Two. The following summarizes how these were met.

Part Two was divided into four sections, each concerned with a phase of the comprehensive approach to theory construction. The first phase to be discussed was theoretical formulation. It was divided into five chapters. Chapter Four reviewed definitions of sociological theory as a prelude to the development of a comprehensive definition and in keeping with Homans' warning that students must be told what theory is before they can be expected to develop it.
The next two chapters dealt with decisions that narrow the scope of interest to a specific problem or area of concern. Thus Chapter Five was concerned with stating one's assumptions, the most important of which was said to be the theorists' theoretical orientation since this often influences the theorists' choice of concepts, types of measures, method of data gathering, types of analysis and the inferences that are drawn. The issues considered in Chapter Five tended to set general boundaries, those in Chapter Six, were far more specific. It was concerned with outlining the specific domain of interest. Activities relevant to this goal included the problem statement, designating the unit of analysis, population and sample.

A substantive empirical example was introduced at this point. The theoretical framework, unit of analysis, population and sample for that example was discussed. This example was intended for use throughout the remainder of the text to illustrate the various points as they were being made. It was hoped that this would give the reader a feel for the method in terms of both "reconstructed logic" and "logic-in-use". As Kaplan (1964) so rightly observed, it is one thing to talk about a method in the abstract but another to apply that method to a real world situation.

The last two chapters of the theoretical formulation section dealt with the parts of propositions and hypotheses.
as well as the propositions, hypotheses, and their systemmatization. Chapter Seven was concerned with specifying concepts and their measures so that they were isomorphic with each other and with their real world counterpart. Furthermore, considerable attention was given to the specification in terms that would clarify the exact nature of the theorist's meanings as opposed to the vague or ambiguous approach that has often been followed in the past.

The emphasis in Chapter Eight was placed on specifying the exact nature of linkages (relationships) between concepts (and measures). Some of the issues which were dealt with here under linkages are often discussed in connection with propositions, but it was felt that more attention needs to be placed on specifying the exact nature of the linkage, so instead of talking about qualitative versus quantitative propositions, the discussion was directed toward the linkage (i.e. qualitative versus quantitative linkages). A variety of approaches to systematization were discussed including causal modeling, the approach taken in the substantive example.

The second phase to be discussed was measurement. Due to the excessive length of the text, it was decided that only the most important bridging aspects of measurement would be discussed here. The use of previously collected data (and the probability that it's use will increase in the future as
costs rise and funds decrease) requires that only those topics which have direct relevance on the analysis and inferential phases be discussed. Therefore a limited number of measurement issues were chosen. These included research design, study design, sampling, data collection techniques, and scaling. Considerable attention was also given to various preliminary assessment techniques for such things as reliability validity, and measurement error.

The third phase of the comprehensive approach to theory construction to be discussed was the analysis phase. The discussion of this phase was divided into three chapters. In Chapter Ten the emphasis was on relating different types of linkages within hypotheses to different types of statistical techniques. Only the bivariate forms of these were discussed in this chapter. Correlation, simple regression, and analysis of variance were chosen for discussion here because they represent the most popular two-variable parametric techniques in current usage.

Multiple regression was the topic of Chapter Eleven. The discussion was begun on a rather elementary level, dealing with the predictive aspects of multiple regression. In the second part of the chapter more complicated procedures dealing with the explanatory side of multiple regression were introduced. This second portion was included for those with more knowledge of statistical procedures than is required to
understand the first part of the chapter.

The last chapter in the analysis phase was on path analysis. The assumptions and a step-by-step procedure were discussed for the beginning student. A section on the Alwin and Hauser approach to direct and indirect effects was included for those with a more sophisticated knowledge of statistics.

Inference was the last of the four phases to be discussed. Since it represents the phase where everything is put together (i.e. conclusions are drawn and generalizations are made from the sample results to the propositions and theories which relate to the population), it was decided that the process of drawing inferences would best be illustrated with an empirical example. In this way, those activities in the formulation and measurement phases which have an influence on inference could also be demonstrated.

Having demonstrated the method as a whole, the discussion in this chapter was turned to assessing the method that was utilized in that example and throughout the second part of this text. This assessment was done in terms of the criteria used to assess the current text books in theory construction (scope and integration) as well as whether or not the objectives laid out in the introduction were met. The conclusion of this assessment was that these criteria and objectives were in fact met.
This concludes the summary of the proposed method. The enormity of the task of compiling, integrating, and writing a "comprehensive" procedure for constructing theory has been acknowledged by many of the authors examined in the first part of this dissertation. Certainly the size of the present volume would tend to substantiate that position, especially since, even here, it was necessary to omit activities which could have been included. Some of these will be mentioned briefly in the next section.

