MARSHALL, Ronald Mark, 1941-
THE SEMANTIC SPECIFICATION BEHAVIOR OF AUDITORS WITH RESPECT TO AUDITING CONCEPTS.
The Ohio State University, Ph.D., 1969
Accounting

University Microfilms, Inc., Ann Arbor, Michigan

© Copyright by
Ronald Mark Marshall
1970
THE SEMANTIC SPECIFICATION BEHAVIOR OF AUDITORS WITH RESPECT TO AUDITING CONCEPTS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Ronald Mark Marshall, B.A., M.B.A.

* * * * * *

The Ohio State University
1969

Approved by

[Signature]
Advisor
Department of Accounting
VITA

March 27, 1941  Born - Cincinnati, Ohio

1963  B.A., The Ohio State University, Columbus, Ohio

1964  M.B.A., The University of Michigan, Ann Arbor, Michigan


1966-1968  Teaching Assistant, Department of Accounting, The Ohio State University, Columbus, Ohio

1968-1969  Ernst & Ernst Research Fellow

1969-1970  Assistant Professor, Department of Accounting and Financial Administration, Michigan State University, East Lansing, Michigan

PUBLICATIONS


# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vita</td>
<td>ii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>v</td>
</tr>
<tr>
<td>List of Illustrations</td>
<td>vi</td>
</tr>
<tr>
<td><strong>Chapter</strong></td>
<td></td>
</tr>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. Semantic Specification Behavior: An Abstract Model</td>
<td>21</td>
</tr>
<tr>
<td>III. Problems of Estimation</td>
<td>57</td>
</tr>
<tr>
<td>IV. An Operational Experiment</td>
<td>76</td>
</tr>
<tr>
<td>V. An Actual Experiment Involving 382 Auditors, 50 Semantic Scales and 3 Audit-Technique Concepts</td>
<td>86</td>
</tr>
<tr>
<td>VI. Evaluation of Experimental Results</td>
<td>107</td>
</tr>
<tr>
<td>VII. Estimations of Experimentally Derived Semantic Factors</td>
<td>134</td>
</tr>
<tr>
<td>VIII. Summary, Limitations and Further Research</td>
<td>143</td>
</tr>
<tr>
<td><strong>Appendix</strong></td>
<td></td>
</tr>
<tr>
<td>A. Verification of Properties of the Model Given in Chapter II, Section 4</td>
<td>151</td>
</tr>
<tr>
<td>B. Verification of Property (1) and (11) of the Estimation Method of Chapter III</td>
<td>155</td>
</tr>
<tr>
<td>C. Instruction Set Used in the Questionnaire of $E^0(Q, \rho^a, \sigma^a)$</td>
<td>156</td>
</tr>
</tbody>
</table>
## APPENDIX

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>D. A listing of the three audit-technique concepts and illustrative</td>
<td>159</td>
</tr>
<tr>
<td>audit procedures used in the questionnaire of $E^O(Q, P^a, C^a)$</td>
<td></td>
</tr>
<tr>
<td>E. A listing of the 50 semantic scales used in the questionnaire of</td>
<td>160</td>
</tr>
<tr>
<td>$E^O(Q, P^a, C^a)$</td>
<td></td>
</tr>
<tr>
<td>F. Factor analysis as a model for semantic specification behavior</td>
<td>163</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>171</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated Latent Roots, $\hat{d}^2_{1,1}$ of $\hat{R}$, Ratios of Estimated Roots to $p$ and Ratios of Cumulative Sums of Estimated Roots to $p$</td>
<td>97</td>
</tr>
<tr>
<td>2. Semantic Content Changes Between the Initial Semantic Groups (ISG's) and the Final Semantic Groups (FSG's)</td>
<td>109</td>
</tr>
<tr>
<td>3. Identification and Interpretation of Semantic Factor Variables Derived from $E^0(Q, P^a, C^a)$</td>
<td>112</td>
</tr>
<tr>
<td>4. Classification of Semantic Variables of $\tilde{z}(4)$ According to Semantic Content</td>
<td>114</td>
</tr>
<tr>
<td>5. Estimated Variances of First Six Principal Component Variables and Cumulative Percentages of Variances to $p$</td>
<td>116</td>
</tr>
<tr>
<td>6. Average Estimated Semantic Factor Scores for Each of the Six Organizational Categories on Each of the Six Semantic Factors</td>
<td>122</td>
</tr>
<tr>
<td>7. Kruskal-Wallis Test Statistic and Level of Rejection of Null Hypothesis of No Difference Among Average Semantic Scores Over Six Organizational Categories for Each of Six Semantic Factors</td>
<td>128</td>
</tr>
<tr>
<td>8. Indicated Rankings of Organizational Categories for Each of the Six Semantic Factors</td>
<td>129</td>
</tr>
<tr>
<td>9. A Sequence of $n_p^-$ Sets of Estimated Significant Regression Weights</td>
<td>138</td>
</tr>
<tr>
<td>10. Summary of Experimental Results</td>
<td>144</td>
</tr>
</tbody>
</table>
### LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Schematic Arrangement of Sets and Functions</td>
<td>29</td>
</tr>
<tr>
<td>2.</td>
<td>A &quot;Pure&quot; Simple Structure Pattern</td>
<td>45</td>
</tr>
<tr>
<td>3.</td>
<td>$\hat{g}^B(1)$</td>
<td>99</td>
</tr>
<tr>
<td>4.</td>
<td>$\hat{g}^B$</td>
<td>104</td>
</tr>
<tr>
<td>5.</td>
<td>A Graphic Profile of the Differential Pattern of Six Semantic Factors Over Each of Six Organizational Categories</td>
<td>123</td>
</tr>
<tr>
<td>6.</td>
<td>A Graphic Profile of Six Organizational Categories Over Six Semantic Factors</td>
<td>131</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

1. The Problem

1.1 The Problem Area

This investigation represents a preliminary study of the problem of providing the management of CPA organizations with information relevant to decisions concerning the control and evaluation of audit behavior. Specifically, the investigation is concerned with information relating to the ways in which an auditor perceives various aspects of his task environment. The motivation for this concern is the assumption that an auditor's perception of his task environment has definite and systematic implications with respect to his behavior in that environment. The primary objective is to investigate one possible framework within which objective information relating to auditor perceptions can be obtained. Such information, for discussion purposes, will be referred to as perception information.

There are at least three decision areas relating to control and evaluation for which perception information is potentially useful and relevant: (1) the evaluation of the extent to which audit behavior (of auditors and in audit
engagements) is effective, (2) the control of the personnel input to the CPA organization, and (3) the maintenance and improvement of the effectiveness of audit behavior. These areas will now be discussed.

Of greatest potential relevance for perception information are decisions of evaluation—decisions concerning the extent to which audit behavior is that desired by the CPA organization. The process by which these decisions are made currently is unknown, although there appears to be, at least outwardly or formally, a substantial dependency upon the existence of generally accepted auditing standards and upon the exercise of professional judgment. The form of this dependency is the apparent assumption that the existence of generally accepted auditing standards is sufficient to reasonably insure that the exercise of professional judgment will result in appropriate or desirable audit behavior.

There are several reasons for inferring such an assumption. First, formal recognition and emphasis are given auditing standards by influential groups which suggests their importance and necessity to the auditing process. For example, the Code of Professional Ethics of the AICPA requires that audits be conducted in conformity with auditing standards (Carey, 1966, p. 62) and the Securities and Exchange Commission gives government support
and recognition by the requirement that for each audit, the auditor state whether such conformity did in fact occur (Carey, 1966, p. 78). Second, at various points in the literature, a connection between auditing standards and performance measurement is suggested. For example, in discussing the distinction between standards and procedures, the AICPA states that standards "... deal with measures of the quality of ... performance ..." (AICPA, p. 15). Similarly, the authors of a well known auditing text refer to standards as "authoritative rules" for measuring the quality of performance (Meigs and Johnson, p. 16). Also in legal circles, standards are apparently viewed as criteria for determining what constitutes appropriate conduct in an audit engagement (Berryman, p. 71). Third, the auditor's opinion itself implies, among other things, the assumption that the judgments involved in a given audit engagement, having been made in accordance with generally accepted auditing standards, are "... such as to justify the opinion" (Carey, 1962, p. 47).

A reasonable question, however, is: do these standards elicit desired professional judgment as opposed to some other type; and if so, how? The implicit assumption of the investigation is that the standards, by themselves, do not systematically and uniformly elicit such desired behavior and that in situations in which desired behavior does result, it is either not recognizable on an
objective basis or the process by which it occurred is not completely understood. There are several reasons for this assumption.

One, the standards do not represent specific and unambiguous criteria against which objective evaluation of behavior can be made. They are in the form of general objectives or goals outlining in non-specific terms the type of behavior desired (Vance, 1950, p. 96). Such generality is apparently deliberate and desired, arising from a fear that officially sponsored audit procedures indicating how standards are to be applied would have a restrictive and undesirable effect (Vance and Netter, 1956, p. 5).

Second, because of their generality and ambiguity, the extent to which they are, in fact, generally accepted and recognized standards of behavior is open to question (Carey, 1966, p. 68). Consequently, the extent to which they elicit systematic and uniform interpretation and behavior must likewise be questioned (Carey, 1965, p. 315). Concern for this point is evidenced by the profession's explicit consideration of such problems as the development of a more effectively stated Code of Professional Ethics, the development of an inter-practice review mechanism to uncover deviations from standards and, relatedly, the development of effective enforcement procedures to insure conformity with standards (Carey, 1965, pp. 331-336).
Finally, due to the complexity of the auditing process, cause and effect knowledge or understanding with respect to judgments and actions taken, the results of judgments and actions and the governing standards of evaluation is at best incomplete. The extent to which such knowledge is felt to be incomplete is, in part, indicated by the persistent and expanding use of statistical sampling. This use represents an attempt to inject greater objectivity into the auditing process at the point at which judgments are made concerning the extent of sampling to be undertaken. The hoped for result is greater knowledge of the effects of such decisions and a more objective basis for evaluating them.

In view of the foregoing, it would seem that there are at least two areas in which steps could be taken to render more objective the evaluative decisions made in the auditing area. The first one is the improvement of cause and effect understanding of the auditing process. The second is the development of specific standards or criteria of evaluation (Slocum, p. 6; Berryman, p. 80). In any evaluation situation, these areas are the prime factors or variables (Thompson, pp. 84-92). For each area, it would seem relevant to investigate the perceptions of an auditor of, for example, such aspects of his task environment as internal control, auditing methodology, audit evidence, financial assertions and generally accepted auditing
standards (Roy and MacNeill, p. 208). This research, however, will be limited to concentrating only on auditor perceptions with respect to the area of auditing methodology.

The second decision area for which perception information has potential relevance is the problem of selecting appropriate audit personnel. The obvious question is: can perception information be used to improve selection procedures? For example, could such information be used by a CPA firm to help it distinguish between those persons, otherwise similar in education and background, who are likely to succeed in the organization, those who are likely to leave the firm and those who are likely not to succeed? Would it be possible, in particular, to predict from perception information how long a prospective auditor would likely stay with the firm and the level in the firm he is most likely to reach? Such questions have obvious implications for the reduction of turnover and associated costs as well as for the possibility of predicting the potential contribution to the firm of a given selection alternative.

And finally, the third area of potential relevance are decisions concerning the maintenance and improvement of the effectiveness of audit behavior by means of appropriate indoctrination, training and supervision, and via effective communication of organizational standards and objectives. For this area, such questions as the following appear
relevant: Do training and indoctrination programs have an effect on auditor perceptions? If so, what effect and can it be related to likely changes in audit behavior? What are the auditor perceptions of organizational standards and objectives? How do such perceptions vary over organizational levels? And finally, in what ways can perception information be used to improve the understanding and communication of these standards and objectives? The potential significance of perception information for this decision area is not at this point known, but it would appear that questions such as these are worth consideration and possible investigation.

As will be seen in the following sections, the investigation will not attempt to study the potential relevance of perception information to all these decision areas. A first task in approaching the entire area and a primary task for this investigation is to determine the feasibility of obtaining perception information in an objective way. To do this, the investigation focuses on a particular type of perception—called semantic perception—and does so by considering only one particular aspect of an auditor's task environment—that having to do with the techniques by which audit evidence is obtained.

A discussion of the approach to be used and some expected difficulties completes the remainder of the first part of this chapter. In the second part of the chapter,
a more precise indication of the nature of the investigation undertaken will be given.

1.2 An Approach to the Problem—The Semantic Differential

Investigating this problem area requires an objective method of measuring an auditor's perception in a given instance—an objective framework within which an auditor's perception toward his task environment can be specified. One way is potentially provided by a psychological instrument, called the semantic differential, designed by Osgood and his associates (Osgood, 1957) to quantitatively measure what Osgood calls the "meaning" of concepts, objects and events for individuals. Basically, the instrument consists of a set of seven-interval rating scales, called semantic scales, with which an experimental subject is requested to rate a set of concepts according to a set of instructions. Each semantic scale is identified by a pair of polar (semantically opposite-in-meaning) adjectives positioned at the ends of the scale. The instructions of the differential tell the subject to rate the concepts on the scales according to what they mean to him insofar as how he perceives them to relate to the polar adjectives identifying the scales.
1.3 A Method of Analysis of Semantic Differential Data—Factor Analysis

The way in which semantic differential data has conventionally been analyzed is by use of factor analysis. Factor analysis as a method of analysis essentially involves "factoring" a sample correlation matrix in an attempt to derive a set of "common factors" which are inferred to be the "sources" of the intercorrelations of the data variables and which, hopefully, are significantly small in number as compared to the number of data variables. For semantic differential data, these factor variables are interpreted as "dimensions of meaning" with respect to the concepts and subjects of a given experimental situation.

Factor analysis as a model specifies that each of a set of p observable random variables is a linear function of a set of k "common factor" variables (k < p), which are independent and unobservable, and a set of p "specific factor" variables, which are also independent and unobservable, and which are "unique" to the observable random variables insofar as "accounting for" that portion of their data variance which is not common or shared.

1.4 Illustrations of Semantic Differential Applications

Using non-technical concepts (Osgood, p. 49), Osgood empirically derived three dominant dimensions of meaning:
an evaluation dimension, a potency dimension and an activity dimension (Osgood, pp. 31-75). These results have been supported in other subsequent semantic differential studies, as for example using theatre concepts (Smith, 1961).

In the area of business organizations, the semantic differential was applied to measure employee morale by measuring the meaning of nine morale concepts (Scott, 1967). The concepts included: "me at work," "my opportunity for growth," "my job," "top management," "company benefits," "my fellow workers," "my pay," and "my working conditions." For the concept "me at work," eight discernible dimensions were derived, examples of which are labeled: vigor, emotionality, personal work, and personal stability. The other concepts likewise yielded discernible dimensions.

Another organizational study focused on the meaning of attributes of jobs for managers and workers. The results indicated different dimensions compared to those of Osgood. Also indicated were differences in the dimensions between workers and managers (Triandis, 1960).

In an earlier study reported by Osgood (pp. 66-68), navy sonar men were asked to judge the meaning of sonar signals. The same three dimensions of evaluation, potency and activity reported in the Osgood studies were derived.

Also reported by Osgood (pp. 68-70; pp. 291-295) is a study involving the meaning of representational and abstract paintings for groups of both artists and
nonartists. Generally, the results also indicated the same three dimensions as in the sonar study.

More recently, the semantic differential has been used to determine the dimensions of meaning for such varied concepts as: televised instruction (Ohnmacht, 1966), performed folksongs (Gray and Wheeler, 1967), social role concepts (Friedman and Gladden, 1964) and human values (Osgood, Ware and Morris, 1961).

1.5 **Difficulties**

Aside from its widespread use and acceptance, there are difficulties with the semantic differential and its associated method of analysis, factor analysis.

First, the differential lacks a theoretical framework in which the concept of "meaning," which the differential operationally defines, is given a sound psychological basis, although a tentative theoretical structure is suggested by Osgood (1957, chapter 1). Without a theoretical structure, the semantic differential reduces to an empirical technique having no abstract or logical basis.

Second, and not unrelated, the method of analysis, factor analysis, upon which the differential has typically depended, has a logical difficulty. The common factor variables are not uniquely definable within the model which means that neither their measurement nor estimation is possible. And this further means that dimensions of
meaning, the object of a semantic differential analysis, are indeterminate in the sense of being represented in a model by entities which are not uniquely definable. This indeterminacy, it should be noted, does not refer to the "rotational" indeterminancy of factor analysis which can be resolved, even if arbitrarily, on an objective basis. It refers instead to a logical problem of defining precisely what is meant within the factor model by a common factor variable or, in the terminology of the semantic differential, by a dimension of meaning.

This logical indeterminacy is demonstrated in Appendix F. For discussions of the problem, see Guttman (1955).

Also in Appendix F, the factor model is briefly indicated. The notation and terminology of the appendix parallels that of Chapter II. It may therefore be more convenient for the interested reader to defer its consideration until after reading Chapter II.

2. Nature of Investigation

2.1 General Approach and Objective

To meet the difficulties of the semantic differential approach, the investigation attempts to construct an abstract rationalization or model for semantic differential behavior. In doing this, the method and model of factor analysis are not used. Instead, a model
is developed which involves the use of principal component analysis and which avoids the logical indeterminacy present in the factor model. It is not the purpose of the investigation, it should be emphasized, to develop a psychological framework for semantic differential behavior. This is beyond the bounds of the investigation. But it is the intention to construct an abstract framework within which semantic behavior can be viewed.

2.2 A Note on Terminology

To avoid unnecessary ambiguities and unintended implications, the term "meaning" will not be used. Instead, the term specification will be employed and the behavior elicited from the use of a semantic differential instrument will be called semantic specification behavior. And instead of "dimension of meaning," the designation—semantic dimension of specification—will be substituted. This terminology will be given precise meaning in terms of experimentation defined both at an idealized, abstract level and an operational level.

2.3 Scope of Investigation

Although the underlying and general concern is with an auditor's perception of his entire task environment, the specific concern of this investigation is limited to that segment of the task environment relating to auditing
methodology—namely to techniques used to obtain audit evidence. Concepts representing techniques of obtaining audit evidence will be used to experimentally represent this portion of the task environment. These concepts will be called audit-technique concepts.

2.4 Purpose of Investigation

The purpose of this study is to investigate the semantic specification behavior of auditors with respect to audit-technique concepts. The specific objectives are:

(1) To develop an abstract model in which semantic specification behavior and semantic dimensions of specification are given explicit meaning and representation.

(2) To conduct actual experimentation involving the semantic specification behavior of a sample of auditors from a national CPA firm for purposes of investigating two things:

(a) The nature and extent to which semantic dimensions of specification, as defined by the model, exist in a particular experimental instance, and

(b) The extent to which such dimensions, as do exist, can be estimated.

2.5 Overview and Content of Chapters

In this section, an overview of the investigation is given, indicating the general nature of the model and the content of the following chapters.
A first task, and the topic of Chapter II, is the development of an abstract model of semantic specification behavior for a population of subjects in general—apart from any contextual situation. The objective of the model is to define semantic specification behavior and semantic dimensions of specification. The model involves:

(a) An idealized experiment and sample space.
(b) A set of observable random variables defined on the idealized sample space and representing semantic specification behavior.
(c) A set of unobservable random variables representing a set of independent semantic dimensions of specification.
(d) A mathematical relationship involving the above sets of random variables.

An idealized experiment, for purposes of this investigation, is one void of empirical content or contextual meaning. The idealized experiment of Chapter II can be viewed as, essentially, an operation or function involving three abstract sets: (1) a set, P, representing a population of subjects whose semantic behavior is the object of concern, (2) a set, C, representing a domain of semantic concepts associated with some segment of the task environment of P, and (3) a set, Q, representing a domain of p semantic specifiers selected from a universe, Q⁺, of p⁺
semantic specifiers relevant to C and P (i.e., \( Q \leq Q^+ \) and \( p \leq p^+ \)).

Given these sets, a particular trial of the idealized experiment, denoted \( E(Q, P, C) \), is defined in Chapter II as one in which a subject from P specifies a "representative" concept, denoted \( \tilde{c} \), of C in terms of each of the \( p \) specifiers of Q. The set, \( S \), representing all possible outcomes of a given trial of \( E(Q, P, C) \) is the (idealized) sample space of \( E(Q, P, C) \). The behavior elicited by \( E(Q, P, C) \) is what the investigation calls **semantic specification behavior**.

The representation in the model of semantic specification behavior is by means of a set of \( p \) observable random variables—called **semantic (specification) variables**—defined on \( S \). The \( r \)th semantic variable denotes a numerical representation of the specification of \( \tilde{c} \) in terms of the \( r \)th specifier of Q for some particular experimental trial.

The model also involves a second set of \( k \) (\( k < p \)) unobservable and uncorrelated random variables—called **semantic factor variables**—which represent a set of independent semantic dimensions of specification postulated to exist for any given Q. The precise nature of these dimensions and factor variables is not, for reasons of brevity, indicated here. As will be seen in Chapter II,
their definition involves another idealized experiment and sample space.

The statement of the model involving these two sets of random variables is:

\[ z = Bg + z^{(2)} \]

where:

a. \( z \) — a \( px1 \) random vector of observable semantic variables.

b. \( g \) — a \( kx1 \) random vector of uncorrelated semantic factors.

c. \( z^{(2)} \) — a \( px1 \) random vector of "residual" random variables.

d. \( B \) — a \( pxk \) matrix containing the covariances between the variables of \( z \) and \( g \) (i.e., \( B = \text{cov}(z, g) = \text{E}(zg^\prime) \), where "\( \text{E} \)" denotes mathematical expectation).

As is seen, the model defines each semantic variable as a linear function of \( k \) independent (statistically) semantic factor variables and a "residual" random variable. A complete statement of this relationship and its properties is given in Chapter II.

The basic parameter of the model is the number, \( k \), of semantic factors. Although an objective basis for estimating \( k \) is not available, a two-stage estimation procedure is suggested in Chapter III. Also in Chapter III, a method for estimating the variables of \( g \), involving the
least-squares criterion, is given. It should be noted in this connection that the model, as will be seen, defines each of the variables of g as a linear function of the p observable variables of z. The objective of the estimation procedure of Chapter III is to determine to what extent the variables of g can be estimated by mutually exclusive subsets of the variables of z, each containing a substantially smaller number of variables than the total number, p.

In Chapter IV, the abstract sets of P, C and Q of Chapter II are given contextual meaning, and an operational experiment, E°(Q, P, C), and sample space, S°, are defined which involve these contextually defined sets. E°(Q, P, C) is of the same form as E(Q, P, C) but is such as to be empirically possible.

The contextual definitions given in Chapter IV are: P is defined as the population of auditors on the audit staff of a national CPA firm; C is defined as the domain of concepts representing audit techniques; and Q is defined as a set of seven-interval semantic (specification) scales of the type typically used in the semantic differential (see section 1.2 of this chapter).

Also in Chapter IV, the model of Chapter II is restated in terms of E°(Q, P, C). Although the restatement is analogous to the statement in Chapter II, it reflects
differences arising from the fact that \( E^O(Q, P, C) \) of Chapter IV is an empirical approximation of \( E(Q, P, C) \) of Chapter II. One difference, for example, arises from the fact that a "representative" concept, \( \bar{c} \), of \( C \) cannot be empirically generated. Rather, an approximation is made by having each subject respond to a sample of concepts from \( C \) and considering the "average" response as the empirical equivalent of a subject's response to \( \bar{c} \). Another difference results from the fact that the specifiers of \( Q \) in Chapter IV allow the subject only a finite number of response choices, whereas the specifiers of \( Q \) in Chapter II are unrestricted and thereby afford the subject an infinite number of response choices. The result is that the random variables of the model in Chapter II are continuous variables; those of Chapter IV are discrete.

The statement of the model for \( E^O(Q, P, C) \) is distinguished from its counterpart in Chapter II by the use of "bar" notation. That is, the model of Chapter IV appears:

\[
\bar{z} = \bar{g}g + \bar{z}^{(2)}
\]

Such notation is used throughout Chapter IV and all subsequent chapters since they are concerned only with \( E^O(Q, P, C) \).