Suggestions for Future Research

My interest in this section will be to outline areas and/or activities which either need to be refined or which merit further consideration than the limitations of the present thesis allowed. I shall deal with the latter category first.

Topics deserving further consideration

A number of topics require greater consideration than was possible here. These topics were in some cases deemed to be either less essential or more sophisticated than the ones that were included. For, instance, the topic of measurement error was discussed rather briefly in various chapters throughout the text. Furthermore, the discussion generally related more to the implications of measurement error than how one would adjust for it. As the expertise of the reader
increases, however, it would be well for him/her to pursue the suggested readings on the errors-in-variables adjustment approach, cited in Chapter Eleven.

The building of composites was another topic that was only lightly touched on in the text. A more thorough treatment would be required if the reader wanted to build his/her own measures. It was felt that this topic was perhaps less essential to the overall goal of the text, since the assumption was made in Chapter Nine that most beginning students would probably rely either on someone else's scale or on previously collected data. Where this assumption is not valid, the reader is strongly advised to do the recommended outside reading mentioned in Chapter Nine.

The topics discussed in this section have received considerable attention elsewhere and generally require no more than additional reading. The next two topics, however, include activities that are in need of either refinement or more extensive development.

Activities which need refinement

Throughout the text, it was often necessary to use more than one term, such as indicator (operational measure), to designate something in order to communicate to as wide an audience as possible. This was necessary because there is no real consistency in either the terminology used to designate a particular meaning (e.g., concept, unit, variable are all
used by different people to mean the same thing) or in the way a particular term is used, (e.g. Dubin uses the term unit to denote a concept whereas most use it to denote what a concept refers to, as in unit of analysis). Certainly at least some systematization of terminology is needed not only for individuals but for sociology as a whole, in order to facilitate communications with one another and to "take stock" of where we are as a discipline.

Another area which needs some refinement is the assessment of assumptions. Sociologists are, in general, making more of an attempt to examine the degree to which various assumptions are being met by their theory and data, at least more attention is being paid to listing one's assumptions. What seems to be lacking is an explication of "how" one can test the various assumptions. An attempt has been made at various points throughout this text to either discuss such procedures or direct the reader to others who have done so. However, it would perhaps be beneficial, if this were done in a more systematic fashion, bringing together the various sources, and illustrating techniques with sample data, so that perhaps the reader could see graphically the difference between the ideal and the real and/or between a reasonable "fit" versus violation of the assumption.
Topics in need of further development

The topics which were discussed above were all ones that required either further reading or systematization. Those which will be included here, are in need of considerably more work. These consist of the assessment of direct and indirect effects and the use of multiple indicators. Both of these topics were discussed in the text, in terms of their present level of development, and their potential benefit, but only after certain problems have been worked out.

The Alwin and Hauser approach to decomposition of direct and indirect effects has been applauded as an improvement over the traditional approach suggested by Duncan (1965). Even Duncan has stated that he prefers their procedure over his own. However, their approach is only applicable to fully (not including the input variables) recursive models, and even then, one often gets results which are impossible to interpret. This approach may well be an improvement over Duncan's but it is far from being the "definitive" approach to the assessment of direct and indirect effects.

Recent work in the area of social indicators has revived interest in the use of multiple indicators. Most of the work on multiple indicators was done several years ago, and then apparently "dropped." When the concepts are fairly homogenous, there seems to be little or nothing to gain from the substitution of the more complicated procedure of
multiple indicators instead of simple composites. However, as one moves further away from homogeneous concepts (as is often the case with social indicators) the utility of the multiple indicator approach increases. As was the case with the procedure above, the present approach to multiple indicators is not without problems. Burke and Schuessler have said that one can use the multiple partial correlation as a path coefficient. However, they are not strictly comparable. They are related to each other but the precise relationship has not yet been worked out. To this point we can say no more than that it is somewhat analogous to the relationship between partial correlation and standardized partial regression (path) coefficients (using single or composite indicators).

Both of these topics offer potential improvements in interpretation and understanding, but considerable development is needed before that potential can be fully realized. There are of course other topics which could have been included here but these seemed to be two of the areas which are in need of immediate concentration, since these procedures are being currently utilized, regardless of their drawbacks. This occurs because even though they are not without problems, they do offer improvements over past techniques.
This text has attempted to demonstrate a need for a more comprehensive and integrated approach to theory construction (Part I). Once that need was established, an alternative approach which purported to fill that need was suggested (Chapter 3) and discussed in detail (Part II). Although it came closer to meeting the two criteria of scope and integration than other current texts, (as shown in the comparison in Chapter 13), this section has attempted to draw attention to various topics which would benefit from more elaboration, refinement or development.

These are included here to stimulate readers to go beyond the present level of development of this text. In doing so it is hoped that the procedure for developing integrated and comprehensive sociological theory can become even more systematic and refined, and with perhaps less effort than was required for this attempt.
Abell, Peter

Allan, G. J. B.

Althauser, R. P.

Alwin, D. F. and R. M. Hauser

Babbie, E. R.

Baker, B. O., C. Hardyck and L. Petrinovich

Bailey, K. D.

Bancroft, T. A.

Bassart, W. H.
1968 "Kant's doctrine of the reciprocity of freedom and reason." International Philosophical Quarterly 8:334-55.

Becker, H. S.
Becker, H. and A. Boskoff  

Becker, H. and B. Geer  

Bell, C. and H. Newby  

Berger, J., B. Cohen, J. Snell and M. Zelditch  

Berger, P. L. and T. Luckman  

Bertrand, A.  

Blalock, H. M., Jr.  

Blalock, H. M., Jr.  

Blalock, H. M., Jr.  

Blalock, H. M., Jr.  

Blalock, H. M., Jr.  
Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr.

Blalock, H. M., Jr. and Ann B. Blalock

Blalock, H. M., Jr., C. S. Wells and L. F. Carter

Blasi, A.

Blumer, Herbert
Bohrnstedt, G. W.

Bohrnstedt, G. W.

Bohrnstedt, G. W. and R. F. Carter

Bonjean, C. M., R. J. Hill, and Dale McLemore

Borgatta, E. F.
1968 "My student, the purist: A lament." Sociological Quarterly 9:29-34.

Borgatta, E. F.

Borgatta, E. F. and G. W. Bohrnstedt
1972 "How one normally constructs good measures." Sociological Methods and Research 1:4-12.

Boudon, R.

Boyle, R.

Boyle, R.
Brown, S. R. and R. W. Taylor  

Buckley, W.  

Bunge, Mario  

Burke, P. J. and K. Schuessler  

Byrne, J. J.  

Campbell, D. T. and D. W. Fiske  

Campbell, D. T. and J. C. Stanley  

Cancian, F.  

Carter, R. F.  

Cartwright, D. S.  

Clearly, T. A., R. L. Linn, and G. W. Walster  
Cohen, Jacob  
1968 "Multiple regression as a general data analytic system." Psychological Bulletin 70:426-444.

Cohen, M.  

Coleman, James S.  
1964 Introduction to Mathematical Sociology.  
New York: Free Press.

Coombs, C. H.  
1953 "Theory and methods of social measurement."  

Coser, Lewis A.  

Costner, H. L.  

Costner, H. L. and H. M. Blalock, Jr.  

Costner, H. L. and E. Schoenberg  

Costner, H. and L. Wager  

Cronbach, L.  

Curtis, R. F. and E. F. Jackson  
Denzin, N.K.

Dillman, D. A.
1972  "Increasing mail questionnaire response in large samples of the general public." *Public Opinion Quarterly* 36: 254-257.

Draper, N.R. and H. Smith

Dubin, Robert

Ducasse, C. J.

Duncan, O. D.

Duncan, O. D.

Duncan, O. D.

Duncan, O.D.

Durkheim, Emile

Durkheim, Emile
1951  *Suicide.* New York: The Free Press.

Edwards, A. C.
Einstein, A.

Etzioni, A.

Felson, M.

Finney, D.J.

Francis, R.

Fybate Lecture Notes

Ghiselli, E.E.

Gibbs, J. P.
1970a "Causation and theory construction." Social Science Quarterly 52:815-826.

Gibbs, J. P.

Gidlow, B.
Glaser, B. G. and A. Strauss  

Goode, W. J. and P. K. Hatt  

Goodman L. A.  

Gordon, R. A.  

Gould, J. and W. L. Kolb  

Gouldner A. W.  

Guttman, Louis  

Hage, Jerald  

Hammond, J. J.  

Heise, D. R.  

Heise, D. R.  