It should be emphasized that the primary statement of the model in both Chapters II and IV is in terms of a
population of subjects. This is to be contrasted against a less fundamental statement of the model in terms of a sample of subjects, which is briefly indicated at the ends of both Chapters II and IV.

The remaining chapters deal with actual experimentation. Chapter V indicates the results of an actual experiment involving 382 auditors from P, 50 semantic scales from Q+ and 3 audit-technique concepts from C. In Chapter VI, an evaluation of the results is given. In Chapter VII, particular estimations of the semantic factors derived from the experimental observations are computed according to the method of Chapter III. Finally, Chapter VIII contains a brief summary, some limitations of the investigation and suggested further research.
CHAPTER II

SEMANTIC SPECIFICATION BEHAVIOR:
AN ABSTRACT MODEL

0. Introduction

This chapter will present an abstract model of semantic specification behavior with respect to semantic concepts. Semantic specification behavior is defined in terms of idealized experiments involving abstract sets of subjects, semantic concepts and semantic specifiers. The model consists of definitions of these sets and postulated relationships among their elements. The model does not include statistical-type considerations (Chapter III) or questions of an empirical nature relating to particular contextual situations (Chapter IV).

The chapter is organized into seven parts. Part 1 is concerned with definitions of elementary sets and concepts (sections 1.1 - 1.5); definitions of two idealized experiments and associated sample spaces (section 1.6); a postulated relationship between the two experiments (sections 1.7 and 1.8); and two random variable functions, one defined on each of the two idealized sample spaces (sections 1.9 and 1.10). Part 2 outlines briefly principal
component analysis of a given set of observable random variables. In part 3, an explicit statement of the model is given involving principal component variables and the random variables of part 1. This is followed in part 4 by properties of the model. In part 5, the model is discussed for purposes of a general interpretation (sections 5.1 - 5.4) and to suggest the potential significance of its basic elements (sections 5.5 - 5.8). Part 6 considers the model for different experimental situations by postulating certain relationships (sections 6.1 and 6.2) and indicating consequent experimental implications (section 6.3). Finally, part 7 presents the model in terms of a sample of experimental observations.

1. Preliminary Definitions and Concepts

1.1 A Population Domain

Let $P$ be a finite set of elements representing a population of subjects whose semantic specification behavior is the object of investigation.

1.2 A Concept Domain

Let $C$ be a finite set of elements representing semantic concepts which function as related and relevant parts of a language system (or sub-system) of $P$. The concepts are related in that they represent distinguishable parts of some meaningful and coherent segment of the
environment of P. They are relevant insofar as constituting parts of the language system to which the subjects of P can and do systematically respond. The set C will be called a concept domain for P.

1.3 A Specifier Domain

Let Q be a finite set of \( p^+ \) elements representing the universe of semantic specifiers relevant to C and P. That is, the elements of \( Q^+ \) are to represent a mechanism or framework in terms of which the subjects of P specify (or qualify) the concepts of C. The set \( Q^+ \) will be called a specifier domain relevant to C and P.

Let Q be any subset of \( Q^+ \) containing \( p \) elements (i.e., \( Q \subseteq Q^+ \), \( p \leq p^+ \)).

1.4 Specification Capacity of Q

Associated with the set Q is the notion of specification capacity (SC) defined as the extent to which the relevant aspects or characteristics of the concepts of C can be "effectively" specified by the subjects of P in terms of the specifiers of Q. Let \( SC(Q) \) denote such capacity for the set Q with respect to C and P.

A more precise indication of SC is given later in the chapter (section 6.3) after the model has been presented and discussed. There, various criteria are given by which a relative comparison or ranking of two (or more) Q's in terms of SC can be made.
The intended interpretation of this notion of capacity is that "part" of the specifiers of a given Q is postulated to represent "effective" specification potential. The remaining "part" is considered to constitute redundant or ineffective specification potential.

It is also intended to imply that maximum SC occurs when Q is the total universe of specifiers—that is, when \( Q = Q^+ \). Any other Q is viewed as possessing less SC, that is

\[
SC(Q) < SC(Q^+) \tag{1.5}
\]

where \( Q \subseteq Q^+ \) and "<" denotes "less than."

**1.5 Semantic Dimensions of Specification**

For any Q, let \( Q^* \) be a finite set of elements representing \( k \) (where \( k < p \)) independent and "ideal" specifiers which are

1. derivable from Q by a unique transformation J—that is,

\[
Q^* = \{ q^*_r \mid q^*_r = J(q), r = 1, \ldots, k < p \},
\]

and which

2. constitute the most efficient and parsimonious representation of SC(Q) in the sense that each element of \( Q^* \) functions as a potential specifier of some independent and unique property or aspect of the concepts of C.

The elements of \( Q^* \) will be referred to as independent **semantic dimensions** of specification relevant to Q, C and P.
1.6 Two Idealized Experiments and
Associated Sample Spaces

For a given $Q$ and $Q^*$, let $E^*(Q^*, P, C)$ represent an
idealized experiment. Define a given trial of $E^*(Q^*, P, C)$
as one in which a subject from $P$ specifies an "idealized"
concept, denoted $\tilde{c}$ (and defined as a perfect or typical
representation of the concepts of $C$), in terms of the $k$
semantic dimensions of $Q^*$. Let $S^*$ be a set representing
all possible specification response outcomes for a given
trial of $E^*(Q^*, P, C)$. That is, $S^*$ will denote the sample
space of $E^*(Q^*, P, C)$ where the ordered $k$-tuple,

$$s^* = (s_1^*, ..., s_r^*, ..., s_k^*),$$

represents an element of $S^*$. The $r$th component, $s_r^*$, of $s^*$
represents the specification of $\tilde{c}$ by a given subject of $P$
in terms of the $r$th semantic dimension, $q_r^*$, of $Q^*$.

In a similar manner, define a given trial of a
second idealized experiment, $E(Q, P, C)$, as one in which a
subject of $P$ specifies $\tilde{c}$, for a given $C$, in terms of the $p$
specifiers of $Q$. Let $S$ be a set representing all possible
specification response outcomes for a given trial of
$E(Q, P, C)$. That is, $S$ will denote the sample space of
$E(Q, P, C)$ where the ordered $p$-tuple,

$$s' = (s_1', ..., s_j', ..., s_p'),$$

represents an element of $S$. The $j$th component, $s_j'$, of $s'$
represents the specification of $\tilde{\sigma}$ by a given subject of $P$ in terms of the $j$th specifier, $q_j$, of $Q$.

The behavior involved in either of these two experiments will be called semantic specification behavior.

1.7 A Postulated Transformation

Of the two experiments, $E^*(Q^*, P, C)$ is to be preferred. Since it is based on $Q^*$, it will yield information in terms of which the subjects of $P$, if it is desired, can be most effectively differentiated.

Suppose, however, $Q$ is known, but the transformation $J$ and hence the set $Q^*$ are unknown. For this situation, which is assumed by the model, only an experiment of the form $E(Q, P, C)$ is possible. Consequently, only the response outcomes of the sample space $S$ are observable. Those of $S^*$ involving $Q^*$ are not. A basic assumption of the model is that the equivalent of the information associated with the set $S^*$ can be determined from the information obtained by conducting the experiment $E(Q, P, C)$.

The way in which this assumption is incorporated into the model is indicated in terms of a numerical representation defined on $S^*$: Let $f$ be a function defined on $S^*$ such that
where: $g_r = f_g(s^*_r) \in \mathbb{R}^l$, the set of real numbers and
\[ E(g_r g^*_s) = 0 \quad r \neq s \quad r, s = 1, \ldots, k \]
(the symbol "E" denotes mathematical expectation).

That is, $g$ is defined as a kxl random vector of continuous and uncorrelated random variables assumed observable for any given trial of $E(Q^*, P, C)$. These variables will be called **semantic factors** relevant to $C$ and $P$ for some particular $Q$.

With this definition, the assumption that information obtainable from $E(Q, P, C)$ is equivalent to that obtainable from $E^*(Q^*, P, C)$ can be expressly stated in the form of a postulated transformation, $K$, such that
\[(1.7.2) \quad g = K(s)\]
where $s \in S$ and $g$ is a random vector of semantic factor variables representing the response specifications of $\tilde{q}$ by a given subject from $P$ in terms of the semantic dimensions of $Q^*$. In other words, the model postulates that $g$, defined originally on $S^*$, is obtainable by a certain transformation of the elements of $S$ arising from $E(Q, P, C)$. The explicit
statement of the model in section 3 gives a precise
definition of this postulated transformation.

1.8 A Schematic Arrangement

A schematic arrangement of the sets and functions
which have been introduced appears in Figure 1. Dotted
lines are used in the diagram to indicate those sets and
functions which are assumed unknowable or unobservable.
Solid lines indicate those which are knowable and observable. Also, the diagram depicts $E(Q, P, C)$ and
$E*(Q*, P, C)$, the two idealized experiments, as transforma-
tions of the elements of $Q, P$ and $C$ and $Q*, P$ and $C$,
respectively.

1.9 Semantic Specification Variables

Let $f_y$ be a function defined on $S$ such that

\[
y = f_y(s) = \begin{pmatrix} y_1 \\ \vdots \\ y_j \\ \vdots \\ y_p \end{pmatrix}
\]

where $y_j = f_y(s_j) \in \mathbb{R}^1$, the set of real numbers. That is, $y$ is defined as a p x 1 random vector of continuous random
variables assumed observable for any given trial of
$E(Q, P, C)$. The random variables of $y$ will be called
semantic (specification) variables.
FIGURE 1
SCHEMATIC ARRANGEMENT OF SETS AND FUNCTIONS

\[ R^k = \{(s_1, \ldots, s_k) \mid s_1 \in \mathbb{R}^l, \ldots, s_k \in \mathbb{R}^l\} \]
It will be assumed that corresponding to the variables of \( y \) is a multivariate distribution function, \( f \), defined on \( S \). A specific form of the distribution need not be assumed. It will, however, be assumed that the first and second moments of the marginal distributions of \( f \) for each of the \( y_j \) exist and, without loss of generality, that 
\[
E(y) = [E(y_j)] = \varnothing_{px1}, \text{ a px1 null vector.}
\]

1.10 Standardized Semantic Variables

Define a new random vector \( z \):
\[
(1.10.1) \quad z = D^{-1/2}y
\]
where:
\[
D_y = \text{diag}(\Sigma_{yy})
\]
\[
\text{var}(y) = E(yy') = \Sigma_{yy}, \text{ the pxp covariance matrix of } y
\]
\[
\text{var}(z) = E(zz') = D^{-1/2}_y \Sigma_{yy} D^{-1/2}_y = R, \text{ the pxp nonsingular, symmetric population correlation matrix.}
\]

That is, \( z \) is defined as a px1 random vector of standardized observable random variables having zero means and unit variance. The random variables of \( z \) will be called standardized semantic (specification) variables.

The statement of the model given in section 3 will be formulated in terms of these standardized variables. It should be noted, however, that there are differences between a model formulated in terms of \( y \) and one in terms of \( z \).
These differences, however, are not felt to be crucial and the choice of formulating the model in terms of \( z \) as opposed to \( y \) was made primarily on the basis of practical considerations.

2. Principal Component Analysis

2.1 Principal Component Variables

The model involves the use of the method of principal components. The nature of this method will now be indicated. For more extensive coverages, see Anderson (1958, chapter 11, 1963, pp. 122-148), Morrison (1967, chapter 7) or Lawley (1963, chapter 4).

The method of principal components essentially involves an orthogonal linear transformation of the variables of \( z \) into a set of \( p \) uncorrelated random variables, \( u_r (r = 1, \ldots, p) \), called the principal component variables of \( z \). The transformation is such that the \( r \)th variable, \( u_r \), has maximum variance of all normalized linear combinations uncorrelated with \( u_1, \ldots, u_{r-1} \).

Let such a transformation be indicated:

\[
(2.1.1) \quad u = L'z
\]

where \( L \) is a \( pxp \) orthogonal linear transformation, \( u \) is a \( pxl \) random vector of principal component variables and where the covariance matrix of \( u \) is
(2.1.2) \( \text{var}(u) = E(uu^\T) = L^\T R L = D_d^2 = \begin{pmatrix} d_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_p^2 \end{pmatrix} \)

It can be shown (Anderson, 1958, p. 276) that (assuming \( R \) is nonsingular)

\[
(2.1.3) \quad d_1^2 > \cdots > d_p^2 > 0
\]

are the latent roots of the determinantal equation

\[
|R - d^2 I| = 0
\]

and that the columns of \( L \) are the corresponding unit-length eigenvectors. That is, \( L^\T L = I \) and the \( r \)th column of \( L \), \( L(r) \), satisfies

\[
(2.1.4) \quad (R - d_r^2 I) L(r) = \phi_{p \times 1} \quad r = 1, \ldots, p.
\]

2.2 Standardized Principal Component Variables

The standardized principal component variables, denoted by the random vector \( x \), are defined:

\[
(2.2.1) \quad x = (D_d^2)^{-1/2} u = D_d^{-1} u
\]

which, using (2.1.1), yields

\[
(2.2.2) \quad x = D_d^{-1} L z
\]

From (2.1.2) and (2.2.2), it follows that

\[
(2.2.3) \quad E(xx^\T) = I_{p \times p} \text{ (an identity matrix)}
\]

indicating that \( x \) is a random vector of \( p \) standardized and uncorrelated random variables.
2.3 Notation

The following notation is defined:

(2.3)a \[ L = (L^{(1)}, L^{(2)}) \]
represents a partitioning of \( L \) such that \( L^{(1)} \) is a \( p \times k \) matrix containing (as columns) the 1st \( k \) unit-length eigenvectors \( (k \neq 0) \) associated with the largest \( k \) latent roots of \( R \) and \( L^{(2)} \) is a \( p \times m \) matrix containing the last \( m \) unit-length eigenvectors \( (m \neq 0) \).

(2.3)b \[ D^2_d = \begin{pmatrix} D^2_{d_1} & 0 \\ 0 & D^2_{d_2} \end{pmatrix} \]
represents a partitioning of \( D^2_d \) such that \( D^2_{d_1} \) is a \( k \times k \) diagonal matrix containing the \( k \) largest latent roots of \( R \) and \( D^2_{d_2} \) is an \( m \times m \) diagonal containing the remaining latent roots.

(2.3)c \[ u = \begin{pmatrix} u^{(1)} \\ u^{(2)} \end{pmatrix} \]
represents a partitioning of \( u \) such that \( u^{(1)} \) is a \( k \times 1 \) vector of the 1st \( k \) principal component variables and \( u^{(2)} \) is an \( m \times 1 \) vector of the remaining principal component variables.

(2.3)d \[ x = \begin{pmatrix} x^{(1)} \\ x^{(2)} \end{pmatrix} = \begin{pmatrix} D_{d_1}^{-1} u^{(1)} \\ D_{d_2}^{-1} u^{(2)} \end{pmatrix} \]
3. A Statement of the Model

For a given experiment \( E(Q, P, C) \), the model can be stated:

\[
(3.1) \quad z = z^{(1)} + z^{(2)}
\]

where:

a. \( P = k + m \)

b. \( z^{(1)} = Bg \)

c. \( z^{(2)} = L(2)D_{d_2}x^{(2)} \)

d. \( B = AT \)

e. \( A = L(1)D_{d_1} \)

f. \( g = T^r x^{(1)} \)

Definition of \( T \)

Let \( B^{(n)} \) be a pxk matrix in which the \( j \)th \( r \)th element is

\[
b_{jr}^{(n)} = \left( \frac{a_{jr}}{\sqrt{\sum_{r=1}^{k} \sum_{j=1}^{p} a_{jr}^2}} \right)^{1/2} \quad r = 1, \ldots, k \quad j = 1, \ldots, p
\]

where \( a_{jr} \) is the \( j \)th \( r \)th element of \( A \). That is, \( B^{(n)} \) is the matrix \( A \) normalized by rows. Let \( T^* \) be an element of the set, \( OT \), of all possible orthogonal transformations of the column vectors of \( B^{(n)} \). Let \( B^*(n) \) be an element of the corresponding set, \( TM \), of all possible transformed matrices where \( B^*(n) = B(n)T^* \). Then define \( T \) as that
orthogonal transformation of OT for which the following function defined on TM is maximized:

\[ V^* = \sum_{r=1}^{k} \left[ \sum_{p=1}^{P} \left( \sum_{j=1}^{J_r} b_r(n)^2 \right)^2 - \left( \sum_{j=1}^{J_r} b_r(n)^2 \right)^2 \right] \]

where \( b_r(n)^2 \) is the \( j \)th \( r \)th element of \( B_r(n) \).

This function is Kaiser's "normal varimax criterion" (see Harman, p. 306; Kaiser, 1958; Morrison, p. 285). The effect of such a transformation criterion is discussed in section 5.6.

The following additional notation is defined:

\[(3.2) a \quad M(1) = E(z(1)z(1)^\top) = \text{var}(z(1))\]

and

\[(3.2) b \quad M(2) = E(z(2)z(2)^\top) = \text{var}(z(2))\]

Using (3.1)b, (3.1) can be written

\[(3.3) \quad z = Bg + z(2)\]

from which it follows that

\[(3.4) \quad R = BB^\top + M(2) = M(1) + M(2)\]

4. Properties of the Model

From (2.3)a-d, (3.1)a-g, (3.2)a-b, (3.3) and (3.4), the following properties are deduced (see Appendix A for proofs):
(i) \[ \text{var}(z) = \text{var}(z^{(1)}) + \text{var}(z^{(2)}) \]
\[ = M^{(1)} + M^{(2)} \]
\[ = L^{(1)}D_{d1}^{2}L^{(1)^{\prime}} + L^{(2)}D_{d2}^{2}L^{(2)^{\prime}} \]
\[ = L \sum_{d}D_{d}^{2}L^{\prime} \]
\[ = R \]

(ii) \[ \text{cov}(z^{(1)}, z^{(2)}) = E(z^{(1)}z^{(2)^{\prime}}) = \varphi_{pp} \]

(iii) a. \[ \text{cov}(z, g) = E(zg^{\prime}) = B \]

b. \[ \text{cov}(z^{(1)}, g) = E(z^{(1)}g^{\prime}) = B \]

c. \[ \text{cov}(z, x^{(1)}) = E(zx^{(1)^{\prime}}) = A \]

d. \[ \text{cov}(z, x^{(2)}) = E(zx^{(2)^{\prime}}) = L^{(2)}D_{d2}^{2} \]

(iv) The amount of variance of the variables of \( z \)
accounted for by those of \( z^{(1)} \) and \( z^{(2)} \) is given by
the traces of the corresponding covariance matrices,
\( M^{(1)} \) and \( M^{(2)} \), which, in turn, are equal to the
traces of \( D_{d1}^{2} \) and \( D_{d2}^{2} \), respectively.\(^1\)

That is,

a. \[ \text{trace} \left( \text{var}(z^{(1)}) \right) = \text{trace}(M^{(1)}) = \text{trace}(D_{d1}^{2}) \]

b. \[ \text{trace} \left( \text{var}(z^{(2)}) \right) = \text{trace}(M^{(2)}) = \text{trace}(D_{d2}^{2}) \]

c. \[ \text{trace} \left( \text{var}(z) \right) = \text{trace}(D_{d}^{2}) = \text{trace}(R) = p \]

\(^1\) The trace of a matrix is defined as the sum of its diagonal elements. Thus, for example, the
\[ \text{trace}(D_{d1}^{2}) = \sum_{i=1}^{k} d_{i}^{2}. \]
5. Discussion of the Model

5.1 Parameters of the Model

The parameters of the model are:

1. The matrix $D$, consisting of the $p$ latent roots, $d_1^2, \ldots, d_p^2$, of $R$.
2. The matrix $L$, containing the corresponding unit-length eigenvectors, $L(1), \ldots, L(p)$.
3. The number, $k$, of semantic factor variables of $g$.

The problems of estimating these parameters will be discussed in Chapter III.

5.2 The Segmentation of $z$

The statement of the model in (3.1) indicates that the random vector $z$ is segmented into two orthogonal parts, $z^{(1)}$ and $z^{(2)}$ (i.e., from property (ii), $\text{cov}(z^{(1)}, z^{(2)}) = \phi_{pp}$). Each part is defined as a linear function of an orthogonal (i.e., independent and uncorrelated) set of random variables. The first part, $z^{(1)}$, consisting of $p$ variables, is defined as a linear function of the semantic factor variables of the random vector $g$. The second part, $z^{(2)}$, consisting also of $p$ variables, is defined as a linear function of the variables of the random vector $x^{(2)}$.

5.3 Terminology

The semantic factor variables of $g$ were defined in section 1.7 as functions on the postulated sample space $S$. 
the set of possible response specifications of $\bar{c}$ in terms of $Q^*$, the most efficient and parsimonious representation of $SC(Q)$—the effective specification potential associated with $Q$. Accordingly, since the variables of $g$ in (3.1) define or "generate" the $p$ variables of $z^{(1)}$, the latter will be referred to as effective specification variables (ESV).

The total variance of the variables of $z^{(1)}$, given by the trace($D_{d1}^2$), will be called effective specification variance (ESVR). It will be interpreted as that amount of the total variance of the variables of $z$ "attributable to" or "explained by" the variables of $g$. In other words, the semantic factor variables of $g$ are interpreted as the "sources" of ESV.

The remaining part of $z$, $z^{(2)}$, is viewed as a residual term representing redundant, unsystematic, random, unreliable or otherwise ineffective specification. Accordingly, the $p$ variables of $z^{(2)}$ will be referred to as residual specification variables (RSV) and the corresponding variance, given by the trace($D_{d2}^2$), residual specification variance (RSVR).

5.4 The Explicit Nature of the Postulated Transformation, $K$

In section 1.7, a transformation, $K$, on $S$ yielding $g$ was postulated. The explicit nature of this transformation, incorporated in the statement of the model in (3.1),
can be seen by substituting into (3.1)f the definition of \( x^{(1)} \) given in (2.2.2) and using (1.9.1) and (1.10.1). This yields

\[
(5.4.1) \quad g = T'x^{(1)}
\]

\[
= T'D^{-1}L(d_1)z
\]

from (2.2.2)

\[
= T'D^{-1}L(d_1)D_y f_y
\]

from (1.10.1)

\[
= T'D^{-1}L(d_1)D_y f_y(s)
\]

from (1.9.1)

From (5.4.1) we see that \( K, \) postulated in (1.7.2), is

\[
(5.4.2) \quad K = T'D^{-1}L(d_1)D_y f_y
\]

That is,

\[
(5.4.3) \quad g = K(s) = f_g(s^*)
\]

The model, thus, postulates the existence of the transformation, \( K, \) defined on the sample space \( S \) which will yield a result, \( g, \) equivalent to that resulting from the function \( f_g \) defined on the sample space \( S^*. \) Consequently, the model postulates that the experiment \( E(Q, P, C), \) yielding the sample space \( S, \) is equivalent to the experiment \( E^*(Q^*, P, C), \) yielding the sample space \( S^*, \) in terms of the resulting semantic factor variables of \( g \) (see Figure 1).

5.5 Potential Significance of Semantic Factors

Since the set of semantic factor variables of \( g \) is a numerical representation of a postulated set of independent semantic dimensions—a set of "ideal" semantic
specifiers, each relating to a unique property or characteristic of the concepts of C—they constitute a potentially significant framework in terms of which to differentiate the subjects of P. That is, they represent potential parameters with which to describe and specify.

Assuming differentiation is possible on a systematic basis, two related problems arise and are mentioned at this point:

(1) How are the semantic factors to be interpreted in a given contextual situation?

(2) Would it be possible to "meaningfully" estimate or predict the semantic factors of g with a subset of the random variables of y?

The first problem is concerned with giving meaning to a semantic factor framework.

The second problem is concerned with a more efficient method of measuring in terms of semantic factors. That is, from (5.4.1), we know that

(5.5.1) \[ g = T^{-1}L(1)^{\dagger}D^{-1}y \]

or that each postulated semantic factor is defined to be a linear function of p semantic variables. The problem is whether a linear function of a subset of y (containing a substantially smaller number than p variables) can be found which closely approximates g. The problem is considered in the next chapter.
A solution to both of these problems depends upon the linear relationship between the semantic factors of \( g \) and the standardized semantic variables of \( z \) (or, equivalently, the unstandardized semantic variables of \( y \)). Such a relationship, in turn, is closely related to the transformation, \( T \), defined in (3.1)\( g \). This transformation will now be discussed.

5.6 The Transformation \( T \)

According to (3.1)\( f \), the transformation \( T \) is part of the definition of \( g \). In fact, its role in that definition is significant in that it is designed to transform the variables of \( x^{(1)} \) (the first \( k \) standardized principal component variables) in such a way as to "simplify," to the extent possible, the linear relationships between the variables of \( x^{(1)} \) and those of \( z \).

The nature of the transformation can best be seen by defining a new matrix \( H \) of order \( p \times k \):

\[
H = (h_{jr}) = (b^*(n)^2 - \theta^*(n))^2 - \theta^*(n))
\]

\( j = 1, \ldots, p; \quad r = 1, \ldots, k \)

\( ^2 \)This assumes that if

\[ \text{cov}(z_i, g_r) \geq \text{cov}(z_j, g_r), \text{ then} \]

\[ \text{var}(y_i) \geq \text{var}(y_j), i, j = 1, \ldots, p; \quad r = 1, \ldots, k. \]

A more complete explanation of this point is given in Chapter III (section 3.3). Also, see section 5.8 of this chapter.
where $\overline{\delta^*(n)} = \frac{1}{p} \sum_{j=1}^{p} b^*(n)$. That is, $H$ is defined such that the $r$th column contains the deviations of the squared components of the $r$th column of $B^*(n)$ (see (3.1)g) from the mean value of these squared components. Let $\mathcal{S}C$ be the set of all such possible matrices. Defining $T$ such that the function $V^*$ in (3.1)h is maximized is equivalent to maximizing the function--trace $H^*H$--defined on $\mathcal{S}C$. That is, it is equivalent to selecting that transformation of the set $\mathcal{Q}T$ (see (3.1)g) such that the sum of the squares of the elements of the resulting matrix $H$ of $\mathcal{S}C$ is maximum.

This can be shown by partitioning $H$ by columns:

(5.6.2) $H = (h_1, \ldots, h_k)$

Then $H^*H$ is

(5.6.3) $H^*H = \left(\begin{array}{c} h_1 \\ \vdots \\ h_k \end{array}\right) \left(\begin{array}{c} h_1, \ldots, h_k \end{array}\right) = \left(\begin{array}{ccc} h_1^2 & \ldots & h_k^2 \\ \vdots & \ddots & \vdots \\ h_k^2 & \ldots & h_k^2 \end{array}\right)$

and its trace is

(5.6.4) $\text{trace}(H^*H) = h_1^2 + \ldots + h_k^2$

$$= \sum_{r=1}^{k} h_r^2$$

$$= \sum_{r=1}^{k} \left[ \sum_{j=1}^{p} b_{jr}^2 \right]$$
And since $p$ is a constant, maximizing $V^*$ is equivalent to maximizing the trace $(H^TH)$.

The effect of the transformation is that the components of the column vectors of $B$ are as close as possible to either unity ($\pm 1$) or zero (Harman, p. 305; Kaiser, p. 190). This follows from the nature of $V^*$ and the fact that the components of $B$ are correlation coefficients reflecting the linear relationships between two sets of standardized variables—those of $z$ and $g$.

5.7 A Simple Structure Pattern

The criterion of $T$, Kaiser's normal varimax criterion, is intended to bring about a transformation of the variables of $x^{(1)}$ such that a "simple structure" pattern (Thurstone, chapter XIV) exists between the variables of $z$ and the resulting transformed variables of $g$. Kaiser's criterion represents an objective interpretation of the concept of "simple structure" as developed by Thurstone.

A "simple structure" pattern in its purest form is one in which (1) each of the semantic variables of $z$ is
"significantly" related, in terms of correlation, to one (or no more than a relative few) of the variables of g and (2) one in which, for a given variable of g, there is a relatively "small" number of the semantic variables of z with which a "significant" linear relationship exists.

Denoting a "significant" relationship as "+" and an "insignificant" one as a blank space, the matrix B in the event of such a "pure" relationship would appear as in Figure 2. In Figure 2, nr represents the number of variables of z relating significantly with the rth variable of g (r = 1, ... k).

5.8 Implications of Simple Structure: Potential Interpretation and Prediction

If the results of the transformation indicate a "simple structure" pattern does exist to a sufficient degree, then potentially "meaningful" interpretation and prediction may be possible in terms of mutually exclusive subsets of the variables of y, wherein each subset is a potential source of interpretation and prediction for one, and only one, semantic factor variable. Whether or not "meaningful" interpretation and prediction are possible for a given semantic factor depends upon the semantic nature or content of the semantic variables in the corresponding subset. This, in turn, involves elements of a particular experimental situation, apart from the abstract
\[ \text{cov}(z, g) = B = \]

FIGURE 2
A "PURE" SIMPLE STRUCTURE PATTERN
model of this chapter. These contextual aspects will be taken up in Chapter IV when an operational experiment is considered.

A point implied by the footnote in section 5.5 should, at this time, be made more explicit. If the variances of the semantic variables of \( y \) are assumed to meet the condition indicated in that footnote, then the matrix \( B \) also reflects the nature of the linear pattern between the variables of \( y \) and those of \( g \). This point has relevance to the problem of predicting the variables of \( g \) with a subset of those of \( y \). The assumption is explained more fully in the next chapter in the context of such a prediction problem.

6. **Investigating the Model for Different Experimental Situations**

6.1 **Three Characteristics of the Segmentation of \( z \)**

There are three important characteristics of the segmentation of \( z \) in (3.1):

1. The proportion of the total specification variance of the variables of \( z \) which is ESVR. That is, the ratio
   \[
   \frac{\text{trace}(D_{d_1}^2)}{\text{trace}(D_d^2)} = \frac{\text{trace}(D_{d_1}^2)}{p}.
   \]
2. The number, \( k \), of semantic factor variables of \( g \).
3. The random vector \( g \) of \( k \) semantic factor variables.

As indicated in section 5.1, the second characteristic, the number, \( k \), of semantic factors, is one of the
parameters of the model to be estimated. Once $k$ is estimated, the $\text{trace}(D_{\text{d1}}^2)/p$ is automatically determined as well as the random vector $g$.

6.2 Postulated Relationships

It will be noted that the statement of the model in (3.1) is with respect to a particular $Q$. It is of interest to consider or postulate what will happen to the model in terms of these three characteristics for different $Q$'s—that is, for different experimental situations. In particular, two experimental situations will be considered:

(i) $Q = Q^+$ (i.e., the experiment $E^+(Q^+, P, C)$

and

(ii) $Q = Q^-$ (i.e., the experiment $E^-(Q^-, P, C)$ where $Q^- \subset Q^+$)

It will be recalled (section 1.3) that $Q^+$ represents the universe of semantic specifiers relevant to $C$ and $P$. And $Q$ is any subset of $Q^+$.

For purposes of the following discussion, the superscript "+" will be used to indicate the model for the situation $Q = Q^+$, the superscript "-" will indicate the model for the situation $Q = Q^- \subset Q^+$ and an absence of a superscript will denote the general case. Thus, the segmentation of $z$ for the three situations would appear:

(i) situation $Q$ (the general case):

$$z = z^{(1)} + z^{(2)}$$

where $z$ is a px1 random vector.
(ii) situation $Q^+$ (the total set):
$$z^+ = z^+ (1) + z^+ (2) \text{ where } z^+ \text{ is a } p^+)x1 \text{ random vector}$$

(iii) situation $Q^-$ (any subset not equal to $Q^+$):
$$z^- = z^- (1) + z^- (2) \text{ where } z^- \text{ is a } p^-)x1 \text{ random vector}$$
and where $Q$ contains $p$ elements, $Q^+$ contains $p^+$ elements
and $Q^-$ contains $p^-$ elements ($p^- < p^+$).

For the two experimental situations, $E^+(Q^+, P, C)$ and $E^-(Q^-, P, C)$, the following relationships with respect to the three segmentation characteristics are postulated:

(6.2.1)a \[ \frac{\text{trace} D^2_{d1}}{p^-} < \frac{\text{trace} D^2_{d1}}{p^+} \]

(6.2.1)b \[ k^- < k^+ \]

(6.2.1)c \[ g_r^- \neq g_r^+ \text{ for } r = 1, \ldots, k^- \text{ and where the } \]
\[ g_r^+(r = 1, \ldots, k^-) \text{ constitute any subset of the } k^+ \text{ variables of } g^+. \]

The relationship postulated in (6.2.1)a means that the portion of the total variance of the variables of $z$ which is ESVR will be greater for the case of $Q^+$ than for $Q^-$. This is interpreted to imply that the "part" of the total specification constituting effective specification will be greater for an experiment, $E^+(Q^+, P, C)$, involving the universe of potential specifiers, $Q^+$, than for an experiment, $E^-(Q^-, P, C)$, involving any subset, $Q^-$, of that universe.

The postulated relationship (6.2.1)b means that the number of the "sources" of ESVR will be larger for situation
Q\(^+\) than for situation Q\(^-\). This suggests that fewer dimensions of specification are potentially identifiable when only a subset of the specifiers of Q\(^+\) are involved in a given experiment.

(6.2.1)c is intended to imply that the semantic dimensions represented by the variables of g\(^-\) for the case Q\(^-\) will not be identical to any corresponding set of semantic dimensions as represented by any subset of k\(^-\) variables of g\(^+\). This result follows from the definition of g as a postulated linear function of the variables of z. That is, from (5.4.1), we have

\[
(6.2.2) \quad g = T^{-1}D^{-1}\(L^{-1}\)z
\]

Thus, since g\(^+\) is a linear function of z\(^+\) involving p\(^+\) variables and g\(^-\) is a linear function of z\(^-\) involving p\(^-\) variables where p\(^-\) < p\(^+\) (i.e., a subset of the variables of z\(^+\)), the corresponding k\(^-\) variables of g\(^+\) and g\(^-\) will necessarily be different.

It is instructive to further consider what, in general, will occur to the postulated relationships (6.2.1)a-c as p\(^-\) → p\(^+\). The expected results can be stated:

(6.2.3)a \quad \lim (\text{trace}^{-1}D^2_{d1}/p^-) = (\text{trace}^+D^2_{d1}/p^+) \quad \text{as} \quad p^- \rightarrow p^+

(6.2.3)b \quad \lim k^- = k^+ \quad \text{as} \quad p^- \rightarrow p^+

(6.2.3)c \quad \lim g^-_r = g^+_r \quad \text{as} \quad p^- \rightarrow p^+ \quad r = 1, \ldots, k^-

where "lim" is to denote a mathematical limit.
These results are equivalent to postulating that
\[(6.2.4) \lim_{p^- + p^+} \{z^- = z^-(1) + z^-(2)\} = \{z^+ = z^+(1) + z^+(2)\},\]
indicating that a general convergence of the two experimental situations is expected as the set \(Q^-\) "approaches" \(Q^+\) in terms of the number of semantic specifiers.

The convergence with respect of (6.2.3)a, it should be noted, need not necessarily occur in a continuous fashion. That is, the effect of adding an additional specifier to \(Q\) may possibly result in decreasing (trace \(D^2_{d1}/p\)) rather than increasing it if, for example, the additional specifier is highly associated with a dimension of specification not yet represented by the set \(Q\).

Another set of postulated relationships concerns comparison of \(Q^-\) with another subset, \(Q^X\), containing \(p^X\) elements where \(p^- = p^X < p^+\) and where some (but not all) of the elements of \(Q^X\) are also contained by \(Q^-\). For such a comparison, the following relationships are postulated as being highly probable:

\[(6.2.5)a \quad (\text{trace} D^2_{d1}/p^-) \neq (\text{trace}^X D^2_{d1}/p^X)\]
\[(6.2.5)b \quad k^- \neq k^X\]
\[(6.2.5)c \quad g^r_r \neq g^X_r \quad r = 1, \ldots, k^- \text{ if } k^- < k^X;\]
\[r = 1, \ldots, k^X \text{ if } k^X < k^-\]

That is, it would be expected that only by chance would the two experiments, \(E^-(Q^-, P, C)\) and \(E^X(Q^X, P, C)\),
yield identical results with respect to these characteristics.

6.3 Experimental Implications of Postulated Relationships

The postulated relationships with respect to the three segmentation characteristics suggest criteria in terms of which the notion of SC introduced in section 1.4 can be made more precise. Although no over-all quantitative measure exists so that a particular Q and associated E(Q, P, C) can be objectively evaluated in terms of SC(Q), three criteria are suggested as indices of such SC. Specifically, as between two sets, say Q^- and Q^x, it would seem logical to suppose that SC(Q^-) > SC(Q^x) if each of the following conditions hold:

(6.3.1a) \((\text{trace}^{-D_d^2/p^-}) \geq (\text{trace}^{x_d^2/p^x})\)

(6.3.1b) \(k^- \geq k^x\)

(6.3.1c) \((\sum_{r=1}^{k^-} |g^-_r - g^+_r| /k^-) \leq (\sum_{r=1}^{k^x} |g^x_r - g^+_r| /k^x)\)

Beyond this situation—say, for example, if only one or two of these conditions were met, it would be difficult to determine, in the absence of additional criteria, which set possessed the greatest amount of SC.

The first criterion suggests that the ratio of ESVR to total experimental variance should be as large as possible, or equivalently, that the amount of ineffective
specification should be minimized. The second criterion indicates that a set Q should be as representative of \( Q^+ \) as possible, given its size. Finally, the last criterion implies that the semantic factors represented by Q should approximate the corresponding factors of \( Q^+ \) to the extent possible.

These criteria indicate at least three factors which theoretically are relevant to the selection of a particular Q for the design of a given \( E(Q, P, C) \). As implied, the logical objective would be to, for a given size of Q, select the one for which the SC is maximum. The difficulty with the criteria insofar as their practical relevance to this design problem is that they are "after the fact." That is, an experiment must be conducted before their specific nature in a given case can be known. Moreover, in the absence of information pertaining to \( Q^+ \) and the best alternative Q, say \( Q^x \), which is the likely situation and the one faced by the investigation, evaluation of a particular Q, say \( Q^- \), and an associated \( E^-(Q^-, P, C) \), must necessarily depend only on \( k^- \), the trace \( \frac{D^2}{d_1} / p^- \) and any criteria external to the model.

7. The Model in Terms of Sample Observations

7.1 Statement of the Model

The model will now be stated in terms of a sample of \( N \) independent, \( p^- \)-dimensional observations arising from \( N \)
trials of the experiment, $E(Q, P, C)$. Letting the $pxN$ matrix $Z$ represent a sample of standardized observations and using "hat" notation to indicate sample estimates, the model appears:

\[ Z = \hat{Z}^{(1)} + \hat{Z}^{(2)} \]

where:

a. $p = \hat{k} + m$

b. $\hat{Z}_{pxN}^{(1)} = \hat{B}\hat{G}$

c. $\hat{Z}_{pxN}^{(2)} = \hat{L}^{(2)}\hat{D}_{d_2}$

d. $\hat{B}_{pxk} = \hat{A}\hat{T}$

e. $\hat{A}_{pxk} = \hat{L}^{(1)}\hat{D}_{d_1}$

f. $\hat{G}_{kxN} = \hat{T}\hat{X}^{(1)}$

g. $T^TT = TT^T = I_{kxk}$, a $kxk$ orthogonal transformation matrix defined according to (3.1)g

h. $\frac{1}{N - 1} ZZ^* = \hat{R}$, a $pxp$ unbiased estimate of $R$

i. $\hat{L}_{pxp} = (\hat{L}^{(1)}, \hat{L}^{(2)})$ represents a partitioning according to (2.3)a of a $pxp$ matrix of sample estimates of the unit-length eigenvectors of $R$
7.2 Properties of the Model

The properties of the model in the case of a sample of observations are analogous to those given in terms of the population in section 4. They will be stated without proof:

\( \hat{R} = \frac{1}{N - 1} ZZ' \)

\( = \frac{1}{N - 1} (\hat{z}(1)\hat{z}(1)' + \hat{z}(2)\hat{z}(2)') \)

\( = \hat{L}^{(1)} \hat{D}_d^{2} \hat{L}^{(1)}' + \hat{L}^{(2)} \hat{D}_d^{2} \hat{L}^{(2)}' \)

\( = \hat{D}_d^{2} \hat{L} \)

\( = \hat{M}(1) + \hat{M}(2) \)

\( \hat{z}(1)\hat{z}(2)' = \phi_{p \times p} \)
(iii) a. $Z \Gamma = B_{pxk}$

b. $Z(1) \Gamma = B_{pxk}$

c. $Z \chi(1) = A_{pxk}$

d. $Z \chi(2) = L(2) D_{d2}$

(iv) a. $\text{trace } M(1) = \text{trace } D_{d1}^2$

b. $\text{trace } M(2) = \text{trace } D_{d2}^2$

c. $\text{trace } R = \text{trace } D_d^2 = p$

7.3 Terminology for Sample Matrices

The terminology, some of which is analogous to that previously introduced (section 5.3), to be used with respect to the sample observations of a given $E(Q, P, C)$ is:

1. $Z$—pxN standardized semantic data matrix
2. $Z(1)$—pxN effective specification matrix
3. $Z(2)$—pxN residual specification matrix
4. $B$—pxk simple structure factor matrix
5. $A$—pxk initial factor matrix
6. $G$—kxN semantic score matrix
7. $\text{trace } D_{d1}^2$—amount of estimated ESVR
8. $\text{trace } D_{d2}^2$—amount of estimated RSVR
9. $k$—the estimated number of semantic factor variables
Of these sample elements, (1), (4), (6), (7), (8) and (9) are the most important of the basic output of E(Q, P, C).
CHAPTER III

PROBLEMS OF ESTIMATION

0. Introduction

The problems of estimation, associated with the model presented in Chapter II, are considered here. The estimation problems can be segmented into two parts: (1) the estimation of the parameters of the model—k, D and L and (2) the estimation of the semantic factor variables of g in terms of subsets of the semantic variables of y. Parts 1 and 2 discuss parameter estimation; part 3 considers semantic factor estimation.

1. Estimating $D_d^2$ and L

The sample estimates of the elements of $D_d^2$, the latent roots of R, for a sample of N independent, p-dimensional observations

$$z_i = \begin{pmatrix} z_{1i} \\ \vdots \\ \vdots \\ z_{pi} \end{pmatrix} \quad i = 1, \ldots, N$$

are defined as the roots, $\hat{d}_1^2, \ldots, \hat{d}_p^2$, of the determinantal equation
\[ | \hat{R} - d^2 I | = 0 \]

where \( \hat{R} \) (section 7, Chapter II) is an unbiased estimate of \( R \), the population correlation matrix.

The sample estimates of the corresponding unit-length eigenvectors of \( L \) are defined as the set of column vectors, \( \hat{L}(1), \ldots, \hat{L}(p) \), satisfying the equations

\[ (\hat{R} - d^2 I) \hat{L}(r) = \phi_{p \times 1} \]

and

\[ \hat{L}(r) \hat{L}(r) = 1 \quad r = 1, \ldots, p. \]

It can be shown (Anderson, 1958, p. 279) that if

(1) the distribution function, \( f \), associated with \( z \) (section 1.9, Chapter II) is assumed a multivariate normal,

(2) \( R \) is nonsingular and

(3) \( N > p \), then the sample estimates \( \hat{D}_d^2 \) and \( \hat{L} \) defined above are maximum likelihood estimates of \( D_d^2 \) and \( L \), respectively.

In the experiment reported in Chapter V, the sample estimates of \( L \) will not be given directly. Rather, the simple structure factor matrix

\[ \hat{B} = \hat{A} \]

\[ = \hat{L}(1) \hat{D}_d \hat{L}(1)^T \]

will be given. Since, as is seen, \( \hat{B} \) represents a transformation, \( T \), of \( (\hat{L}(1)^T \hat{D}_d \hat{L}(1)) \)--that is, of the matrix \( \hat{L}(1)^T \hat{D}_d \hat{L}(1) \) scaled by columns—and since \( T \) itself is a function of
(\hat{L}^{(1)}_{d_1}), reporting \hat{B} is equivalent to an indirect reporting of \hat{L}.

The elements of \(D_d^2\) will be reported in full.

2. Estimating the Number, \(k\), of Semantic Factor Variables

2.1 Nature of the Estimation Problem

As can be seen from the statement of the model in Chapter II (section 3), the point of segmentation of \(z\), or the number, \(k\), of semantic factor variables, is not unique. That is, \(k\) can take on any integer from 1 to \(p-1\), since \(p = k + m\) and \(k, m \neq 0\). Consequently, \(k\) must be estimated and there are \(p-1\) estimation possibilities.

A completely objective statistical framework is not available for estimating \(k\). Rather, a two-stage process is suggested in which in stage 1, an initial estimate, \(\hat{k}(i)\), is objectively determined according to a specified criterion. This is followed in stage 2 by a subjective determination of a final estimate, \(\hat{k}\), of \(k\) (where \(\hat{k} \leq \hat{k}(i)\)) by means of subjective interpretation of the linear pattern between the semantic variables of \(z\) and the semantic factors of \(g\). The suggested two-stage estimation process will now be discussed.
2.2 Initial Estimation

Let the initial estimate, \( \hat{k}(1) \), of \( k \) be the number of latent roots of \( R \) greater than one. In other words, consider only the transformation of those principal component variables with variances greater than unity and call these initial semantic factor variables. Letting \( g(1) \) represent the set of \( \hat{k}(1) \) initial semantic factor variables and using the notation "\((1)\)" to denote initial estimates, the model in terms of the segmentation of \( z \) is indicated:

\[
(2.2.1) \quad z = B(1)g(1) + z^{(2)}(1)
\]

where:

a. \( g(i) = T'(1)D^{-1}_d(i)u^{(1)}(1) \)

b. \( z^{(2)}(i) = L^{(2)}(i)u^{(2)}(i) \)

c. \( B(i) = L^{(1)}(i)D_d(i)T(i) \)

d. \( u^{(1)}(1) \) -- a \( \hat{k}(1)x1 \) random vector of principal component variables whose variances, \( d_1^2, \ldots, d_{\hat{k}(1)}^2 \), are greater than unity

e. \( u^{(2)}(1) \) -- a \( [p - \hat{k}(1)]x1 \) random vector of principal component variables whose variances, \( d_{\hat{k}(1)+1}^2, \ldots, d_p^2 \), are less than or equal to unity.

f. trace \( D_d^2(i) \) -- initial ESVR
The rationale underlying this step is the assumption that any principal component variable having a variance less than or equal to that of any one of the semantic variables of $z$ (i.e., less than or equal to unity) is not of any special significance. Consequently, it would not contribute, in any significant way, to a definition of $g$. That is, it is assumed that if each of the variables of $u^{(1)}(i)$ is "significant" insofar as indicated by its variance, then each of the variables of

$$g(i) = T' (1) D^{-1}_d (1) u^{(1)}(1)$$

is a potentially "significant" semantic factor variable.

(Stage 2)

2.3 **Initial Semantic Groups**

Let $i z(r), (r = 1, \ldots, k' \leq k(i))$, be a set of $n_r(i) \times 1$ random vectors representing subsets of the semantic variables of $z$, each containing $n_r(i)$ variables, $i z_j(r), (j = 1, \ldots, n_r(i) < p)$, having the following properties:

(i) The $i z_j(r)$ are those variables of $z$ which are semantically related.

(ii) The $i z_j(r)$ are relatively small in number as compared to $p$.