Heise, D. R.  
Heise, D. R. and G. W. Bohrnstedt

Hempel, Carl G.

Hillary, George
1955 "Definitions of community: Areas of agreement." Rural Sociology 20:111-123.

Homans, G. C.

Homans, G. C.

Hume, David

Hunter, Floyd

Inkeles, Alex

Isajiw, W. W.

Jacobson, A. and N. Lalu

Jorgskog, Karl
Kaplan, A.

Kelsall, R.K., A. Poole and A. Kuhn

Kerlinger, F.N.

Kerlinger, F. N. and E. J. Pedhazur

Kinch, J. W.

Kish, L.

Kish, L.

Klonglan, G., G. Beal, J. Bohlen and C. Mulford
1966 "Prediction of local civil defense directors' role performance: Minn. Georgia and Mass." Rural Sociology Report #52 Department of Sociology and Anthropology, Iowa State University, Ames, Iowa.

Klonglan, G., B. Yep, C. Mulford and D. A. Dillman

Kudner, G. F. and H. W. Richardson

Kuhn, T.
Labovitz, S.
1965 "Methods for control with small sample size."

Labovitz, S.
1967 "Some observations on measurement and statistics."
   Social Forces 46:151-160.

Labovitz, S.
1968 "Commentary: Reply to Champion and Morris."

Labovitz, S.
1970a "The assignment of numbers to rank order categories."

Labovitz, S.
1970b "The nunutility of tests of significance, the
   significance of tests of significance reconsidered."

Labovitz, S.
1972 "Statistical usage in sociology."
   Sociological Methods and Research 1:13-35.

Labovitz, S. and R. Hagedorn
   Book Company.

Land, K.
1969 "Principles of path analysis."

Laumann, E.O. and J.S. House
1974 "Reply to Dr. Felson's comment."

Lehmann, T. and T. R. Young
1974 "From conflict theory to conflict methodology: an emerging
   paradigm for sociology."

Lerner, D.

Li, C. C.
Lin, Nan

Locke, H. J. and K. M. Wallace
1959 "Short marital adjustment and prediction tests: Their reliability and validity." Marriage and Family Living 21:251-255.

Lotz, Roy

Lyons, M. and T.M. Carter
1971 "Comment on Boyle's path analysis and ordinal data." American Journal of Sociology 76:1112-1132.

McGinnis, Robert

Madden, E.H.

McNall, Scott

March, J. G. and H. A. Simon

Matson, F. N.

Mayer, L. S.

Mayeske, G. W., C. E. Wisler, A. E. Beaton, F. D. Weinfeld, W. N. Cohen, T. Okada, J. M. Proshak, and K. A. Tabler
Meehan, E.J.

Meltzer, L.

Menzel, H.
1950 "Comment on Robinson's 'Ecological Correlations and the behavior of Individuals.'" American Sociological Review 15:674.

Merton, R.

Merton, R.

Miller, Delbert C.

Miller, M. K. and C. S. Stokes

Mills, C. Wright

Mood, A. M.

Movahedi, S. and R. H. Ogles

Morris, R. N.
Morris, E. N.  
1970  "Multiple correlation and ordinally scaled data."  
Social Forces 48: 299-311.

Muir, D. E.  
1969  "Searching data for predictive variables: A set theoretical approach."  

Mulford, C., G. Klonglan and P. Schmitz  
Ames, Iowa: Iowa State University Press.

Mulford, C., G. Klonglan, R. Warren and P. Schmitz  
1972  "A causal model of effectiveness in organizations."  
Social Science Research 1: 61-78.

Mullins, N.C.  
1974  "Theory construction from available materials:  
a system for organizing and presenting propositions."  

Myrdal, G.  

Nagel, Ernest  


Namboodiri, N. K. and N. K. Lalu  
1971  "The average of several simple regression estimates  
as an alternative to the multiple regression estimate in postcensal and intercensal population estimation: a case study."  

Nie, N., C. Hull, J. Jenkins, K. Steinbrenner and D. Bent  
1975  Statistical Package for the Social Sciences,  
Ofshe, R. and R. Anderson  

Parsons, E. J. and T. S. Medford  

Parsons, T.  

Parsons, T.  

Parsons, T. and E. A. Shills  

Parsons, Talcott, E. Shils, K. Naegele, and J. Pitts  

Phillips, D.L.  

Phillips, D.L. and K.J. Clancy  

Phillips, D. L. and K. J. Clancy  

Poggie, J.J.  