(iii) $i z_j(r) \neq i z_h(s)$, $r \neq s$, $j = 1, \ldots, n_r(i)$; $h = 1, \ldots, n_s(i)$, $r, s = 1, \ldots, k'$ (That is, a given variable belongs to only one subset, or none at all.)
(iv) The \( i z_j(r) \) have "significant" linear relationships only with \( g_r(i) \), the rth initial semantic factor variable of \( g(i) \).

(v) The linear relationships between the \( i z_j(r) \) and \( g_r(i) \) are such that

\[
(2.3.1) \quad \text{cov}[i z_1(r), g_r(i)] \geq \ldots \geq \text{cov}[i z_n_r(i), g_r(i)] \geq \text{cov}[z_j, g_r(i)]
\]

\[ j = 1, \ldots, p - n_r(1) \]

where the \( z_j \) are those variables of \( z \) not in \( i z(r) \).

The rth random vector, \( i z(r) \), representing the rth subset of \( z \), will be called the rth initial semantic group (ISG).

2.4 Final Estimation

Using this definition, define the final estimate, \( \hat{k} \), of \( k \) as the number, \( k' \), of ISG's. That is,

\[
\hat{k} = k' < k(i)
\]

The model at this final stage of estimation then becomes identical to that given in Chapter II (section 3). That is, at this final stage of estimation we have

\[
p = \hat{k} + m \quad \text{and}
\]

\[
(2.4.1) \quad z = Bg + z(2)
\]

where \( B \), \( g \) and \( z(2) \) are as defined in (3.1)b-h of Chapter II.
2.5 Final Semantic Groups

Corresponding to the definition of the ISG's, let \( z(r), (r = 1, \ldots, k = k' < k(1)) \), be a set of \( n_r \times 1 \) random vectors representing subsets of the semantic variables of \( z \), each containing \( n_r \) variables, \( z_j(r) (j = 1, \ldots, n_r < p) \), having the following properties:

(i) The \( z_j(r) \) are those variables of \( z \) which are semantically related.

(ii) The \( z_j(r) \) are relatively small in number as compared to \( p \).

(iii) \( z_j(r) \neq z_h(s) \) for \( j = 1, \ldots, n_r; h = 1, \ldots, n_s \), \( r, s = 1, \ldots, k \).

(iv) The \( z_j(r) \) have "significant" linear relationships only with \( g_r \), the \( r \)th semantic factor variable of \( g \).

(v) The linear relationships between \( z_j(r) \) and \( g_r \) are such that

\[
(2.5.1) \quad \text{cov}(z_1(r), g_r) \geq \ldots \geq \text{cov}(z_{n_r}(r), g_r) \geq \text{cov}(z_j, g_r) \\
\phantom{(2.5.1)} \quad j = 1, \ldots, p - n_r
\]

where the \( z_j \) are those variables of \( z \) not in \( z(r) \).

The \( r \)th random vector, \( z(r) \), representing the \( r \)th subset of \( z \), will be called the \( r \)th final semantic group (FSG).

It should be noted that the random vectors \( z(r) \) and \( z(r), (r = 1, \ldots, k' = k) \), need not necessarily contain the same number of semantic variables. That is,
But the content of \( _i z(r) \) and \( z(r) \) is expected to overlap or possibly be identical. More precisely it is expected that

\[
_i z(r) \subseteq z(r)
\]

(2.5.2) or

\[
_i z(r) \supseteq z(r) \quad r = 1, \ldots, k' = \hat{k}.
\]

If \( _i z(r) = z(r) \), \( (r = 1, \ldots, \hat{k} = k') \), and if \( \hat{k} = k' = \hat{k}(i) \), then it would follow that

\[
\begin{align*}
    z &= B(1)g(1) + z^{(2)}(1) \\
    &= Bg + z^{(2)}
\end{align*}
\]

which is to imply that the estimates of stage 1 and 2 are identical.

2.6 Terminology

It should be noted that

\[
\begin{align*}
    \sum_{r=1}^{k'} n_r(i) &= n^*(i) \leq p \quad \text{and} \\
    \sum_{r=1}^{\hat{k}} n_r &= n^* \leq p
\end{align*}
\]

which is to say that not all the variables of \( z \) are necessarily a part of a given semantic group.

A variable which can be classified in a semantic group (that is, the variables \( _i z_j(r) \) of \( _i z(r) \) or the variables \( z_j(r) \) of \( z(r) \)) will be called a pure identifiable semantic variable. A variable which cannot be classified in a semantic group but which is "significantly" related to
only one of the semantic factors (of \( g(i) \)) will be referred to as a pure unidentifiable semantic variable. Finally, a variable which cannot be classified in a semantic group or which is "significantly" related to two or more semantic factor variables (of \( g(i) \) or \( g \))—will be called a complex semantic variable.

2.7 Semantic Groups and Simple Structure

It is apparent from the discussion of simple structure in Chapter II (section 5.7) that the number of semantic groups (\( k' = \hat{k} \)) and the nature of these groups—the extent of semantic relatedness (property (i)), the number of variables of each group (property (ii)) and the "significance" of the relationships between the variables of the groups and the respective semantic factor variables (of \( g(i) \) or \( g \)) (property (iv))—depend upon the degree to which a "simple structure" pattern prevails between the variables of \( z \) and those of \( g(i) \) or \( g \). A "pure" simple structure pattern, as depicted in Figure 2, would yield a set of \( \hat{k}(i) \) semantic groups (i.e., \( \hat{k} = k' = \hat{k}(i) \)) in which all of the variables are contained—that is, it would yield a situation in which all of the variables are pure identifiable semantic variables. To the extent a "simple structure" pattern does not exist, some of the variables of \( z \) would be complex and the number, \( k' = \hat{k} \), of semantic groups would likely be less than \( \hat{k}(i) \).
Furthermore, as discussed in Chapter II (section 5.8), the extent to which a simple structure pattern exists and the nature of the pattern insofar as the semantic content of the semantic groups will affect the extent to which the semantic factors, either those of \( g(i) \) or \( g \), can be "meaningfully" interpreted or identified. And, as will be seen in part 3 of this chapter, it will also affect the degree to which "meaningful" estimation of the variables of \( g \) is possible.


3.1 Objective

The objective of the estimation method given below is to determine to what extent the semantic factor variables can be predicted by mutually exclusive and semantically related subsets of semantic variables. The method involves the least-squares criterion and the estimation of each semantic factor is treated as a separate problem.

3.2 Potential Predictor Variables

Let \( y(r), (r = 1, \ldots, k) \) be a set of \( n_r \times 1 \) random vectors representing subsets of the (unstandardized) semantic variables of \( y \), each containing \( n_r \), \( y_j(r), (j = 1, \ldots, n_r < n_r < p) \), such that
\( y(r) = [\text{var}(y_j(r))]^{1/2} z_j(r) \)

\( j = 1, \ldots, n_r^* \leq n_r \)

\( r = 1, \ldots, k \)

The \( y_j(r) \) variables of \( y(r) \) will be called potential semantic predictor variables of the \( r \)th semantic factor variable, \( g_r \), of \( g \).

3.3 Properties of Potential Predictor Variables

Since the variables of the \( y_j(r) \) are defined in (3.2.1) as simple scalings of a subset of the variables of the \( z \) (i.e., \( n_r^* < n_r \)), the equivalent of the properties (i)', (ii)', and (iii)' defining the variables of the \( z(r) \) (section 2.5) are relevant to the variables of the \( y(r) \). That is,

(i)' The \( y_j(r) \) are those variables of \( y \) which are semantically related.

(ii)' The \( y_j(r) \) are relatively small in number as compared to \( p \).

(iii)' \( y_j(r) \neq y_h(s) \quad r, s = 1, \ldots, n_r^* ; \quad h = 1, \ldots, n_s^* \)

The equivalent of properties (iv)' and (v)' is also applicable if it is assumed that

\[
\text{if } \text{cov}(z_i, g_r) \geq \text{cov}(z_j, g_r) \]

\[
\text{then } \text{var}(y_i) \geq \text{var}(y_j) \]

\( i, j = 1, \ldots, p \)

\( r = 1, \ldots, k \)
That is, assuming the condition stated in (3.3.1), then

(iv) The $y_j(r)$ have "significant" linear relationships only with $g_r$, the $r$th semantic factor variable of $g$.

(v) The linear relationships between the $y_j(r)$ and $g_r$ are such that

$$\text{cov}(y_1(r), g_r) \geq \ldots \geq \text{cov}(y_{n^*_r}(r), g_r) \geq \text{cov}(y_j, g_r)$$

where the $y_j$ are those variables of $y$ not in $y(r)$.

Property (iv) follows from the fact that

(3.3.2) $y = D_y^{-1/2} z$

from which we have

(3.3.3) $\text{cov}(y, g) = D_y^{-1/2} \text{cov}(z, g)$

$$= D_y^{-1/2} B$$

$$= \Sigma_{yg}, \text{ a } pxk \text{ covariance matrix.}$$

From (3.3.3), we see that

$$\text{cov}(y_j, g_r) = [\text{var}(y_j)]^{1/2} \text{cov}(z_j, g_r), r = 1, \ldots, k$$

which means that if the linear relationship between $z_j$ and $g_r$ is "significant" ("insignificant"), then the linear relationship between $y_j$ and $g_r$ is "significant" ("insignificant"). Thus, we have property (iv).

Property (v) follows from condition (3.3.1) by noting that since
\[ \text{cov}(z_1(r), g_r) \geq \ldots \geq \text{cov}(z_{n_r}(r), g_r) \geq \text{cov}(z_j, g_r), \]

then

\[ (\text{var } y_1(r))^{1/2} \text{cov}(z_1(r), g_r) \geq \ldots \]

\[ \ldots \geq (\text{var } y_{n_r}(r))^{1/2} \text{cov}(z_{n_r}(r), g_r) \]

\[ \geq (\text{var } z_j)^{1/2} \text{cov}(z_j, g_r) \]

which is equivalent to

\[ \text{cov}(y_1(r), g_r) \geq \ldots \geq \text{cov}(y_{n_r}(r), g_r) \geq \text{cov}(y_j, g_r) \]

or property (v) *.

The condition stated in (3.3.1) was noted in Chapter II (sections 5.5 and 5.8). The condition will be assumed in deriving estimation equations from the data of the experiment indicated in Chapter IV.

3.4 Notation

(a) \[ \begin{bmatrix} y(r) \\ g_r \end{bmatrix} \begin{bmatrix} y(r) \\ g_r \end{bmatrix}^{\prime} = E \begin{bmatrix} y(r)y(r)^{\prime} & y(r)g_r^{\prime} \\ g_r^{\prime}y & g_r \end{bmatrix} = \begin{bmatrix} \Sigma_{yy} & \Sigma_{yg} \\ \Sigma_{yg} & \Sigma_{gg} \end{bmatrix} \begin{bmatrix} r(r) \\ 1 \end{bmatrix} \]

(b) \( g_r \)--an estimation of the rth semantic factor variable, \( g_r \).
(c) $R_r^*$, $y(r)$—the multiple correlation coefficient between $g_r$ and $g_r^*$, $r = 1, \ldots, k$.

(d) $W(r)$—an $n_r \times 1$ column vector of regression weights for predicting the $r$th semantic factor with the $n_r^*$ semantic variables of $y(r)$.

3.5 **Linear Least-Squares Estimation**

This method involves determining a linear regression of $g_r$ on $y(r)$. The general model for such a linear estimation can be stated:

\[(3.5.1) \quad g_r = g_r^* + e_r \quad r = 1, \ldots, k\]

where

a. $g_r^* = W(r)^\top y(r)$

b. $W(r)^\top = \text{cov}(g_r, y(r)) \left[\text{cov}(y(r), y(r))\right]^{-1}$

   \[= E(g_r y(r)^\top) \left[E(y(r)y(r)^\top)\right]^{-1}\]

\[= \frac{r(r)^\top}{y g \quad y y}\]

(i.e., $W(r)^\top$ is a $l \times n_r^*$ row vector of regression weights for the regression of $g_r$ on $y(r)$, Anderson, 1958, p. 28).

c. $E(e_r^2) = \sigma^2$

d. $E(e_r) = 0$

3.6 **Properties of the Estimation**

From (3.5.1), the following properties are deduced (see Appendix B):
(i) \( \text{cov}(e_r, y(r)) = \varphi_{1 \times n_r} \), a null vector

(ii) \( \text{var}(g_r) = \sum_{yg} (r)^{-1} \sum_{yy} (r) + \sigma^2 = 1 \)

(iii) \( E(g^*_r g_r) = E(W(r)^* y(r) g_r) \)
     \[ = W(r)^* \sum_{yg} (r) \]
     \[ = \sum_{yg} (r)^{-1} \sum_{yy} (r) \]

(iv) \[ \frac{E(g^*_r g_r)}{[\text{var}(g_r)]^{1/2}[\text{var}(g_r)]^{1/2}} = \frac{[\sum_{yg} (r)^{-1} \sum_{yy} (r)]}{[\sum_{yg} (r)^{-1} \sum_{yy} (r)]^{1/2}} \]
     \[ = \frac{[\sum_{yg} (r)^{-1} \sum_{yy} (r)]^{1/2}}{[\sum_{yg} (r)^{-1} \sum_{yy} (r)]^{1/2}} \]
     \[ = R_r^* \cdot y(r) \]

(v) \( \text{var}(g_r) = R_r^* \cdot y(r) + \sigma^2 \)

or

\( \sigma^2 = 1 - R_r^* \cdot y(r) \)

indicating the portion of the variance not explained by the regression of \( g_r \) on \( y(r) \).

3.7 Evaluation of Method

The effectiveness of this method of estimation can be judged by these factors:

(1) The semantic relatedness of the semantic predictor variables of the \( y(r) \).
(2) The portion of the variance of the $g_r$ explained or predicted by the $g^*_r$ (i.e., the squared multiple-correlation coefficients, $R^*_r$).

(3) The number, $n^*_r$, of "significant" semantic predictor variables required for the estimation (where $n^*_r \leq n_r^* \leq n_r < p$).

Evaluating the predictor variables on the basis of these factors involves subjective judgement of the elements in a given experimental situation. Objective criteria are not available by which to determine the degree of relatedness, whether the portion of predicted variance is sufficient or whether the number of variables required is appropriate.

3.8 A Sequence of Significant Regression Equations

The number of "significant" predictor variables for a given variable of $g$ will depend on the statistical considerations incorporated in the computational procedure for determining the regression coefficients. A step-wise procedure is to be used which involves a t-test criterion for determining the significance of the regression coefficients (Johnston, p. 130). Making the appropriate normality assumption concerning the distribution of the $e_r$ (Johnston, pp. 115-116), a sequence of $n_r^*$ regression equations ($n_r^* \leq n_r$) is to be developed on the basis of the
t-test criterion such that each equation of the sequence contains a different number of significant predictor variables than the previous equation in the sequence and such that the same set of significant predictor variables appears only once. The "best" estimation equation of a given sequence will be that for which one of two things occur: (1) all the predictor variables have been considered as either significant or insignificant according to the t-test criterion or (2) the increase in the multiple correlation coefficient from going to the next equation in the sequence (i.e., adding another variable to the regression) is not considered large enough to justify the added variable.

For predicting the rth semantic factor, such a sequence of $n_r^*\cdash$ regression equations is indicated:

\[
g_r^* = \sum_{i=1}^{W_1(r)} \imath_{11}^{(r)} y_1(r) + \ldots + \sum_{i=n_1}^{W_1(r)} \imath_{n_1}^{(r)} y_{n_1}(r)
\]

\[
g_r^* = \sum_{i=1}^{W_2(r)} n_{11}^{(r)} y_1(r) + \ldots + \sum_{i=n_2}^{W_2(r)} n_{n_2}^{(r)} y_{n_2}(r)
\]

\[
g_r^* = \sum_{i=1}^{W_3(r)} n_{11}^{(r)} y_1(r) + \ldots + \sum_{i=n_3}^{W_3(r)} n_{n_3}^{(r)} y_{n_3}(r)
\]

\[
g_r^* = \sum_{i=1}^{W_4(r)} n_{11}^{(r)} y_1(r) + \ldots + \sum_{i=n_4}^{W_4(r)} n_{n_4}^{(r)} y_{n_4}(r)
\]

\[
g_r^* = \sum_{i=1}^{W_5(r)} n_{11}^{(r)} y_1(r) + \ldots + \sum_{i=n_5}^{W_5(r)} n_{n_5}^{(r)} y_{n_5}(r)
\]

where it is assumed that (1) each of the sets of regression weights of each of $n_r^*$ equations in the sequence is
significant according to the t-test criterion,
(2) $n^*_r \leq n^{**}_r \leq n^*_r$ and (3) the $n^*_r$th element of the sequence
represents that point for which no further "significant"
increase in the multiple correlation coefficient is possi­
ble. Given these assumptions, the $n^*_r$th equation in the
above sequence would represent the "best" regression equa­
tion for the prediction of $g_r$—"best" in the sense of being
"significant" according to the t-test criterion and
"efficient" in the sense of accounting for or explaining
as much of the variance of $g_r$ with as few predictor
variables as possible.

In Chapter VII, estimation equations will be
developed for the data of the experiment described in
Chapter IV.

4. Summary of Terminology of Chapter III

The following is a summary of the terminology
introduced in Chapter III:
(1) $B(1)$—initial simple structure factor matrix
    (section 2.2).
(2) $g(1)$—a $k(1)x1$ vector of initial semantic factor
    variables (section 2.2).
(3) $\hat{k}(1)$—the initial estimate of $k$ (section 2.2).
(4) $T(1)$—the initial transformation matrix, $T$
    (section 2.2).
(5) $\text{trace}[D_1^2]$ (i) --- initial ESVR (section 2.2).

(6) $z(r)$ -- an $n_r \times 1$ vector representing the $r$th ISG (section 2.3).

(7) ISG -- initial semantic group (section 2.3).

(8) $k^*$ -- the number of ISG's (section 2.3).

(9) $\hat{k}$ -- the final estimate of $k$ (section 2.4).

(10) $z(r)$ -- an $n_r \times 1$ vector representing the $r$th FSG (section 2.5).

(11) FSG -- final semantic group (section 2.5).

(12) $y(r)$ -- an $n_r' \times 1$ vector ($n_r' < n_r$) of potential predictor variables (section 3.2).

(13) $g_r^*$ -- the least-squares estimation of $g_r$ (section 3.4).

(14) $W^{(r)}$ -- an $n_r' \times 1$ vector of regression weights for predicting $g_r$ with the variables of $y(r)$ (section 3.4).

(15) $n_r''$ -- the number of significant predictors according to $t$-test criterion ($n_r'' < n_r'$) (section 3.8).

(16) $n_r^*$ -- the number of significant predictor variables of "best" estimation equation ($n_r^* < n_r''$) (section 3.8).
CHAPTER IV

AN OPERATIONAL EXPERIMENT

0. Introduction

In Chapter II (section 1.6), an "idealized" experiment, E(Q, P, C), was defined involving the semantic specification by a subject from P of a perfect representation, \( \bar{c} \), of C on each of the p specifiers of Q. This chapter will define an operational experiment of the form E(Q, P, C)—one that is experimentally possible and for which the elements of P, C and Q are given specific contextual definition.

1. Contextual Definitions of P, C and Q

1.1 A Population Domain of Auditors

Let the population domain, P, be a finite set of auditors from the audit staff of a national CPA firm. It is intended that P encompass non-CPA's as well as CPA's, but exclude auditors whose primary function is non-audit (e.g., management or tax services).

Let \( P^a \), containing N elements, be a subset of P representing a particular experimental sample of auditors.
1.2 A Concept Domain of Audit-Technique Concepts

Let the concept domain, $C$, be a finite set of semantic concepts that represent that segment of the task environment of $P$ relating to auditing methodology—specifically, the techniques by which audit evidence is obtained. The concepts of $C$ will be called audit-technique concepts. Let $c_k$ be the $k$th concept of $C$ where $k = 1, \ldots, m$.

Let $C^a$, containing $h$ elements, ($h \leq m$), be a subset of $C$ representing a particular experimental sample of audit-technique concepts.

1.3 A Specifier Domain of Semantic (Specification) Scales

Let the specifier domain, $Q^+$, relevant to $C$ and $P$ be a finite set of semantic (specification) scales. A semantic scale is a seven-interval rating scale identified by a pair of polar (semantically opposite-in-meaning) adjectives positioned at the ends of the scale.

Let $Q$, containing $p$ elements, be a subset of $Q$ (where $Q \subseteq Q^+$, $p \leq p^+$) representing a particular experimental sample of semantic (specification) scales.

The semantic scale was developed by Osgood and his associates (Osgood, 1957) for use in a general psychological instrument, called the semantic differential, designed to quantitatively measure the "meaning" of concepts, objects or events for individuals. As will be indicated
(section 2.2), such an instrument, in the form of a questionnaire, constitutes the empirical framework of the operational experiment (section 2.1).

The response categories, $r_{c_t}$ ($t = 1, \ldots, 7$), of a semantic scale are defined by how they relate to the pair of polar adjectives (Osgood, pp. 28-29):

- $r_{c_1}$ very closely related to $X$
- $r_{c_2}$ quite closely related to $X$
- $r_{c_3}$ only slightly related to $X$
- $r_{c_4}$ neither $X$ nor $Y$
- $r_{c_5}$ only slightly related to $Y$
- $r_{c_6}$ quite closely related to $Y$
- $r_{c_7}$ very closely related to $Y$

The way in which these scales are to be experimentally used will depend on the set of experimental instructions, to be discussed in section 2.2.

2. An Operational Experiment and Sample Space

2.1 Definition

For a given $Q$, let $E^0(Q, P, C)$ represent an operational experiment— one that is experimentally possible.
Define a given trial of $E^0(Q, P, C)$ as one in which an auditor of $P$ specifies, according to a certain set of instructions (see section 2.2), the $k$th concept, $c_k$, of $C$ using each of the $p$ semantic scales of $Q$. Let $S^0$ be a set representing all possible specification response outcomes for a given trial of $E^0(Q, P, C)$. That is $S^0$ will denote the sample space of $E^0(Q, P, C)$ where the ordered $p$-tuple, 

$$s^0_k = (s^0_{1k}, \ldots, s^0_{jk}, \ldots, s^0_{pk}),$$

is an element of $S^0$. The $j$th component, $s^0_{jk}$, of $s^0$ represents the specification of the $k$th concept of $C$ on the $j$th semantic scale, $q_j$, of $Q$.

The behavior elicited from a given trial of $E^0(Q, P, C)$ will be called the **semantic specification behavior** of an auditor of $P$ relevant to a concept of $C$.

### 2.2 Administrative Form of $E^0(Q, P, C)$

The administrative form of $E^0(Q, P, C)$ is a questionnaire consisting of the concepts to be specified, the semantic scales by means of which specification is to be made and a set of specification instructions. The instructions indicate to the auditor how he is to experimentally use the scales in specifying the concepts.

The instruction set used for the current investigation is given in Appendix C. It represents an adaptation of a set of instructions suggested by
Osgood (pp. 82-84) for general use in a semantic
differential instrument.

2.3 An Actual Experiment

Let \( E^O(Q, P^a, C^a) \) represent an actual experiment of
the form \( E^O(Q, P, C) \) involving the experimental samples \( Q, P^a \) and \( C^a \). The behavior elicited from a given trial of
\( E^O(Q, P^a, C^a) \) will be called experimental semantic specifi­
cation behavior of an auditor from \( P^a \) relevant to a concept
from \( C^a \). In Chapter V, a particular set of samples—\( Q, P^a \)
and \( C^a \)—used in the investigation is indicated.

3. Definition of Semantic Variables

3.1 Experimental Semantic (Specification)

Variables

Let \( f_{y_k} \) be a function defined on \( S^o \):

\[
y_k = f_{y_k}(s^o) = \begin{pmatrix} y_{1k} \\ \vdots \\ y_{jk} \\ \vdots \\ y_{pk} \end{pmatrix}
\]

where \( y_{jk} = f_{y_k}(s^o_{jk}) \in R^I = \{ I \mid I = 1, \ldots, 7 \} \). That is,
\( y_k \) is defined as a px1 random vector of (discrete) random
variables assumed observable for any trial of \( E^O(Q, P, C) \)
and for which the possible values are positive integers
ranging from 1 to 7. The random variables of \( y_k \) will be
called experimental semantic (specification) variables.
3.2 Analytical Semantic (Specification) Variables

Let \( \vec{y} \) be a pxl random vector defined:

\[
\vec{y} = \sum_{k=1}^{h} y_k = \begin{pmatrix}
\vec{y}_1 \\
\vdots \\
\vec{y}_j \\
\vdots \\
\vec{y}_p
\end{pmatrix}
\]

where \( \vec{y}_j = \sum_{k=1}^{h} y_{jk} \in \mathbb{R}^I = \{I \mid I = 1, \ldots, (hx7) \} \) and where \( h \leq m \). That is, \( \vec{y} \) is defined as a pxl random vector of (discrete) random variables derivable from \( h \) repeated trials of \( E^0(Q, P, C) \) for any given subject of \( P \). The possible values of the random variables \( \vec{y}_j \) are the positive integers ranging between 1 and \( (hx7) \). The random variables of \( \vec{y} \) will be called analytical semantic (specification) variables.

As before (Chapter II, section 1.9), it will be assumed that corresponding to the variables of \( y_k \) is a multivariate (discrete) distribution function, \( f_e \), defined on \( S^0 \). It will also be assumed that the first and second moments of the marginal distributions of \( f_e \) for each of the \( y_{jk} \) exist and that \( E(y_k) = [E(y_{jk})] = \phi_{pxl} \), a pxl null vector.

Although experimentally, \( \vec{y} \) and \( y_k \) are random vectors of discrete random variables, they will be treated, theoretically, as if they were continuous random variables.
Such treatment reflects the assumption that the semantic specification of concepts by subjects involves a continuum—that an "ideal" experimental semantic scale would be continuous and allow the subject an infinite number of possible specification choices. The fact that this assumed theoretical state cannot be precisely duplicated experimentally reflects upon the accuracy of $E^0(Q, P, C)$ as an approximation of $E(Q, P, C)$.

It will be further assumed that $\bar{y}$ is a numerical approximation of a specification response of a given auditor of $P$ that would occur if the auditor were requested to specify an "idealized" or "typical" representation of the concepts of $C$ using each of the $p$ semantic scales of $Q$. That is, $\bar{y}$ is the experimental approximation of the random vector $y$ which was defined (Chapter II, section 1.9) as a function on the "idealized" sample space $S$ of the "idealized" experiment $E(Q, P, C)$, involving specification of $\bar{c}$, defined as a "perfect" or "typical" representation of the concepts of $C$.

3.3 Standardized (Analytical) Semantic Variables

As before (Chapter II, section 1.10), define a new random vector $\tilde{z}$:

$$\tilde{z} = D^{-1/2}y$$
where:

\[ D_y = \text{diag}(\Sigma_{yy}) \]

\[ \text{var}(\tilde{y}) = E(\tilde{y}\tilde{y}') = \Sigma_{\tilde{y}\tilde{y}} \text{, the p}\times\text{p covariance matrix of } \tilde{y} \]

\[ \text{var}(\tilde{z}) = E(\tilde{z}\tilde{z}') = D^{-1/2}\Sigma_{\tilde{y}\tilde{y}}D^{-1/2} = R, \text{ the p}\times\text{p nonsingular, symmetric population correlation matrix.} \]

That is, \( \tilde{z} \) is defined as a pxl random vector of standardized (discrete) random variables derivable from \( h \) repeated trials of \( E^0(\Omega, \Pi, \chi) \) for any given subject of \( \Pi \) and having zero means and unit variances. The random variables of \( \tilde{z} \) will be called standardized (analytical) semantic variables.

4. The Model for the Operational Experiment

4.1 A Statement of the Model for \( E^0(\Omega, \Pi, \chi) \)

The statement of the model for \( h \) repeated trials of \( E^0(\Omega, \Pi, \chi) \) is exactly analogous to the model for \( E(\Omega, \Pi, \chi) \) of Chapter II (section 3). At this point, it will be partially stated for completeness:

\[ \tilde{z} = \tilde{z}(1) + \tilde{z}(2) = \tilde{E}g + \tilde{E}(2)D_{a^2}\tilde{x}(2) \]

where, as before, the random vector \( \tilde{z} \) is split into two orthogonal parts, \( \tilde{z}(1) \) and \( \tilde{z}(2) \), where \( \tilde{z}(1) \), representing
effective specification, is a linear function of the semantic factors of \( \bar{g} \) and \( \bar{z}^{(2)} \), representing ineffective specification, is a linear function of the variables of the random vector \( \bar{x}^{(2)} \). For the complete statement of the model, see Chapter II (section 3).

It will be noted that the elements in (4.1.1) are distinguished from their counterparts in Chapter II by the use of "bar" notation. This notation is used in all subsequent chapters since in these chapters, \( E^0(Q, P, C) \), not \( E(Q, P, C) \), is involved.

4.2 A Statement of the Model for \( E^0(Q, P^a, C^a) \)

The statement of the model for \( E^0(Q, P^a, C^a) \), a particular experiment involving \( h \) samples of \( N \) independent, \( p \)-dimensional observations arising from \( N \times h \) trials of \( E^0(Q, P^a, C^a) \), is exactly analogous to the model for a sample of observations involving \( E(Q, P, C) \) (Chapter II, section 7.1). At this point a partial statement of the model will be given for completeness, using (as before) "hat" notation to indicate sample estimates:

\[
\hat{\bar{z}} = \hat{\bar{z}}^{(1)} + \hat{\bar{z}}^{(2)} = \hat{\bar{B}} \hat{\bar{z}}^{(2)} + \hat{\bar{L}} \hat{\bar{m}} \hat{\bar{m}} N
\]

where \( \bar{z}_N \), \( \hat{\bar{z}}^{(1)} \), \( \hat{\bar{z}}^{(2)} \), \( \hat{\bar{B}} \), and \( \hat{k} \) are as defined in section 7 of Chapter II.
5. **Summary of Terminology of Chapter IV**

The following is a summary of the terminology introduced in Chapter IV.

1. \( P^a \) -- an experimental sample of auditors from \( P \) (section 1.1).
2. \( C^a \) -- an experimental sample of audit-technique concepts from \( C \) (section 1.2).
3. \( Q \) -- an experimental sample of semantic scales from \( Q^+ \) (section 1.3).
4. \( E^o(Q, P, C) \) -- an operational experiment (section 2.1).
5. \( E^o(Q, P^a, C^a) \) -- an actual experiment involving the samples \( Q, P^a \) and \( C^a \) (section 2.3).
6. \( y_k \) -- a pxl vector of experimental semantic (specification) variables (section 3.1).
7. \( \tilde{y} \) -- a pxl vector of analytical semantic (specification) variables (section 3.2).
8. \( \tilde{z} \) -- a pxl vector of standardized (analytical) semantic variables (section 3.3).
CHAPTER V

AN ACTUAL EXPERIMENT INVOLVING 382 AUDITORS,
50 SEMANTIC SCALES AND 3 AUDIT-
TECHNIQUE CONCEPTS

0. Introduction

This chapter describes an actual experiment and the results. In part 1, the experimental samples, $P^a$, $C^a$ and $Q$, used in the experiment are indicated. In part 2, the administration of the experiment is discussed. And in part 3, the results of the experiment are presented.

1. The Experimental Samples: $P^a$, $C^a$ and $Q$

1.1 An Experimental Sample of Auditors ($P^a$)

A sample, $P^a$, of 382 auditors was selected from the domain, $P$, of auditors of the audit staff of Ernst & Ernst, a national CPA firm. The selection was made, using a modified random sampling procedure, from the six organizational categories of the firm and from 29 of its 103 offices, covering three office-size categories. The six organizational categories are: partner, manager, supervisor, senior, in-charge and staff accountant. The office size
categories are: large (over 100), medium (between 30 and 100) and small (under 30).

The distribution of the sample over the six organizational categories is:

<table>
<thead>
<tr>
<th>Organizational Category</th>
<th>Number of Auditors Sampled from Each Organizational Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner</td>
<td>60</td>
</tr>
<tr>
<td>Manager</td>
<td>66</td>
</tr>
<tr>
<td>Supervisor</td>
<td>73</td>
</tr>
<tr>
<td>Senior</td>
<td>48</td>
</tr>
<tr>
<td>In-Charge</td>
<td>59</td>
</tr>
<tr>
<td>Staff</td>
<td>76</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>382</strong></td>
</tr>
</tbody>
</table>

The distribution of the sample over the three office-size categories and the number of offices of each category sampled are:

<table>
<thead>
<tr>
<th>Office Size</th>
<th>Number of Offices Sampled</th>
<th>Number of Auditors Sampled from Each Office-Size Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (over 100)</td>
<td>3</td>
<td>111</td>
</tr>
<tr>
<td>Medium (between 30 and 100)</td>
<td>6</td>
<td>94</td>
</tr>
<tr>
<td>Small (under 30)</td>
<td>20</td>
<td>177</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>29</strong></td>
<td><strong>382</strong></td>
</tr>
</tbody>
</table>
1.2 An Experimental Sample of Audit-Technique Concepts (C^a)

A sample, C^a, of three concepts was selected from the domain, C, of audit-technique concepts. The concepts selected are:

(1) Physical Examination & Count
(2) Inquiry
(3) Confirmation

The concepts were selected from a listing of ten audit-technique concepts given by Mautz (p. 100). The concepts not chosen are: examination of authoritative documents and comparison with the record, recomputation, retracing bookkeeping procedures, scanning, examination of subsidiary record, correlation with related information and observation. The entire set of ten concepts, although not a complete or unique representation of C, appears to be a reasonably comprehensive representation from which to select a sample.

The three concepts selected were chosen primarily due to their importance and relevance in the auditing profession as evidenced by their inclusion in one of the standards of field work. Specifically, the third standard is stated:

Sufficient competent evidential matter is to be obtained through inspection, observation, inquiries, and confirmations to afford a reasonable basis for an opinion regarding the financial statements under
examination (AICPA, Statement on Auditing Procedure No. 33).

It would seem that on the basis of this professional standard, the concepts selected represent related and relevant parts of the auditor's professional language system to which he can and does systematically respond (see Chapter II, section 1.2, for the abstract definition of C).

The selection of three concepts rather than some larger number was made for purposes of limiting the response time of each subject to about 15 minutes.

Each of the concepts was accompanied in the questionnaire by an illustrative audit procedure to help clarify, as indicated in the instruction set (Appendix C), the intended nature of the audit technique represented by the concept. A listing of the three concepts and illustrative audit procedures is given in Appendix D.

1.3 An Experimental Sample of Semantic (Specification) Scales (Q)

A sample, Q, of 50 semantic (specification) scales was selected from the domain, Q⁺, of semantic scales. The selection of polar adjectives by which the scales are identified consisted of two phases: (1) a systematic search of Roget's Thesaurus (revised edition, 1962) covering all the logical categories of classification included in the Thesaurus, and (2) a pre-study investigation of the
polar adjectives obtained in phase 1 for purposes of selecting a final set of 50 adjective pairs.

The result of the search of the Thesaurus was an initial listing of approximately 400 polar adjectives. This listing was reduced to 110 adjective pairs for use in phase 2.

The criteria of this selection and reduction process were: (1) relevancy of the specifiers for specifying audit-technique concepts and (2) the representativeness of the specifiers insofar as representing as many of the semantic dimensions associated with $Q^+$ as possible.

The selection and reduction of this phase was admittedly subjective. The criteria of relevancy and representativeness could not be objectively incorporated into the process, although an attempt was made to be as consistent as possible. To introduce some objectivity into the final selection phase (phase 2), the 110 semantic scales resulting from phase 1 were used in a pre-study experiment involving nine persons, each having at least one year of auditing experience, from the accounting department at Ohio State. Each subject was asked to specify five audit-technique concepts on each of the 110 semantic scales of phase 1. The specification was done according to the instruction set given in Appendix C.

The five concepts specified by each subject were randomly selected from Mautz's listing of ten
audit-technique concepts given in the previous section. This resulted in each subject specifying a different set of concepts. The reason for selecting five rather than using all ten of the concepts was to limit the completion time of each pre-study subject to no more than an hour. It is assumed that a random sample of five provided a reasonable simulation of the stimuli represented by the entire set of the ten concepts.

The reduction of these 110 scales to a final list of 50 was done by computing the sample variance of the specification responses on each of the 110 scales over the subjects and concepts of the pre-study. On the basis of this variance, the scales were ranked from high to low and the top 50 selected as Q.

A complete listing of these 50 semantic scales is given in Appendix E.

The reasoning behind this procedure is that the response variance of a scale is assumed to be an index of its relevance to the subjects in specifying the concepts. The distribution of responses for a relevant scale under this assumption would be expected to appear as a bi-modal distribution:
On the other hand, the distribution of responses for an irrelevant scale would be expected to exhibit a disproportionate number of responses in the center of the scale:

The variance of a scale with a bi-modal distribution is obviously greater than the variance of a scale with this second distribution and according to the variance criterion, a scale with a bi-modal distribution would be ranked first in terms of assumed relevance.

It should be noted that the location of central tendency of a scale's distribution is not necessarily taken into account by the variance criterion as perhaps it should be. For example, consider a third distribution drawn:
This distribution is drawn so as to exhibit a variance equal to the variance of the second distribution above. Both of the associated scales on the basis of this variance would be ranked the same and considered irrelevant. But it is clear that the scale with this third distribution is more relevant than the scale of the second distribution, although probably not as relevant as the scale with the first distribution. Fortunately, the data of the pre-study indicate that the central tendencies of the response distributions of the scales over the concepts for each subject are, for the most part, at the center of the scales. This indicates that use of the variance criterion is valid. These data together with the standard deviations of the response distributions by subject are indicated:
### Means of Response Distributions Taken Over Concepts

<table>
<thead>
<tr>
<th>Subject</th>
<th>Means of Response Distributions Taken Over Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+.09</td>
</tr>
<tr>
<td>2</td>
<td>+.11</td>
</tr>
<tr>
<td>3</td>
<td>-.054</td>
</tr>
<tr>
<td>4</td>
<td>+.054</td>
</tr>
<tr>
<td>5</td>
<td>-.014</td>
</tr>
<tr>
<td>6</td>
<td>+.07</td>
</tr>
<tr>
<td>7</td>
<td>+.09</td>
</tr>
<tr>
<td>8</td>
<td>+.072</td>
</tr>
<tr>
<td>9</td>
<td>-.010</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Standard Deviations of Response Distributions Taken Over Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>1.99</td>
</tr>
<tr>
<td></td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>1.06</td>
</tr>
<tr>
<td></td>
<td>2.45</td>
</tr>
<tr>
<td></td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>1.28</td>
</tr>
<tr>
<td></td>
<td>1.67</td>
</tr>
</tbody>
</table>

2. **Administration of $E^0(Q, P^a, C^a)$**

Administration of $E^0(Q, P^a, C^a)$ consisted of having the auditors of $P^a$ (section 1.1) complete a questionnaire (Chapter IV, section 2.2) in which they were asked to specify, according to a set of instructions (Appendix C), each of three audit-technique concepts of $C^a$ (Appendix D) on each of the 50 semantic scales of $Q$ (Appendix E).

Questionnaires totaling 420 were sent in bulk shipments to 30 offices selected for the sample. In each office, a person of authority, which had been designated by the partner of that office, distributed, collected, and returned the questionnaires to the writer within a designated time period. Specific instructions were included in
each shipment to each office indicating the number of questionnaires enclosed, to whom they were to be distributed and when they were to be collected and returned.

As indicated in section 1.1, Pa, the final sample obtained, consisted of 382 auditors. This represents just above a 90% return of the questionnaires sent and reflects the extent of cooperation given by Ernst & Ernst to the experiment.

The distribution of the mailed questionnaires over the organizational and size categories, and the number of offices involved in each size category are:

<table>
<thead>
<tr>
<th>Organizational Category</th>
<th>Number of Questionnaires Mailed to Each Organizational Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner</td>
<td>75</td>
</tr>
<tr>
<td>Manager</td>
<td>78</td>
</tr>
<tr>
<td>Supervisor</td>
<td>76</td>
</tr>
<tr>
<td>Senior</td>
<td>56</td>
</tr>
<tr>
<td>In-Charge</td>
<td>53</td>
</tr>
<tr>
<td>Staff</td>
<td>82</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>420</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Office Size</th>
<th>Number of Offices to which Questionnaires were Mailed</th>
<th>Number of Questionnaires Mailed to Each Office-Size Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (over 100)</td>
<td>3</td>
<td>142</td>
</tr>
<tr>
<td>Medium (between 30 and 100)</td>
<td>6</td>
<td>94</td>
</tr>
<tr>
<td>Small (under 30)</td>
<td>21</td>
<td>184</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>30</strong></td>
<td><strong>420</strong></td>
</tr>
</tbody>
</table>
These data can be compared to those in section 1.1 relating to the actual sample, $p^a$.

3. Experimental Results

3.1 Estimated Latent Roots of $\hat{R}$

From the observations of $E^o(Q, p^a, c^a)$, the sample estimates, $\hat{\delta}_{ij}^2$, $(i = 1, \ldots, 50)$, of the latent roots of $\hat{R}$, the correlation matrix, are given in Table 1. Also indicated are: (1) the ratios of the latent root estimates to $p$, the total experimental variance of the observations from $E^o(Q, p^a, c^a)$ and (2) the ratios of the cumulative sums of these estimates to $p$.

3.2 Initial Estimate, $\hat{k}(i)$, of the Number, $k$, of Semantic Factor Variables

According to the two-stage estimation procedure of Chapter III (section 2.2), the initial estimate, $\hat{k}(i)$, of the number, $k$, of semantic factor variables is given by the number of estimated roots of $\hat{R}$ greater than unity. Accordingly, from Table 1, $\hat{k}(1) = 11$. Also, from Table 1, the estimated initial ESVR, given by

$$\text{trace} \hat{\delta}_{d_1}^2 (1) = \frac{11}{5} \hat{\delta}_{d_1}^2,$$

is 28.90. The proportion of this estimated variance to the total experimental variance of the observations from $E^o(Q, p^a, c^a)$, given by $\text{trace} \hat{\delta}_{d_1}^2 (1)/50$, is .578.
### TABLE 1

<table>
<thead>
<tr>
<th>Latent Root</th>
<th>Estimated Latent Roots ($d_1^2$)</th>
<th>Ratios of Estimated Latent Roots to p ($d_1^2/p$)</th>
<th>Ratios of Cumulative Sums ofLatent Root Estimates to p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.45</td>
<td>0.209</td>
<td>0.209</td>
</tr>
<tr>
<td>2</td>
<td>4.31</td>
<td>0.086</td>
<td>0.299</td>
</tr>
<tr>
<td>3</td>
<td>2.58</td>
<td>0.052</td>
<td>0.347</td>
</tr>
<tr>
<td>4</td>
<td>1.99</td>
<td>0.040</td>
<td>0.387</td>
</tr>
<tr>
<td>5</td>
<td>1.94</td>
<td>0.038</td>
<td>0.425</td>
</tr>
<tr>
<td>6</td>
<td>1.62</td>
<td>0.033</td>
<td>0.458</td>
</tr>
<tr>
<td>7</td>
<td>1.37</td>
<td>0.027</td>
<td>0.485</td>
</tr>
<tr>
<td>8</td>
<td>1.28</td>
<td>0.026</td>
<td>0.511</td>
</tr>
<tr>
<td>9</td>
<td>1.16</td>
<td>0.023</td>
<td>0.534</td>
</tr>
<tr>
<td>10</td>
<td>1.13</td>
<td>0.023</td>
<td>0.557</td>
</tr>
<tr>
<td>11</td>
<td>1.07</td>
<td>0.021</td>
<td>0.578</td>
</tr>
<tr>
<td>12</td>
<td>1.00</td>
<td>0.020</td>
<td>0.598</td>
</tr>
<tr>
<td>13</td>
<td>.98</td>
<td>0.020</td>
<td>.618</td>
</tr>
<tr>
<td>14</td>
<td>.96</td>
<td>0.019</td>
<td>.637</td>
</tr>
<tr>
<td>15</td>
<td>.91</td>
<td>0.018</td>
<td>.655</td>
</tr>
<tr>
<td>16</td>
<td>.87</td>
<td>0.017</td>
<td>.672</td>
</tr>
<tr>
<td>17</td>
<td>.86</td>
<td>0.017</td>
<td>.689</td>
</tr>
<tr>
<td>18</td>
<td>.80</td>
<td>0.016</td>
<td>.705</td>
</tr>
<tr>
<td>19</td>
<td>.78</td>
<td>0.016</td>
<td>.721</td>
</tr>
<tr>
<td>20</td>
<td>.76</td>
<td>0.015</td>
<td>.736</td>
</tr>
<tr>
<td>21</td>
<td>.72</td>
<td>0.014</td>
<td>.750</td>
</tr>
<tr>
<td>22</td>
<td>.70</td>
<td>0.015</td>
<td>.765</td>
</tr>
<tr>
<td>23</td>
<td>.67</td>
<td>0.013</td>
<td>.778</td>
</tr>
<tr>
<td>24</td>
<td>.66</td>
<td>0.013</td>
<td>.791</td>
</tr>
<tr>
<td>25</td>
<td>.64</td>
<td>0.013</td>
<td>.804</td>
</tr>
<tr>
<td>26</td>
<td>.63</td>
<td>0.012</td>
<td>.816</td>
</tr>
<tr>
<td>27</td>
<td>.61</td>
<td>0.015</td>
<td>.829</td>
</tr>
<tr>
<td>28</td>
<td>.60</td>
<td>0.012</td>
<td>.841</td>
</tr>
<tr>
<td>29</td>
<td>.55</td>
<td>0.011</td>
<td>.852</td>
</tr>
<tr>
<td>30</td>
<td>.53</td>
<td>0.010</td>
<td>.862</td>
</tr>
<tr>
<td>31</td>
<td>.49</td>
<td>0.010</td>
<td>.872</td>
</tr>
<tr>
<td>32</td>
<td>.48</td>
<td>0.010</td>
<td>.882</td>
</tr>
<tr>
<td>33</td>
<td>.47</td>
<td>0.009</td>
<td>.891</td>
</tr>
<tr>
<td>34</td>
<td>.44</td>
<td>0.009</td>
<td>.890</td>
</tr>
<tr>
<td>35</td>
<td>.43</td>
<td>0.009</td>
<td>.899</td>
</tr>
<tr>
<td>36</td>
<td>.42</td>
<td>0.008</td>
<td>.907</td>
</tr>
<tr>
<td>37</td>
<td>.41</td>
<td>0.008</td>
<td>.917</td>
</tr>
<tr>
<td>38</td>
<td>.39</td>
<td>0.008</td>
<td>.925</td>
</tr>
<tr>
<td>39</td>
<td>.39</td>
<td>0.008</td>
<td>.933</td>
</tr>
<tr>
<td>40</td>
<td>.37</td>
<td>0.007</td>
<td>.941</td>
</tr>
<tr>
<td>41</td>
<td>.35</td>
<td>0.007</td>
<td>.948</td>
</tr>
<tr>
<td>42</td>
<td>.35</td>
<td>0.007</td>
<td>.955</td>
</tr>
<tr>
<td>43</td>
<td>.32</td>
<td>0.006</td>
<td>.962</td>
</tr>
<tr>
<td>44</td>
<td>.30</td>
<td>0.006</td>
<td>.968</td>
</tr>
<tr>
<td>45</td>
<td>.27</td>
<td>0.005</td>
<td>.974</td>
</tr>
<tr>
<td>46</td>
<td>.27</td>
<td>0.005</td>
<td>.979</td>
</tr>
<tr>
<td>47</td>
<td>.23</td>
<td>0.004</td>
<td>.985</td>
</tr>
<tr>
<td>48</td>
<td>.20</td>
<td>0.004</td>
<td>.989</td>
</tr>
<tr>
<td>49</td>
<td>.19</td>
<td>0.004</td>
<td>.993</td>
</tr>
<tr>
<td>50</td>
<td>.14</td>
<td>0.003</td>
<td>1.000</td>
</tr>
</tbody>
</table>
3.3 Estimated Initial Simple Structure Factor Matrix, \( \hat{B}(1) \)

Corresponding to \( \hat{k}(1) \), is the 50x11 matrix \( \hat{B}(1) \)--the sample estimate of the initial simple structure factor matrix, \( \hat{B}(i) \). It indicates the pattern of linear relationships between the standardized (analytical) semantic variables of \( \hat{z} \) and the initial semantic factor variables of the \( 11 \times 1 \) random vector \( \tilde{g}(i) \).