Polk, K.  
Pollatschik, M.A.
1974 "A quick method for evaluating causal models."
Social Science Research 3:45-60.

Popper, K.R.

Porter, L. and E. Lawler
1968 "What job attitudes tell about motivation."

Pym, J.S.
Sociological Analysis 2:94-106.

Reichenbach, H.
1953 "The philosophical significance of the theory of relativity."

Reynolds, P.D.

Reynolds, P.D.
1974 "Non-parametric partial correlation and causal analysis."
Sociological Methods and Research 3:376-392.

Rhoads, J.
1971 "On causality in social science."
Sociological Inquiry 41:27-37.

Riley, Matilda

Ritzer, G.
Boston: Allyn and Bacon, Inc.

Robin, Stanley S.
1965 "Procedure for securing returns to mail questionnaires."
Sociology and Social Research 50:24-35.
Robinson, W. S.

Robinson, W. S.

Ross, Ralph

Ruby, Lionel

Saunders, D.R.

Schuessler, K.

Schwartz, M. S. and C. G. Schwartz

Scott, W.

Selltiz, C., M. Jahoda, M. Deutsch, and S. Cook

Selltiz, C., L. Wrightsman, and S. Cook

Selvin, H. C.

Shanin, Teodor
Shils, E. and H. Finch

Short, J. R., Jr.

Simon, H. A.

Sjoberg, G. and R. Nett

Skidmore, William

Sklair, L.

Snedecor, G. W. and W. G. Cochran

Somers, Robert H.

Songquist, J.A.

Sorte, Michael A.
1972 "Replication as a verification technique in survey research." Sociological Quarterly 218-227.

Sorokin, P.

Specht, D. and R. Warren
Stokes, C. S. and M. K. Miller

Sullivan, J.

Sullivan, J. L.

Sullivan, J. L.

Tannenbaum, A. and J. Bachman

Tausky, Curt

Taylor, K. and J. Frideres

Theil, H.

Thomlinson, R.

Tittle, C. R. and R. J. Hill
Turk, Austin

Turner, J.H.

Turner, J. and A. Beeghley

Upshaw, H. S.

Van de Geer, J. P.

Vroom, V.

Wallace, W. L. (ed.).

Wallace, Walter

Walster, G. W. and T. A. Clearly

Warner, W., M. Meeker and K. Ells

Warren, R., C. Faisal, C. Mulford, G. Klonglan and W. Fuller
1974 "Re-examining a causal model of effectiveness in organizations for measurement error." Unpublished paper. Department of Sociology and Anthropology, Iowa State University, Ames, Iowa.
Warren, R., J. Keller and W. Fuller

Warren, R., G. Klonglan and M. Sabri

Warren, R., G. Klonglan, and S. Tai

Warren, R., C. Mulford, and J. Winkelpleck


Warren, R. D. and J. M Winkelpleck

Warshay, L. H.

Webb, E., D. Campbell, R. Schwartz and L. Sechrest

Wheaton, B., B. Muthen, D. Alwin and G. Summers
Williams, A. F. and H. Wechsler
1970 "The mail survey: Methods to minimize bias owing to incomplete response (a research note)." Sociology and Social Research 54:533-535.

Wilson, E. K.
1966 Sociology, Rules, Roles and Relationships.
Homewood, Ill.: Dorsey Press.

Wilson, Thomas P.

Wisler, C. E.

Wright, Sewall

Wright, Sewall

Yuchtman, E. and S. Seashore

Zeller, R. and Z. Levine

Zetterberg, Hans
Totowa, New Jersey: Bedminster Press.

Zito, George V.
ACKNOWLEDGEMENT

As with most things that individuals accomplish, this dissertation exists because I received a great deal of encouragement and direction from those around me. Thanks must be given to Latifah, Najeeb and Jehan for enduring without their mother for extended periods, and to Kirsten, for not only enduring but also helping. My appreciation is also extended to Dr. Richard Warren, and Dr. Gerald Klonglan, and Dr. Charles Mulford with whom I have worked closely over the last few years both as a student and a colleague. Many of the ideas in this dissertation sprang from conversations with these individuals. I would also like to thank Dr. John Hartman who set me to thinking about this topic while I was working on my Master's Degree. Very special thanks go to my husband, Abdulla, who has been my mainstay, my co-worker, and the stimulus to go on whenever the road looked too long. Finally, last but far from least, goes appreciation to my parents who have always been there in times of "crisis".