3.4 Initial Semantic Groupings

A permuted form of \( \hat{B}(1) \), denoted \( \hat{B}^*(1) \), is given in Figure 3. The permutation was made by interchanging the rows of \( \hat{B}(1) \) so as to form as many ISG's (initial semantic groups) as possible in accordance with the properties of ISG's given in Chapter III (section 2.3).

In presenting \( \hat{B}^*(1) \) in Figure 3, it is arbitrarily assumed that only those correlation coefficients greater than .20 are potentially "significant" from the standpoint of either interpretation or estimation of the semantic factors of \( \tilde{g}(i) \). Consequently, only coefficients greater than .20 are shown.

Also, it will be noted that the row vectors of \( \hat{B}^*(1) \) are semantically identified by the polar adjective pairs used to identify the corresponding semantic variables.

As can be seen from Figure 3, six ISG's, \( \hat{z}(r) \), \( (r = 1, \ldots, k' = 6) \), containing 29 of the 50 semantic
<table>
<thead>
<tr>
<th>Corresponding Semantic Scales</th>
<th>Pure Identifiable Semantic Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{P}_1$</td>
<td>1 2 3 4 5 6 7 8 9 10 11</td>
</tr>
<tr>
<td>16 required--optional</td>
<td>79</td>
</tr>
<tr>
<td>43 indispensable--dispensable</td>
<td>76</td>
</tr>
<tr>
<td>40 necessary--unnecessary</td>
<td>73 25</td>
</tr>
<tr>
<td>5 unavoidable--avoidable</td>
<td>69 31 20 -26</td>
</tr>
<tr>
<td>41 important--unimportant</td>
<td>63 38 22 -24</td>
</tr>
<tr>
<td>$\tilde{P}_2$</td>
<td></td>
</tr>
<tr>
<td>35 unmonotonous--monotonous</td>
<td>81</td>
</tr>
<tr>
<td>29 interesting--boring</td>
<td>73 23</td>
</tr>
<tr>
<td>45 untedious--tedious</td>
<td>70 21</td>
</tr>
<tr>
<td>49 stimulating--unstimulating</td>
<td>69 -22</td>
</tr>
<tr>
<td>50 unrepetitive--repetitive</td>
<td>68</td>
</tr>
<tr>
<td>46 challenging--unchallenging</td>
<td>65 37 25</td>
</tr>
<tr>
<td>$\tilde{P}_3$</td>
<td></td>
</tr>
<tr>
<td>25 complete--incomplete</td>
<td>89</td>
</tr>
<tr>
<td>24 sufficient--insufficient</td>
<td>86</td>
</tr>
<tr>
<td>8 conclusive--inconclusive</td>
<td>62 -31</td>
</tr>
<tr>
<td>33 infallible--fallible</td>
<td>58 -38</td>
</tr>
<tr>
<td>$\tilde{P}_4$</td>
<td></td>
</tr>
<tr>
<td>20 orderly--disorderly</td>
<td>70 29</td>
</tr>
<tr>
<td>47 productive--unproductive</td>
<td>27 64 -25</td>
</tr>
<tr>
<td>22 useful--useless</td>
<td>36 64</td>
</tr>
<tr>
<td>27 systematic--unsystematic</td>
<td>63 21</td>
</tr>
<tr>
<td>38 effective--ineffective</td>
<td>29 -23 59 23</td>
</tr>
<tr>
<td>14 helpful--obstructive</td>
<td>22 58</td>
</tr>
<tr>
<td>34 practical--impractical</td>
<td>36 57 -24</td>
</tr>
<tr>
<td>19 verifiable--unverifiable</td>
<td>53 -21 23 -24</td>
</tr>
<tr>
<td>30 consistent--inconsistent</td>
<td>46 -20 33 -33</td>
</tr>
</tbody>
</table>

**FIGURE 3**

$\tilde{P}_0^{(1)}$
<table>
<thead>
<tr>
<th>Corresponding Semantic Scales</th>
<th>Pure Identifiable Semantic Variation</th>
<th>Pure Unidentifiable Semantic Variation</th>
<th>Complex Semantic Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>1 demanding--undemanding</td>
<td>21</td>
<td>2 safe--risky</td>
</tr>
<tr>
<td></td>
<td>11 inconvenient--convenient</td>
<td>63</td>
<td>3 restrictive--unrestrictive</td>
</tr>
<tr>
<td></td>
<td>12 complex--simple</td>
<td>21</td>
<td>4 imaginative--unimaginative</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td></td>
<td>6 penetrating--superficial</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td></td>
<td>12 logical--illogical</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td></td>
<td>13 sophisticated--unsophistic</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td></td>
<td>23 varied--routine</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>38</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>39</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>55</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>56</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>57</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>58</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>59</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>61</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>66</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>68</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>71</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>72</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>75</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>76</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>81</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>82</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>83</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>84</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>85</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 1—Continued.
### Complex Semantic Variables

<table>
<thead>
<tr>
<th>Complex Semantic Scales</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>active--passive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-34</td>
<td></td>
<td></td>
<td></td>
<td>47</td>
<td>25</td>
</tr>
<tr>
<td>multi-purposed--single-purposed</td>
<td>24</td>
<td>35</td>
<td>21</td>
<td>-34</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>significant--insignificant</td>
<td>41</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exploratory--unexploratory</td>
<td>21</td>
<td>28</td>
<td>21</td>
<td>35</td>
<td>-27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>deliberate--random</td>
<td></td>
<td>32</td>
<td>27</td>
<td>38</td>
<td>-32</td>
<td>-32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note 1:** Only those correlation coefficients > .20 are shown.

**Note 2:** The coefficients are rounded to nearest 1/100th.
variables, have resulted from such a permutation of \( \hat{B}(1) \).

That is, 29 of the 50 variables of \( \tilde{z} \) are classified as pure identifiable semantic variables. Of the remaining variables, 9 are classified pure unidentifiable semantic variables and 12 as complex semantic variables (Chapter III, section 2.6). These remaining variables are shown in the last 21 rows of \( \hat{B}(1) \).

3.5 Final Estimate, \( \hat{k} \), of the Number, \( \hat{k} \), of Semantic Factor Variables

According to the two-stage estimation procedure of Chapter III (section 2.4), the final estimate, \( \hat{k} \), is defined as the number, \( k' \), of ISG's. Accordingly, for the observations of \( E^0(Q, P^a, C^a) \), \( \hat{k} = 6 \).

Using Table 1, the estimated final ESVR, given by

\[
\text{trace } D_{11}^2 = \sum_{i=1}^{6} \hat{\alpha}_i^2,
\]

is 22.88. The proportion of the estimated ESVR to the total experimental variance of the observations of \( E^0(Q, P^a, C^a) \), given by trace \( D_{11}^2 / 50 \), is .458.

3.6 Estimated Simple Structure Factor Matrix

Corresponding to the final estimate, \( \hat{k} = 6 \), of \( k \) is the 50x6 matrix \( \hat{B} \)--the sample estimate of the simple structure factor matrix, \( \bar{B} \). It indicates the pattern of linear relationships between the standardized (analytical) semantic
variables of $\tilde{z}$ and the six semantic factor variables of the $6 \times 1$ random vector $\tilde{g}$.

### 3.7 Final Semantic Groupings

A permuted form of $\hat{B}$, denoted $\hat{B}^*$, is shown in Figure 4. The permutation was made by interchanging the rows of $\hat{B}$ so as to form six FSG's (final semantic groups) according to the properties of FSG's given in Chapter III (section 2.6).

In presenting $\hat{B}$ in Figure 4, it is arbitrarily assumed (as in Figure 3) that only those correlation coefficients greater than .20 are potentially "significant" from the standpoint of either interpretation or estimation of the semantic factors of $\tilde{g}$. Consequently, only coefficients greater than .20 are shown.

From Figure 4, it can be seen that the six FSG's $\tilde{z}(r), (r = 1, \ldots, k^* = k = 6)$, contain all but ten of the semantic variables of $\tilde{z}$. That is, 40 of the 50 variables of $\tilde{z}$ are classified as pure identifiable semantic variables. The remaining ten variables (shown in the last rows of $\hat{B}^*$) are complex semantic variables. There are no pure unidentifiable variables (see Chapter III, section 2.6).
### Corresponding Semantic Scales

<table>
<thead>
<tr>
<th>Pure Identifiable Semantic Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{z}_j(1) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 indispensable--dispensable</td>
<td>74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43 required--optional</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 unavoidable--avoidable</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40 necessary--unnecessary</td>
<td>69</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41 important--unimportant</td>
<td>61</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bar{z}_j(2) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35 unmonotonous--monotonous</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29 interesting--boring</td>
<td>72</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45 untedious--tedious</td>
<td>69</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49 stimulating--unstimulating</td>
<td>68</td>
<td>21</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 unrepetitive--repetitive</td>
<td>67</td>
<td>-23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46 challenging--unchallenging</td>
<td>66</td>
<td></td>
<td></td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 varied--routine</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bar{z}_j(3) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 complete--incomplete</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 sufficient--insufficient</td>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 conclusive--inconclusive</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33 infallible--fallible</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 safe--risky</td>
<td>51</td>
<td>-33</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 penetrating--superficial</td>
<td>44</td>
<td>-21</td>
<td>-21</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**FIGURE 4**
### Corresponding Semantic Scales

Pure Identifiable Semantic Variables

<table>
<thead>
<tr>
<th>Pure Identifiable Semantic Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_j(4)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30 consistent--inconsistent</td>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 orderly--disorderly</td>
<td>66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47 productive--unproductive</td>
<td>60</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39 direct--indirect</td>
<td>60</td>
<td>37</td>
<td>-24</td>
<td></td>
<td></td>
<td>57</td>
</tr>
<tr>
<td>38 effective--ineffective</td>
<td></td>
<td>47</td>
<td></td>
<td>56</td>
<td>60</td>
<td>37</td>
</tr>
<tr>
<td>36 significant--insignificant</td>
<td></td>
<td>43</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 helpful--obstructive</td>
<td></td>
<td>26</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 useful--useless</td>
<td></td>
<td>47</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 logical--illogical</td>
<td></td>
<td>46</td>
<td>54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34 practical--impractical</td>
<td></td>
<td>44</td>
<td>52</td>
<td>-23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27 systematic--unsystematic</td>
<td></td>
<td></td>
<td></td>
<td>51</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>42 concrete--abstract</td>
<td></td>
<td></td>
<td></td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48 deliberate--random</td>
<td></td>
<td></td>
<td></td>
<td>49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 verifiable--unverifiable</td>
<td></td>
<td></td>
<td></td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28 rational--intuitive</td>
<td></td>
<td></td>
<td></td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Z_j(5)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 demanding--undemanding</td>
<td>62</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 inconvenient--convenient</td>
<td>-28</td>
<td>59</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37 complex--simple</td>
<td>55</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FIGURE 4--Continued.
<table>
<thead>
<tr>
<th>Pure Identifiable Semantic Variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{\tau}_j(6) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26 broad--narrow</td>
<td>-22</td>
<td></td>
<td></td>
<td>65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 universal--particular</td>
<td></td>
<td></td>
<td></td>
<td>62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 general--specific</td>
<td>28</td>
<td>-26</td>
<td>23</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32 multi-purposed--single-purposed</td>
<td>-27</td>
<td>-30</td>
<td>-22</td>
<td>42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complex Semantic Variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \bar{\tau}_j )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 adaptable--unadaptable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>3 unrestrictive--restrictive</td>
<td>26</td>
<td>27</td>
<td>-27</td>
<td>-25</td>
<td>-24</td>
<td></td>
</tr>
<tr>
<td>4 imaginative--unimaginative</td>
<td></td>
<td>40</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 supporting--unsupporting</td>
<td></td>
<td></td>
<td>-37</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 sound--unsound</td>
<td>37</td>
<td>-35</td>
<td>43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 sophisticated--unsophisticated</td>
<td></td>
<td>37</td>
<td>-21</td>
<td>34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 objective--subjective</td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 formal--informal</td>
<td></td>
<td>34</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 active--passive</td>
<td></td>
<td>34</td>
<td>32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44 exploratory--unexploratory</td>
<td></td>
<td>31</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note 1: Only those correlation coefficients > .20 are given.

Note 2: The coefficients are rounded to nearest 1/100th.

FIGURE 4--Continued.
CHAPER VI

EVALUATION OF EXPERIMENTAL RESULTS

0. Introduction

In this chapter, the results of $E^O(Q, P^a, C^a)$ are discussed. Specifically, an evaluation of the semantic factors derived from the observations of $E^O(Q, P^a, C^a)$ is made. In part 1, an interpretation and tentative identification of the semantic factors is given in terms of the semantic content of the FSG's. In part 2, the semantic factors obtained are evaluated on an internal basis, using the estimated variances of the corresponding principal component variables from which the factors are derived; and on an external basis, using a Kruskal-Wallis one-way analysis of variance to test the hypothesis of no difference among average semantic scores over the six organizational categories.

1. Interpretation of Experimental Results

1.1 Comparison of the Semantic Content of the ISG's and FSG's

An examination of the two sets of semantic groups (the ISG's and the FSG's) reveals that their semantic content is similar. The primary difference is that the
FSG's contain 11 more semantic variables than the ISG's. Also, five semantic variables have been reclassified from pure unidentifiable variables at the ISG stage (stage 1) to complex variables at the FSG stage (stage 2). A summary of these semantic content changes appears in Table 2.

From Table 2, the following relationships between the respective ISG's and FSG's are deduced:

\[ \bar{z}(1) = z(1) \]
\[ \bar{z}(2) \subseteq \bar{z}(2) \]
\[ \bar{z}(3) \subseteq \bar{z}(3) \]
\[ \bar{z}(4) \subseteq \bar{z}(4) \]
\[ \bar{z}(5) = z(5) \]
\[ \bar{z}(6) \subseteq \bar{z}(6) \]

As will be noticed from Table 2, a substantial change occurred between \( \bar{z}(4) \) and \( z(4) \), the latter containing six additional semantic variables.

1.2 Evaluation of Semantic Groups

The formulation of the semantic groups shown in Figures 3 and 4 is on the basis of the properties stated in the definition of an "ideal" semantic group in Chapter III (section 2.3 and 2.5). For convenience, the properties are restated at this point. The semantic variables of the rth semantic group (initial or final) are:
<table>
<thead>
<tr>
<th>( \bar{z}_j )</th>
<th>Semantic Identification</th>
<th>ISG's</th>
<th>FSG's</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>adaptable--unadaptable</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>2</td>
<td>safe--risky</td>
<td>X</td>
<td>( \bar{z}(3) )</td>
</tr>
<tr>
<td>6</td>
<td>penetrating--superficial</td>
<td>X</td>
<td>( \bar{z}(3) )</td>
</tr>
<tr>
<td>7</td>
<td>supporting--unsupporting</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>9</td>
<td>general--specific</td>
<td>X</td>
<td>( \bar{z}(6) )</td>
</tr>
<tr>
<td>10</td>
<td>sound--unsound</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>12</td>
<td>logical--illogical</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
<tr>
<td>17</td>
<td>objective--subjective</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>18</td>
<td>formal--informal</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>23</td>
<td>varied--routine</td>
<td>X</td>
<td>( \bar{z}(2) )</td>
</tr>
<tr>
<td>28</td>
<td>rational--intuitive</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
<tr>
<td>32</td>
<td>multi-purposed--single-purposed</td>
<td>X</td>
<td>( \bar{z}(6) )</td>
</tr>
<tr>
<td>36</td>
<td>significant--insignificant</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
<tr>
<td>39</td>
<td>direct--indirect</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
<tr>
<td>42</td>
<td>concrete--abstract</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
<tr>
<td>48</td>
<td>deliberate--random</td>
<td>X</td>
<td>( \bar{z}(4) )</td>
</tr>
</tbody>
</table>
(1) semantically related;
(2) small in number compared to \( p \);
(3) in no other semantic group;
(4) "significantly" related to only the \( r \)th semantic factor; and
(5) linearly related to the \( r \)th semantic factor to an extent greater than or equal to any variable not in the \( r \)th semantic group.

The extent to which these properties are satisfied by the semantic groups shown in Figures 3 and 4 is now discussed. The first property is reasonably satisfied, except, perhaps, for the fourth FSG, \( \tilde{z}(4) \). The semantic content of \( \tilde{z}(4) \) is more "diluted" than its counterpart \( \tilde{z}(4)' \) in that the added variables, those in \( \tilde{z}(4) \) but not in \( \tilde{z}(4)' \) (Table 2), contribute to an additional semantic aspect (see section 1.3). Such semantic dilution makes interpretation of the corresponding semantic factor, \( \tilde{g}_h \), difficult.

Property (2) also is reasonably satisfied, except, again, for \( \tilde{z}(4) \). All semantic groups, except \( \tilde{z}(4) \), have less than eight variables. \( \tilde{z}(4) \) has 15.

Property (3) is completely satisfied for all groups, both initial and final.

On the basis of the average magnitudes of the linear relations between the variables of a given group and the semantic factors, property (4) is reasonably satisfied.
For example, from Figure 4, it would appear reasonable to conclude that each of the variables of \( z(1) \) is "significantly" related only to \( \tilde{F}_1 \), since in only two instances (\( \tilde{z}_{j0}(1) \) and \( \tilde{z}_{j1}(1) \)), do the linear relationships with another semantic factor (in this case, \( \tilde{g}_h \)) exceed .20 in magnitude. The variables of the remaining FSG's and those of the ISG's reflect a similar pattern of linear relation, although, in some cases, to a lesser degree.

Property (5) is satisfied for all groups except \( _1 \tilde{z}(4) \) and \( \tilde{z}(4) \). For \( _1 \tilde{z}(4) \), the inclusion of \( \tilde{z}_3 \) represents a deviation since

\[
\text{cov}(\tilde{z}_3, \tilde{g}_4(1)) = .46 < \text{cov}(\tilde{z}_{12}, \tilde{g}_4(1)) = .48,
\]

where \( \tilde{z}_{12} \) is classified as a complex variable. For \( \tilde{z}(4) \), the inclusion of \( \tilde{z}_{28} \) and \( \tilde{z}_{19} \) represents a deviation since

\[
\text{cov}(\tilde{z}_{28}, \tilde{g}_4) = .45 < \text{cov}(\tilde{z}_{19}, \tilde{g}_4) = .48 < \text{cov}(\tilde{z}_{h1}, \tilde{g}_h) = .49
\]

where \( \tilde{z}_{h1} \leq \tilde{z}(1) \).

1.3 Semantic Interpretation and Identification of Semantic Factors

Of the two sets of semantic groups, the semantic content of the FSG's is of greatest interest from the standpoint of the interpretation and identification of the six semantic factors of \( \tilde{g} \). One such interpretation and identification is suggested in Table 3.

Of the six semantic factors, \( \tilde{g}_h \), identified as a "logic-utility" factor, is the most difficult to interpret
<table>
<thead>
<tr>
<th>Semantic Factor Variables</th>
<th>Identification of Semantic Factor Variables</th>
<th>Interpretation of Semantic Content of Corresponding FSG</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{g}_1$ Necessity Factor</td>
<td>Indicates the necessity and indispensability of a technique to an auditing situation.</td>
<td></td>
</tr>
<tr>
<td>$\tilde{g}_2$ Interest Factor</td>
<td>Indicates the interest, stimulation and challenge of a technique as an audit task.</td>
<td></td>
</tr>
<tr>
<td>$\tilde{g}_3$ Validity-Strength Factor</td>
<td>Indicates the validity, sufficiency and strength of a technique as a methodological tool.</td>
<td></td>
</tr>
<tr>
<td>$\tilde{g}_4$ Logic-Utility Factor</td>
<td>Indicates the orderliness, effectiveness and practicability of a technique.</td>
<td></td>
</tr>
<tr>
<td>$\tilde{g}_5$ Difficulty Factor</td>
<td>Indicates the rigor and complexity of a technique.</td>
<td></td>
</tr>
<tr>
<td>$\tilde{g}_6$ Scope Factor</td>
<td>Indicates the generality and versatility of a technique.</td>
<td></td>
</tr>
</tbody>
</table>
since it appears to represent at least two semantic aspects: a "logic" aspect relating to order and rationality and a "utility" aspect relating to effectiveness and productiveness. In Table 4 the variables of $\hat{z}(4)$, except for two ($\tilde{z}_{19}^{(4)}$ and $\tilde{z}_{36}^{(4)}$), are classified according to these two semantic aspects. $\tilde{z}_{19}^{(4)}$ and $\tilde{z}_{36}^{(4)}$, not clearly related to either aspect, are shown separately.

From Table 4 and an examination of Figure 4, it would appear that the "logic" aspect is dominant, both from the standpoint of the number of variables and the clarity of their linear relations with $\tilde{z}_4$. That is, the eight variables associated with the "logic" aspect appear to be primarily related to $\tilde{z}_4$ (with the exception of $\tilde{z}_{12}^{(4)}$), whereas the five variables corresponding to the "utility" aspect appear to be, at least to some extent, also related to $\tilde{z}_1$. This, of course, reflects upon the extent to which $\tilde{z}(4)$ satisfies property (4) of the definition of a semantic group (see section 1.2).

2. Evaluation of Semantic Factors

2.1 Internal Evaluation

An internal index of the relative importance of semantic factors is possible on the basis of the estimated variances of the principal component variables, $\hat{u}_r^{(1)}$ ($r = 1, \ldots, k$), from which the factors are derived (see Chapter III, section 2.2). For the observations of
### TABLE 4
CLASSIFICATION OF SEMANTIC VARIABLES OF $z(4)$
ACCORDING TO SEMANTIC CONTENT

<table>
<thead>
<tr>
<th>Logic Aspect</th>
<th>Utility Aspect</th>
<th>Unclassified</th>
</tr>
</thead>
<tbody>
<tr>
<td>$z_{30}$ consistent--inconsistent</td>
<td>$z_{47}$ productive--unproductive</td>
<td>$z_{36}$ significant--insignificant</td>
</tr>
<tr>
<td>$z_{20}$ orderly--disorderly</td>
<td>$z_{38}$ effective--ineffective</td>
<td>$z_{19}$ verifiable--unverifiable</td>
</tr>
<tr>
<td>$z_{12}$ logical--illogical</td>
<td>$z_{14}$ helpful--obstructive</td>
<td></td>
</tr>
<tr>
<td>$z_{27}$ systematic--unsystematic</td>
<td>$z_{22}$ useful--useless</td>
<td></td>
</tr>
<tr>
<td>$z_{48}$ deliberate--random</td>
<td>$z_{34}$ practical--impractical</td>
<td></td>
</tr>
<tr>
<td>$z_{28}$ rational--intuitive</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{39}$ direct--indirect</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$z_{42}$ concrete--abstract</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

114
E^0(Q, P^a, C^a), Table 5, using the information given in Table 1 (section 3.1) and recalling that
\[ \text{var}(\bar{u}_r^{(1)}) = \frac{\hat{\sigma}^2}{d_r} \quad r = 1, \ldots, 6, \]
presents the estimated variances of each of the first six principal component variables, the percentage of these estimated variances to the total experimental variance, \( p \), and the cumulative percentages of these estimates to the total.

As noted (Chapter V, section 3.5), the final cumulative percentage, 45.8%, is the percentage of estimated ESVR to the total experimental specification variance, \( p \). ESVR, it is recalled (Chapter II, section 5.3), is the total variance of the variables of \( z^1 \) and is interpreted as that amount of the total variance of the variables of \( z \) "attributable to" or "explained by" the variables of \( \overline{z} \). For the particular observations of \( E^0(Q, P^a, C^a) \), 45.8% of the total experimental variance is "attributable to" six (semantic factor) variables, representing 12%, in number, of the original 50 semantic variables.

In interpreting and evaluating semantic factor results, it should be emphasized that the results of a particular experiment, such as \( E^0(Q, P^a, C^a) \), are dependent upon the particular experimental sample of semantic specifiers involved, such as Q. As discussed in Chapter II (section 6), another sample of 50 specifiers, say \( Q^x \), would
<table>
<thead>
<tr>
<th>Principal Component Variables $\bar{u}_r^{(1)}$</th>
<th>Estimated Variances of Principal Component Variables $\hat{\text{var}}(\bar{u}_r^{(1)})$</th>
<th>Percentage of Estimated Variances to Total Experimental Variance $\frac{\hat{\text{var}}(\bar{u}_r^{(1)})}{p} \times 100$</th>
<th>Cumulative Percentages of Estimated Variances to Total Experimental Variance $\frac{\hat{\text{var}}(\bar{u}_r^{(1)})}{p} \times 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.45</td>
<td>20.9%</td>
<td>20.9%</td>
</tr>
<tr>
<td>2</td>
<td>4.31</td>
<td>8.6</td>
<td>29.5</td>
</tr>
<tr>
<td>3</td>
<td>2.58</td>
<td>5.2</td>
<td>34.7</td>
</tr>
<tr>
<td>4</td>
<td>1.99</td>
<td>4.0</td>
<td>38.7</td>
</tr>
<tr>
<td>5</td>
<td>1.94</td>
<td>3.8</td>
<td>42.5</td>
</tr>
<tr>
<td>6</td>
<td>1.62</td>
<td>3.3</td>
<td>45.8</td>
</tr>
</tbody>
</table>
likely involve a different proportion of ESVR
\((\text{trace}^X \bar{D} \bar{D}^2 / p^X)\), a different number \((k^X)\) of semantic factors
and different semantic factors \((e_\pi^X)\) (see Chapter II, section 6.2). Or, equivalently, another selection, \(Q^X\),
would likely result in a different "amount" of SC (specification capacity—see Chapter II, section 1.4). That is, it
would be expected that for another selection, \(Q^X\),
\[ SC(Q) \neq SC(Q^X). \]

Also, the relative importance of semantic factors
of two experiments, say \(E^O(Q, P^a, C^a)\) and \(E^O(Q^X, P^a, C^a)\),
as evidenced by the estimated variances of the corresponding principal component variables, would likely differ.

Furthermore, it would be expected that the
magnitude of the estimated variances (of the first \(k\)
principal component variables) would reflect upon the
"stability" of the derived semantic factors or, equivalently, upon the extent to which they approximate any of
the actual semantic factors of \(\bar{g}^+\). That is, it seems
logical to suppose that the closer the approximation, the
more stable the corresponding factor in representing
systematic semantic specification behavior and the more
interpretive its meaning both in terms of semantic content
and in the extent to which it relates to external criteria.

For \(E^O(Q, P^a, C^a)\), and on the basis of this
reasoning, it would appear that \(\bar{g}_1\), the first semantic
factor, is relatively stable. The variance of its corresponding principal component, $\bar{u}_1^{(1)}$, accounts for $1/5$ of the total experimental variance. The variance of the second principal component variable, $\bar{u}_2^{(1)}$, is less than half of the variance of $\bar{u}_1^{(1)}$ and consequently, $\bar{g}_2$ appears to represent a much less stable semantic factor. The remaining factors appear even less stable than $\bar{g}_2$ and consequently are likely not to approximate to the extent of $\bar{g}_2$ any of the semantic factors of $\bar{g}^+$. It is difficult, however, to meaningfully assess, in the absence of additional criteria, the importance, in any absolute sense, of any of these factors. As will be seen in section 4.5, the importance of $\bar{g}_4$ appears greater than that of $\bar{g}_2$ or $\bar{g}_3$, insofar as is indicated by the extent to which it relates to an external criterion.

In assessing the magnitude of the estimated ESVR of $E^0(Q, P^a, C^a)$, it is relevant to recall that one of the two criteria used in selecting the semantic scales of $Q$ was completeness of the sample in reflecting as many of the semantic dimensions associated with $Q^+$ as possible (Chapter V, section 1.3). In using such a criterion for determining $Q$, it would be expected that the larger the number, $k^+$, of actual semantic dimensions associated with $Q^+$ (Chapter II, section 1.5), the smaller the ESVR and corresponding estimate and conversely, the smaller the
number, $k^+$, the larger the ESVR and corresponding estimate. For example, if $k^+ = p - 1$, then it would be expected that a set of semantic scales selected so as to represent as many of these $k^+$ semantic dimensions as possible would be spread so "thin" as to quite likely not yield any definite or clearly defined semantic groups and corresponding semantic factors. And consequently, the ESVR and corresponding estimate would constitute a relatively small proportion of the total specification variance, $p$. At the other extreme, if $k^+ = 1$ and is significantly less than the number, $p$, of semantic scales, then it would be expected that a definite and dominant semantic group and corresponding factor would result and that the ESVR and corresponding estimate would constitute a relatively large proportion of $p$.

As implied in this discussion and previously in Chapter II (section 6.3), what is required for complete and sound internal evaluation (i.e., excluding external assessment criteria) is the actual proportion of ESVR to total variance ($\text{trace} D_{q,l}^2 / p^+$), the actual number of semantic factor variables ($k^+$) and the actual semantic factors themselves ($g_{r}^{+}, r = 1, ..., p^+$), estimates of which are available only by conducting an experiment of the form $E^+(Q^+, P, C)$, involving the universe, $Q^+$, of semantic specifiers relevant to $P$ and $C$. A theoretically less
desirable but more practical alternative is to conduct a sequence of experiments of the form $E^O(Q, P, C)$, each involving a $Q$ which represents an improvement over those $Q$'s used in previous experiments in the sense of more efficiently and completely representing the semantic factors indicated in those experiments.

2.2 External Evaluation

A potential value of the derivation of semantic factor variables is the extent to which they differentiate auditors on the basis of specified criteria. For the experimental sample, $P_a$, of auditors of $E^O(Q, P_a, C_a)$, one criterion of differentiation is the organizational category from which the auditor comes. Do the semantic factors appear to differentiate on the basis of this organizational criterion? A preliminary investigation of this question was made by computing the average estimated semantic factor scores, denoted $\hat{\mathbf{g}}_r$ ($1, r = 1, \ldots, 6$), for the auditors of each of the six organizational categories (from which the selection of $P_a$ was made) for each of the six semantic factors. Recalling that (Chapter II, section 6.2)

\begin{equation}
(2.2.1) \quad \hat{\mathbf{g}} = \mathbf{T}^{-1}L^L(1)^d \hat{\mathbf{z}},
\end{equation}

the computational equation for the estimated semantic factor scores of $E^O(Q, P_a, C_a)$ is:
(2.2.2) \[ \hat{\gamma}_N = \hat{\gamma}_k \hat{\gamma}_d \hat{\gamma}_l \hat{\gamma}_p \hat{\gamma}_N \]

where \( \hat{\gamma}_k, \hat{\gamma}_d, \hat{\gamma}_l, \hat{\gamma}_p \) and \( \hat{\gamma}_N \) are the sample analogues to the model as defined in the population (see Chapter II, section 7; Chapter IV, section 4).

From the estimated semantic scores computed according to (2.2.2), the average estimated scores by organizational category and semantic factor were computed. These data appear in Table 6. To more readily see the pattern of differentiation reflected by these averages, a graphic profile of them is given in Figure 5. In Figure 5, the graphic profiles are shown in three graphs, two profiles per graph.

The profiles of Figure 5 suggest two questions concerning the population average semantic scores, \( \bar{\bar{\gamma}}_r \)

(i, r = 1, ..., 6):

(1) To what extent do the \( \bar{\bar{\gamma}}_r \) differentiate over the six organizational categories?

(2) What is the pattern of differentiation, to the extent it exists, from the standpoint of the ranking of the organizational categories as indicated by the positions of the \( \bar{\bar{\gamma}}_r \) on the scale representing the semantic factors?
### TABLE 6

**AVERAGE ESTIMATED SEMANTIC FACTOR SCORES FOR EACH OF THE SIX ORGANIZATIONAL CATEGORIES ON EACH OF THE SIX SEMANTIC FACTORS**

<table>
<thead>
<tr>
<th>Organizational Category</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partner</td>
<td>+.266</td>
<td>+.200</td>
<td>+.058</td>
<td>+.043</td>
<td>+.033</td>
<td>+.222</td>
</tr>
<tr>
<td>Manager</td>
<td>+.336</td>
<td>+.068</td>
<td>-.015</td>
<td>-.040</td>
<td>+.165</td>
<td>-.125</td>
</tr>
<tr>
<td>Supervisor</td>
<td>+.092</td>
<td>-.116</td>
<td>-.123</td>
<td>-.029</td>
<td>-.090</td>
<td>+.036</td>
</tr>
<tr>
<td>Senior</td>
<td>-.213</td>
<td>-.044</td>
<td>-.072</td>
<td>-.205</td>
<td>-.112</td>
<td>-.108</td>
</tr>
<tr>
<td>In-Charge</td>
<td>-.368</td>
<td>-.010</td>
<td>-.053</td>
<td>+.106</td>
<td>-.010</td>
<td>+.035</td>
</tr>
<tr>
<td>Staff</td>
<td>-.170</td>
<td>-.068</td>
<td>-.086</td>
<td>-.230</td>
<td>-.005</td>
<td>-.060</td>
</tr>
</tbody>
</table>

Where: $\hat{\gamma}_r^{i}$ is the average estimated semantic factor score for the $i$th organizational category ($i = 1, \ldots, 6$) on the $r$th semantic factor ($r = 1, \ldots, 6$).
FIGURE 5

A GRAPHIC PROFILE OF THE DIFFERENTIAL PATTERN OF SIX SEMANTIC FACTORS OVER EACH OF SIX ORGANIZATIONAL CATEGORIES.
FIGURE 5—Continued.

Partner Manager Supervisor Senior In-Charge Staff
Difficulty Factor ($\bar{R}_5$)
Scope Factor ($\bar{R}_6$)

Partner Manager Supervisor Senior In-Charge Staff

FIGURE 5—Continued.
2.2a Extent of Differentiation Over Organizational Categories

The first question was approached by asking whether the (population) average semantic factor scores of a given semantic factor differed significantly over the six organizational categories. Specifically, six null hypotheses, $H_0^r (r = 1, ..., 6)$, were formulated, one for each semantic factor:

$$H_0^r : \bar{\bar{z}}_1 = ... = \bar{\bar{z}}_6, \quad r = 1, ..., 6.$$

The corresponding alternative hypotheses, $H_1^r (r = 1, ..., 6)$, are stated:

$$H_1^r : \bar{\bar{z}}_1 \neq ... \neq \bar{\bar{z}}_6, \quad r = 1, ..., 6.$$

The method used to test these hypotheses was the Kruskal-Wallis one-way analysis of variance by ranks for $h$ independent samples (Johnson and Leone, p. 271 and Siegel, pp. 184-192). The statistic of the method basically involves the squared deviations of rank orders (of in this case, estimated semantic factor scores) from a mean rank value. The statistic is distributed approximately as a chi square with degrees of freedom, $h - 1$ ($h$, in this case, being 6, the number of organizational categories), for a sample in which the numbers in each of the $h$ groups is sufficiently large (Siegel, p. 185).

Use of this non-parametric test rather than a corresponding parametric test avoids assumptions of
normality and homogeneity of variance (Siegel, p. 189).
The asymptotic power-efficiency of the test is 95.5% (Siegel, p. 193).

The results of the test—the test statistics and levels at which the $H_0$ can be rejected—appear in Table 7.

From Table 7, only two of the factors, $\bar{z}_1$ and $\bar{z}_4$ differentiate to any significant extent over the categories. $\bar{z}_1$ differentiates most with a rejection level of .001, followed by $\bar{z}_4$, with a rejection level of .05. The remaining factors do not differentiate to any extent as evidenced by the unreasonably high rejection levels of these factors.

2.2b Pattern of Differentiation Over Organizational Categories

The pattern of differentiation from the standpoint of the ranking of the organizational categories on the basis of average estimated semantic scores varies according to the particular semantic factor and organizational category considered. On an a priori and intuitive basis, one would like to see a consistent ranking from the positive side of the scale to the negative side (or vice versa) of partner, manager, supervisor, senior, in-charge and staff. There is no inherent reason for the ranking to be from positive to negative or the other way around, but there is reason to expect consistency.
TABLE 7

KRUSKAL-WALLIS TEST STATISTIC AND LEVEL OF REJECTION
OF NULL HYPOTHESIS OF NO DIFFERENCE AMONG
AVERAGE SEMANTIC SCORES OVER SIX
ORGANIZATIONAL CATEGORIES
FOR EACH OF SIX
SEMANTIC FACTORS

<table>
<thead>
<tr>
<th>Semantic Factor</th>
<th>Kruskal-Wallis Test Statistic</th>
<th>Probability Level at Which Hypothesis, $H_0$, Can be Rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Necessity Factor ($g_1$)</td>
<td>26.46</td>
<td>.001</td>
</tr>
<tr>
<td>Interest Factor ($g_2$)</td>
<td>.23</td>
<td>.90</td>
</tr>
<tr>
<td>Validity-Strength Factor ($g_3$)</td>
<td>3.24</td>
<td>.70</td>
</tr>
<tr>
<td>Logic-Utility Factor ($g_4$)</td>
<td>12.42</td>
<td>.05</td>
</tr>
<tr>
<td>Difficulty Factor ($g_5$)</td>
<td>2.93</td>
<td>.80</td>
</tr>
<tr>
<td>Scope Factor ($g_6$)</td>
<td>4.34</td>
<td>.70</td>
</tr>
</tbody>
</table>

Using the information of Table 6, Table 8 indicates the rankings of the categories according to the average estimated semantic scores for each of the six semantic factors. In Table 8, the rankings have been "positioned" so as to distinguish between positive and negative averages. Also, those categories which deviate from a "desirable"
## TABLE 8
INDICATED RANKINGS OF ORGANIZATIONAL CATEGORIES
FOR EACH OF THE SIX SEMANTIC FACTORS

<table>
<thead>
<tr>
<th>$\bar{g}_1$</th>
<th>$\bar{g}_4$</th>
<th>$\bar{g}_2$</th>
<th>$\bar{g}_3$</th>
<th>$\bar{g}_5$</th>
<th>$\bar{g}_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Manager)</td>
<td>Partner</td>
<td>Partner</td>
<td>[Manager]</td>
<td>[Supervisor]</td>
<td></td>
</tr>
<tr>
<td>Partner</td>
<td>[In-Charge]</td>
<td>Manager</td>
<td>Partner</td>
<td>Partner</td>
<td>[In-Charge]</td>
</tr>
<tr>
<td>Supervisor</td>
<td>(Staff)</td>
<td>[In-Charge]</td>
<td>[Staff]</td>
<td>[Staff]</td>
<td></td>
</tr>
<tr>
<td>[Staff]</td>
<td>(Supervisor)</td>
<td>[In-Charge]</td>
<td>[Staff]</td>
<td>[Staff]</td>
<td></td>
</tr>
<tr>
<td>Senior</td>
<td>Manager</td>
<td>(Senior)</td>
<td>[In-Charge]</td>
<td>[In-Charge]</td>
<td>[Senior]</td>
</tr>
<tr>
<td>In-Charge</td>
<td>Senior</td>
<td>[Staff]</td>
<td>(Senior)</td>
<td>Supervisor</td>
<td>Manager</td>
</tr>
<tr>
<td>Staff</td>
<td>Supervisor</td>
<td>[Staff]</td>
<td>Senior</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*a* Positive averages.

*b* Negative averages.

*c* ( )—indicates a "slight" deviation in ranking.

*d* [ ]—indicates a "significant" deviation in ranking.
ranking (see above) are shown in parentheses. To more readily see the pattern of these rankings, a graphic profile of each organizational category over the six semantic factors is given in Figure 6.

From an inspection of Table 8 and Figure 6, two observations can be made. First, $\bar{g}_1$ and $\bar{g}_4$ exhibit the most consistent and desirable ranking of the six factors. For $\bar{g}_1$, the most conspicuous deviation is the staff category, ranking ahead (or above) the senior and in-charge categories. The deviation of the manager category appears slight and, in fact, may not be significantly different (statistically) from the partner category. For $\bar{g}_4$, the in-charge category appears "significantly" out of line, the supervisor category appearing only slightly so. For the remaining factors, the deviations are more numerous and significant with an almost complete breakdown of ranking for $\bar{g}_6$.

Second, for the first four semantic factors, the rankings appear to be most consistent for the upper organizational levels. For $\bar{g}_1$ and $\bar{g}_4$, if one ignores the apparently "slight" deviations, a somewhat consistent ranking is evident for the categories of partner through senior. For $\bar{g}_2$ and $\bar{g}_3$, such consistency appears for the categories of partner through supervisor. No such pattern is evident in the case of $\bar{g}_5$ or $\bar{g}_6$. 
FIGURE 6
A GRAPHIC PROFILE OF SIX ORGANIZATIONAL CATEGORIES OVER SIX SEMANTIC FACTORS
FIGURE 6—Continued.
Over all, it would seem safe to conclude that factors $\tilde{g}_1^*$ and $\tilde{g}_4^*$ possess a reasonable potential for differentiation across organizational categories, from the standpoint of both the extent of differentiation and the pattern of differentiation. This seems especially so for $\tilde{g}_1^*$, for which the Krushal-Wallis analysis of variance indicates significant differences in average estimated semantic scores across the six categories.

It should be noted that this tentative observation pertains only to one possible criterion for which investigation of semantic differentiating potential can be made. Furthermore, the data introduced indicate nothing about differentiating potential within organizational categories. And finally, in regard to the concept domain $C$, nothing can be concluded regarding differential specification behavior relative to specific concepts of $C$. All this suggests further questions for investigation and indicates the need for additional data and analysis before less tentative and more general observations can be made.
CHAPTER VII

ESTIMATIONS OF EXPERIMENTALLY
DERIVED SEMANTIC FACTORS

0. Introduction

The estimations of the six semantic factor variables derived from the observations of $E^0(Q, P^a, C^a)$ are considered in this chapter. The estimations are made according to the linear least-squares estimation method of Chapter III (part 3). In part 1, the semantic predictor variables used in the estimation are given. In part 2, the estimated significant regression weights for each of the six semantic factor variables is presented. In part 3, a set of the "best" estimation equations for the six semantic factors is suggested.

1. Semantic Predictor Variables

According to the definition given in Chapter III (section 3.2), the semantic predictor variables of the semantic factors, $\bar{g}_r$, of $\bar{g}$ are the random variables, $\bar{y}_j(r)$, $(j = 1, \ldots, n_r^2 < n_r < p)$, of the $\bar{y}(r)$.
where: \( \tilde{v}_j(r) = \left[ \text{var}(\tilde{v}_j(r)) \right]^{1/2} \tilde{v}_j(r) \) \( j = 1, \ldots, n_r^* \leq n_r \\
 \tilde{v}_j(r) \neq \tilde{v}_j(r) \) \( r = 1, \ldots, k. \)

For the semantic variables of \( F^0(Q, P^a, C^a) \), the following predictor sets, \( \tilde{v}(r), r = 1, \ldots, 6, \) are defined.

\[
\tilde{v}(1) = \begin{pmatrix}
\tilde{v}_{16}(1) & \text{indispensable--dispensable} \\
\tilde{v}_{13}(1) & \text{required--optional} \\
\tilde{v}_{00}(1) & \text{necessary--unnecessary} \\
\tilde{v}_{5}(1) & \text{unavoidable--avoidable} \\
\tilde{v}_{41}(1) & \text{important--unimportant}
\end{pmatrix}
\]

\[
\tilde{v}(2) = \begin{pmatrix}
\tilde{v}_{35}(2) & \text{unmonotonous--monotonous} \\
\tilde{v}_{45}(2) & \text{untedious--tedious} \\
\tilde{v}_{49}(2) & \text{stimulating--unstimulating} \\
\tilde{v}_{50}(2) & \text{unrepetitive--repetitive} \\
\tilde{v}_{23}(2) & \text{varied--routine} \\
\tilde{v}_{29}(2) & \text{interesting--boring} \\
\tilde{v}_{46}(2) & \text{challenging--unchallenging}
\end{pmatrix}
\]

\[
\tilde{v}(3) = \begin{pmatrix}
\tilde{v}_{25}(3) & \text{complete--incomplete} \\
\tilde{v}_{8}(3) & \text{conclusive--inconclusive} \\
\tilde{v}_{33}(3) & \text{infallible--fallible} \\
\tilde{v}_{2}(3) & \text{safe--risky} \\
\tilde{v}_{24}(3) & \text{sufficient--insufficient} \\
\tilde{v}_{6}(3) & \text{penetrating--superficial}
\end{pmatrix}
\]
It is noted that with the exception of the
4th predictor set, $\bar{y}(4)$, all sets are equivalent in content
to the FSG's, $\bar{z}(1)$, $\bar{z}(2)$, $\bar{z}(3)$, $\bar{z}(5)$ and $\bar{z}(6)$. The set
$\bar{y}(4)$ contains only those variables of $\bar{z}(4)$ relating to the
"logic" aspect of the semantic content of $\bar{z}(4)$ (Table 5).
The primary reason for selecting these variables
for $\bar{y}(4)$ is that they appear to exhibit a more clearly
defined linear relationship with $\bar{z}_4$ (with the possible
exception of $\bar{y}_{12}(4)$) than those variables relating to the
"utility" aspect of the semantic content of \( z(4) \). Most of these latter variables are also, and to some extent, related to \( \tilde{z}_1 \) and thus do not reflect a relatively pure and unambiguous relationship with \( \tilde{z}_d \).

2. Significant Regression Weights

In accordance with the method outlined in Chapter III (part 3), a sequence of \( n_r^{''} \) regression equations is developed for each of the semantic factors derived from \( E^O(\Omega, P^a, C^a) \), such that each equation in each sequence contains a different set of significant (according to the t-test criterion) predictor variables. The \( n_r^{''} \) sets of estimated significant regression weights corresponding to these \( n_r^{''} \) regression equations are given in Table 9 for each of the six semantic factors. Also indicated are the estimated multiple correlation coefficients relevant to each of the \( n_r^{''} \) regression equations for each of the six semantic factors.

3. Suggested "Best" Estimation Equations

In Chapter III (section 3.8), a "best" estimation equation was defined as that equation in the sequence for which no "significant" increase in the \( i^{R^2}_r \cdot \tilde{y}(r) \), \((i = 1, \ldots, n_r^{''})\) is possible. Based on an examination of the \( i^{R^2}_r \cdot \tilde{y}(r) \) in Table 9, the following are suggested as the "best" estimation equations for the six semantic factors:
TABLE 9
A SEQUENCE OF $n_r^{-1}$ SETS OF ESTIMATED SIGNIFICANT REGRESSION WEIGHTS

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>$\hat{w}_j(r)$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}_j(1)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>.03</td>
<td>.21</td>
<td>.14</td>
<td>.13</td>
<td>.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td></td>
<td>.10</td>
<td>.11</td>
<td>.06</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>.14</td>
<td>.16</td>
<td>.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>.05</td>
<td>.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.03</td>
</tr>
<tr>
<td>$iR_1 \cdot \bar{y}(1)$</td>
<td>.071</td>
<td>.654</td>
<td>.715</td>
<td>.738</td>
<td>.741</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{y}_j(2)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
<td>.01</td>
<td>.20</td>
<td>.21</td>
<td>.13</td>
<td>.11</td>
<td>.08</td>
<td>.07</td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
<td>.18</td>
<td>.12</td>
<td>.11</td>
<td>.12</td>
<td>.13</td>
<td>.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td></td>
<td>.09</td>
<td>.11</td>
<td>.11</td>
<td>.11</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>.11</td>
<td>.10</td>
<td>.09</td>
<td>.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.04</td>
<td>.04</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.05</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>$iR_2 \cdot \bar{y}(2)$</td>
<td>.031</td>
<td>.736</td>
<td>.795</td>
<td>.868</td>
<td>.887</td>
<td>.896</td>
<td>.897</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

$^a$See footnote at the end of the table.
TABLE 9—Continued.

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>Estimated Significant Regression Weights$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{y}_j(r)$</td>
<td>1    2    3    4    5    6    7    8</td>
</tr>
<tr>
<td>$\bar{y}_j(3)$</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.02  0.20  0.13  0.14  0.08  0.06</td>
</tr>
<tr>
<td>8</td>
<td>0.14  0.17  0.13  0.13  0.14</td>
</tr>
<tr>
<td>33</td>
<td>0.09  0.10  0.09  0.08</td>
</tr>
<tr>
<td>2</td>
<td>0.06  0.06  0.06</td>
</tr>
<tr>
<td>24</td>
<td>0.08  0.08</td>
</tr>
<tr>
<td>6</td>
<td>0.03</td>
</tr>
<tr>
<td>$i^R3 \cdot \bar{y}(3)$</td>
<td>0.048  0.742  0.826  0.859  0.879  0.888</td>
</tr>
<tr>
<td>$\bar{y}_j(4)$</td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>0.03  0.26  0.17  0.14  0.14  0.12  0.12  0.12</td>
</tr>
<tr>
<td>20</td>
<td>0.11  0.14  0.16  0.09  0.10  0.09  0.09</td>
</tr>
<tr>
<td>42</td>
<td>0.15  0.12  0.13  0.11  0.11  0.11</td>
</tr>
<tr>
<td>48</td>
<td>0.08  0.10  0.08  0.08  0.08</td>
</tr>
<tr>
<td>30</td>
<td>0.08  0.08  0.07  0.07</td>
</tr>
<tr>
<td>28</td>
<td>0.06  0.06  0.06</td>
</tr>
<tr>
<td>27</td>
<td>0.02  0.02</td>
</tr>
<tr>
<td>12</td>
<td>0.00</td>
</tr>
<tr>
<td>$i^R4 \cdot \bar{y}(4)$</td>
<td>0.036  0.518  0.628  0.685  0.718  0.747  0.750  0.750</td>
</tr>
</tbody>
</table>

$^a$See footnote at the end of the table.
TABLE 9—Continued.

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>Estimated Significant Regression Weights&lt;sup&gt;a&lt;/sup&gt;</th>
<th>$\hat{y}_j(r)$</th>
<th>$\hat{w}_{1}(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>xz 1 2 3 4   5 6 7 8</td>
<td></td>
</tr>
<tr>
<td>(\bar{y}_j(5))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.02 .15 .11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.14 .15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{w}_{5}^2)</td>
<td>.034 .536 .583</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\bar{y}_j(6))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.01 .07 .05 .05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-.06 -.09 -.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>.06 .05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\hat{w}_{6}^2)</td>
<td>.032 .109 .140 .141</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Where: \(\hat{w}_{i}(r)\) represents the ith weight of the ith set of the sequence and \(\hat{w}_{i}^2\) represents the estimated multiple correlation coefficient relevant to the ith estimating equation of \(\bar{y}_j(r) (i = 1, ..., n^r < n_r) \ (r = 1, ..., 6)\).
\[ \hat{g}^*_1 = .14\bar{y}_{16}(1) + .11\bar{y}_{43}(1) + .14\bar{y}_{40}(1) \]

where \((\hat{\gamma}^*_{21}.\bar{y}(1) = .715)\)

\[ \hat{g}^*_2 = .13\bar{y}_{35}(2) + .11\bar{y}_{45}(2) + .11\bar{y}_{49}(2) + .11\bar{y}_{50}(2) \]

where \((\hat{\gamma}^*_{22}.\bar{y}(2) = .868)\)

\[ \hat{g}^*_3 = .14\bar{y}_{25}(3) + .17\bar{y}_{8}(3) + .09\bar{y}_{33}(3) \]

where \((\hat{\gamma}^*_{31}.\bar{y}(3) = .826)\)

\[ \hat{g}^*_4 = .14\bar{y}_{39}(4) + .16\bar{y}_{20}(4) + .12\bar{y}_{42}(4) + .08\bar{y}_{48}(4) \]

where \((\hat{\gamma}^*_{42}.\bar{y}(4) = .685)\)

\[ \hat{g}^*_5 = .11\bar{y}_{11}(5) + .15\bar{y}_{15}(5) + .06\bar{y}_{37}(5) \]

where \((\hat{\gamma}^*_{51}.\bar{y}(5) = .583)\)

\[ \hat{g}^*_6 = .05\bar{y}_{21}(6) - .08\bar{y}_{9}(6) + .05\bar{y}_{26}(6) \]

where \((\hat{\gamma}^*_{61}.\bar{y}(6) = .140)\)

and where from Chapter III (section 3.5)

\[ \hat{g}^*_r = \hat{w}(r)\bar{y}(r) \quad r = 1, \ldots, 6. \]

It is evident that the estimation, \(\hat{g}^*_6\), of \(\bar{g}_6\) is not an effective estimation in that only 14% of the variance of \(\bar{g}_6\) is explained or predicted.

The remaining estimations appear more reasonable.

The estimation of \(\bar{g}_1\) accounts for about 71% of the
variance of $\bar{g}_1$ using three variables; that of $\bar{g}_2$ and $\bar{g}_3$ account for over 80% of the variance to be predicted using four and three variables, respectively; the estimation of $\bar{g}_4$ involves four variables and predicts about 68% of the variance of $\bar{g}_4$; and finally about 58% of the variance of $\bar{g}_5$ is estimated using three variables.
CHAPTER VIII

SUMMARY, LIMITATIONS AND FURTHER RESEARCH

0. Introduction

In this chapter, the results of the investigation are briefly summarized (part 1), some of the limitations of the investigation are discussed (part 2), and some possible further research is suggested (part 3).

1. Summary of Results

Six semantic factors are tentatively identified from the sample of observations of $E^0(Q, P^a, C^a)$. A summary of these factors and aspects relevant to them is shown in Table 10.

Of the six semantic factors derived, $g^1_1$, the necessity factor, is dominant and it would seem reasonable to hypothesize that: (1) one of the semantic factors of $g^+$ is a type of necessity factor representing a dimension of specification associated with $Q^+$ (the universe of semantic scales) and having to do with the extent to which an audit technique is felt to be indispensable in a given auditing situation, and (2) $g^1_1$ represents a reasonably close approximation of this semantic factor (of $g^+$) and
TABLE 10
SUMMARY OF EXPERIMENTAL RESULTS

<table>
<thead>
<tr>
<th>Semantic Factor Variable</th>
<th>Suggested Semantic Identification (Table 3)</th>
<th>Index of Internal Importance ( \frac{\text{var}(u^1)}{p} \times 100 ) (Table 5)</th>
<th>Differentiating Potential Over Organizational Categories According to Kruskal-Wallis Test of Differences of Average Semantic Scores (Table 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{g}_1 )</td>
<td>Necessity Factor</td>
<td>20.9%</td>
<td>Significant ((a = .001))</td>
</tr>
<tr>
<td>( \bar{g}_2 )</td>
<td>Interest Factor</td>
<td>3.6</td>
<td>Insignificant ((a &gt; .90))</td>
</tr>
<tr>
<td>( \bar{g}_3 )</td>
<td>Validity-Strength Factor</td>
<td>5.2</td>
<td>Insignificant ((a = .70))</td>
</tr>
<tr>
<td>( \bar{g}_4 )</td>
<td>Logic-Utility Factor</td>
<td>4.0</td>
<td>Significant ((a = .05))</td>
</tr>
<tr>
<td>( \bar{g}_5 )</td>
<td>Difficulty Factor</td>
<td>3.8</td>
<td>Insignificant ((a = .80))</td>
</tr>
<tr>
<td>( \bar{g}_6 )</td>
<td>Scope Factor</td>
<td>3.3</td>
<td>Insignificant ((a = .70))</td>
</tr>
<tr>
<td>Semantic Factor Variable</td>
<td>Number of &quot;Significant&quot; Deviations in the Rankings of Organizational Categories on the Basis of Average Semantic Factor Scores (Table 8)</td>
<td>Percentage of Variance Accounted for by &quot;Best&quot; Estimation Equation $\tilde{\gamma}_2$ $\frac{\sum r \cdot y(r)}{\sum y^2} \times 100$ (Table 9)</td>
<td>Number of Significant Predictor Variables of the &quot;Best&quot; Estimation $n^*$ (Table 9)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>$\tilde{g}_1$</td>
<td>1</td>
<td>71.5%</td>
<td>3</td>
</tr>
<tr>
<td>$\tilde{g}_2$</td>
<td>2</td>
<td>86.8</td>
<td>4</td>
</tr>
<tr>
<td>$\tilde{g}_3$</td>
<td>2</td>
<td>82.6</td>
<td>3</td>
</tr>
<tr>
<td>$\tilde{g}_4$</td>
<td>1</td>
<td>68.5</td>
<td>4</td>
</tr>
<tr>
<td>$\tilde{g}_5$</td>
<td>3</td>
<td>58.3</td>
<td>3</td>
</tr>
<tr>
<td>$\tilde{g}_6$</td>
<td>4</td>
<td>14.0</td>
<td>3</td>
</tr>
</tbody>
</table>
corresponding semantic dimension. This postulation is based on the fact that the estimated variance of the corresponding principal component variable, $\tilde{u}_1^{(1)}$, accounts for $1/5$ of the total experimental variance, the factor differentiates significantly on at least one external criterion (organizational category) and it exhibits a reasonably consistent ranking of organizational categories (on the basis of estimated average semantic scores) as compared to a ranking one might feel desirable (with only one significant deviation). Moreover, more than 70% of its variance is accounted for in an estimation involving three semantic predictor variables, all of which are significant predictors.

Beyond this tentative postulation, nothing can be conclusively stated in the absence of further data regarding the remaining derived semantic factors. $\tilde{g}_4$, the logic-utility factor, appears to differentiate significantly over the organizational categories and has only one significant deviation in the ranking of these categories, but the estimated variance of the corresponding principal component variable, $\tilde{u}_4^{(1)}$, is only 4% of the total variance. Moreover, only 68.5% of its variance is accounted for in an estimation involving four significant predictor variables. It appears further experimentation is necessary before these remaining and tentatively identified semantic
factors can be postulated, with a reasonable degree of confidence, as stable and meaningful approximations of any of the actual semantic factors of $g^+$ and corresponding semantic dimensions.

2. Limitations of the Investigation

The limitations of the investigation include:

(1) The investigation is exploratory in nature and hence lacks the statistical sophistication and objectivity present in the data analysis of studies which are less preliminary. This is, of course, also due to the nature of the model developed for the investigation. This lack of objectivity and statistical sophistication is most conspicuous in the process of estimating the number, $k$, of semantic factors (Chapter III), a process which at one point (stage 2 of the process) is highly subjective.

(2) The investigation is of the type that ideally requires a sequence of experiments, each representing an improvement in terms of incorporating the results of previous experimentation. This is particularly true with respect to the selection of the set, $Q$, of semantic specifiers. Each experiment should involve a $Q$ which reflects, insofar as its semantic content, the semantic factor variables identified from the previous $Q$ and corresponding experiment, with the underlying
goal being the ultimate identification and
approximation of the total set of actual semantic
factors of $g^+$, representing the set of semantic
dimensions associated with the universe, $Q^+$, of
semantic specifiers.

(3) If the investigation involves only one experiment,
external criteria must be brought into the evaluation
phase in order that anything tentatively conclusive
can be stated. For the current study, this represents
a weakness insofar as only one external criterion was
used in the evaluation. The criterion (organizational
categories from which the auditors were obtained)
chosen, however, is by no means an insignificant one
if one assumes that differences in semantic perception
in terms of semantic factors should result and depend
upon the organizational category of the auditors
involved. It represents a first step in an evaluation
of such factors.

3. Possible Further Research

As stated in section 1 and implied in section 2,
further experimentation is needed involving different $Q$'s.
A logical next step in this regard is the selection of a
$Q$ so as to represent as much as possible, on the basis of
semantic content, the six semantic factors tentatively
identified in the current study. Also, consideration
should be given to using a different number, p, of semantic scales. For \( R^0(Q, P^a, C^a) \), the selection of \( p = 50 \) was arbitrary and more definite results might be obtained by selecting \( p > 50 \).

The definition of \( C \) as the domain of audit-technique concepts is also, to some extent, arbitrary. There are other equally important domains representing other segments of an auditor's task environment. For example, a domain of "audit-evidence" concepts, representing the results of the application of audit techniques (i.e., representing different types of audit evidences such as statements by independent third parties, authoritative documents, etc.), could be semantically investigated. Or, perhaps, a domain of "financial assertion" concepts, representing the audit problems to be solved by the auditing process, could be defined and studied. No doubt, other domains could be found relating to other portions of the task environment.

Once a set of semantic factors has been reasonably identified as an approximation of the actual set of factor variables for a given \( C \) and \( P \), they should then be explored to determine what, if anything, they mean. Or equivalently, they should be studied to determine what attributes of an auditor are being specified when measurements or estimations of the factor variables are made and to determine the
implications, if any, with respect to audit behavior. Such questions as the following appear relevant to this problem:

(1) How do the semantic perceptions of auditors vary according to such things as age, quantity and quality of experience, type of education, size of office, etc.?

(2) How do semantic perceptions relate to the type of audit methodology selected in a given auditing situation, to the way in which that methodology is applied and to the nature of resulting audit opinion and judgment?

(3) How do semantic perceptions of an auditor relate to the organization's evaluation of his effectiveness (or lack of effectiveness)?

(4) What is the effect, if any, of formal "staff training" programs utilized by the CPA organization? Is there a significant change and if so, in what direction?

(5) To what extent do semantic perceptions vary among CPA organizations?

These questions do not constitute an exhaustive listing but are suggestive of the types of research questions which could be investigated.
APPENDIX A

VERIFICATION OF PROPERTIES OF THE MODEL
GIVEN IN CHAPTER II, SECTION 1

\begin{align*}
\text{(1)} \\
n \text{var}(z) &= \text{var}(z^{(1)}) + \text{var}(z^{(2)}) \\
&= E(z^{(1)}z^{(1)}) + E(z^{(2)}z^{(2)}) \\
&= E(bgk^TB^T) + E(L^{(2)}d_2x^{(2)}x^{(2)})D_dL^{(2)} \\
&= BE(Tx^{(1)}x^{(1)}TB^T + L^{(2)}d_2E(x^{(2)}x^{(2)})D_dL^{(2)}) \\
&= BT^TB^T + L^{(2)}D_d^2L^{(2)} \\
&= ATT^TA^T + L^{(2)}D_d^2L^{(2)} \\
&= L^{(1)}D_d^2L^{(1)} + L^{(2)}D_d^2L^{(2)} \\
&= (L^{(1)}, L^{(2)}) \begin{pmatrix}
D_{d_1}^2 & 0 \\
0 & D_{d_2}^2
\end{pmatrix}
\begin{pmatrix}
L^{(1)} \\
L^{(2)}
\end{pmatrix} \\
&= LD_d^2L^T \\
&= LD_dE(xx^T)D_dL^T \\
&= E[(LD_dx)(LD_dx)^T]
\end{align*}
\( \text{cov}(z(1), z(2)) = E(z(1)z(2)') \)

\[
= E[(Bg)(L(2)D_d x(2)')] \\
= E(BT'x(1)x(2)'D_d L(2)') \\
= \phi_{pxp}, \text{ since by definition of } x, \\
E(x(1)x(2)') = \phi_{kxm}
\]

(iii)

a. \( \text{cov}(z, g) = E(zg') \)

\[
= E[(z(1) + z(2)x(1)')] \\
= E(BT'x(1)x(1)' + L(2)D_d x(2)x(1)') \\
= BT'E(x(1)x(1)')T \\
= BT'T \\
= B
\]

b. \( \text{cov}(z(1), g) = E(z(1)g') \)

\[
= E(Bgg') \\
= E(BT'x(1)x(1)'T) \\
= B
\]
c. \( \text{cov}(z, x^{(1)}) = E(zx^{(1)}) \)
\[ = E[(z^{(1)} + z^{(2)})x^{(1)}] \]
\[ = E(BT'x^{(1)}x^{(1)}) + L^{(2)}D_{d_2}x^{(2)}x^{(1)} \]
\[ = BT'E(x^{(1)}x^{(1)}) \]
\[ = ATT' \]
\[ = A \]

d. \( \text{cov}(z, x^{(2)}) = E(zx^{(2)}) \)
\[ = E[(z^{(1)} + z^{(2)})x^{(2)}] \]
\[ = E(BT'x^{(1)}x^{(2)}) + L^{(2)}D_{d_2}x^{(2)}x^{(2)} \]
\[ = L^{(2)}D_{d_2}E(x^{(2)}x^{(2)}) \]
\[ = L^{(2)}D_{d_2} \]

(iv)

a. \( \text{trace(var}(z^{(1)}) = \text{trace}(M^{(1)}) \)
\[ = \text{trace}(BB') \]
\[ = \text{trace}(B'B) \]
\[ = \text{trace}(T'A'AT) \]
\[ = \text{trace}(T'D_{d_1}L^{(1)}'L^{(1)}D_{d_1}T) \]
\[ = \text{trace}(T'D_{d_1}^2T) \]
\[ = \text{trace}(D_{d_1}^2) \]
b. \( \text{trace}(\text{var}(z^{(2)})) = \text{trace}(M^{(2)}) \)
   \[ = \text{trace}(E(z^{(2)}z^{(2) 
   \hat{\cdot}})) \]
   \[ = \text{trace}(L^{(2)}D_{d2}^2L^{(2) \hat{\cdot}}) \]
   \[ = \text{trace}[(L^{(2)}D_{d2})(D_{d2}L^{(2) \hat{\cdot}})] \]
   \[ = \text{trace}[(D_{d2}L^{(2) \hat{\cdot}})(L^{(2)}D_{d2})] \]
   \[ = \text{trace}(D_{d2}^2) \]

\[ = \text{trace}(D_{d1}^2) + \text{trace}(D_{d2}^2) \]
   \[ = \text{trace}(D_d^2) \]
   \[ = \text{trace}(R) \]
   \[ = p \]
APPENDIX B

VERIFICATION OF PROPERTY (1) AND (11) OF THE ESTIMATION METHOD OF CHAPTER III

(1) \( \text{cov}(e_r, y(r)) = E(e_r y(r))' \)

\[
= E(g_r - g_r^*) y(r)'
\]

\[
= E(g_r y(r))' - E(g_r^* y(r))'
\]

\[
= \Sigma_{yg} (r) y(r)' - \Sigma_{yg} (r) y(r) y(r)'
\]

\[
= \Sigma (r) y(r)' - \Sigma (r) y(r) y(r)'
\]

\[
= \Sigma_{yg} (r) y(r) + \sigma^2
\]

\[
= 1
\]

(11) \( \text{var}(g_r) = E(g_r^* + e_r)(g_r^* + e_r)' \)

\[
= E(g_r^* g_r^* + e_r e_r^* + g_r e_r^* + e_r e_r^*)
\]

\[
= E(W(r)^' y(r) y(r) W(r)) + E(e_r e_r)
\]

\[
= \Sigma (r) y(r) y(r) y(r) y(r) y(r) + \sigma^2
\]

\[
= \Sigma_{yg} (r) y(r) y(r) y(r) + \sigma^2
\]

\[
= 1
\]
APPENDIX C

INSTRUCTION SET USED IN THE QUESTIONNAIRE

OF $E^O(q, p^a, c^a)$

General

The immediate objective of this questionnaire is to elicit your reactions to certain auditing concepts by having you judge them against a series of descriptive scales.

In making these judgments, you are asked to respond on the basis of what the concepts mean to you as an auditor. This is important since the results of your judgments are to be used in the development of an operational instrument designed to measure the meanings of auditing concepts for auditors.

The Concepts

The questionnaire consists of three concepts. Each concept is selected to represent a particular audit technique used by auditors for obtaining audit evidence. The three concepts are:

1. Physical Examination and Count
2. Inquiry
3. Confirmation

Format

Each concept requires two pages of scales and a new concept is introduced every third page.

Immediately beneath each concept is an illustrative audit procedure involving the use of the particular audit technique represented by the concept. You are requested, however, to judge the concept and not the audit procedure.
The latter is given only for purposes of clarifying the intended nature of the technique represented by the concept.

How To Use The Scales

If you feel that the concept at the top of the page is VERY CLOSELY RELATED to one end of the scale, you should place your check-mark:

complex : X:___:___:___:___:___: simple

OR

complex : ___:___:___:___:___:___: X: simple

If you feel that the concept is QUITE CLOSELY RELATED to one or the other end of the scale (but not extremely), you should place your checkmark:

boring : ___:___:___:___:___:___: interesting

OR

boring : ___:___:___:___:___:___: X:___: interesting

If the concept seems ONLY SLIGHTLY RELATED to one end as opposed to the other end (but not really neutral), then you should check:

cautious : ___:___:___:___: ___:___: X:___: rash

OR

cautious : ___:___:___:___: ___:___: X:___: rash

The direction toward which you check, of course, depends upon which of the two ends of the scale seem most characteristic of the concept you are judging.

If you consider the concept to be NEUTRAL on the scale, both sides of the scale EQUALLY ASSOCIATED with the concept, or if the scale is COMPLETELY IRRELEVANT, unrelated to the concept, then check in the middle space:

unnecessary : ___:___:___:___:___:___: necessary
IMPORTANT: (1) Place your check-marks in the middle of spaces, not on the boundaries.

\[ \text{THIS \quad \text{NOT}} \]
\[ X_1 : X_2 : \ldots : X_n \]

(2) Be sure you check every scale for every concept—do not omit any.

(3) Never put more than one check-mark on a single scale.

Approach To Be Taken In Filling Out Questionnaire

These factors are important to your approach in completing the questionnaire:

1. Do not look back and forth through the scales. Do not try to remember how you checked a particular scale for some previous concept in the questionnaire.

2. Make each scale a separate and independent judgment.

3. Work at fairly high speed. Do not puzzle over individual items. It is your first impression, your immediate feeling about the items that is desired.

Completion Time

Although it is estimated that it will take about 15 minutes to complete the questionnaire, do not be concerned if it takes you either more or less time. This is simply an estimate. To help improve the estimate, please note below the approximate time it takes you to finish.

Approximate completion time: _________ minutes
APPENDIX D

A LISTING OF THE THREE AUDIT-TECHNIQUE CONCEPTS
AND ILLUSTRATIVE AUDIT PROCEDURES USED
IN THE QUESTIONNAIRE OF $E^0(q, P^a, C^a)$

<table>
<thead>
<tr>
<th>Concept</th>
<th>Illustrative Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Physical Examination</td>
<td>Examine and count the asset (e.g., securities) in the presence of the custodian, undertaking simultaneous count or control of all other related $L.$'s (e.g., cash funds).</td>
</tr>
<tr>
<td>and Count</td>
<td></td>
</tr>
<tr>
<td>2. Inquiry</td>
<td>Determine company and industry policies and practices by making suitable inquiries of management.</td>
</tr>
<tr>
<td>3. Confirmation</td>
<td>Confirm account balance(s) and aspects thereof (e.g., validity, accuracy, description of item) by written communication with an outside party.</td>
</tr>
</tbody>
</table>
APPENDIX E

A LISTING OF THE 50 SEMANTIC SCALES USED IN THE QUESTIONNAIRE OF $E^0(Q, p^a, c^a)$

unadaptable :__:__:__:__:__:__:__: adaptable
risky :__:__:__:__:__:__:__: safe
restrictive :__:__:__:__:__:__:__: unrestrictive
imaginative :__:__:__:__:__:__:__: unimaginative
avoidable :__:__:__:__:__:__:__: unavoidable
penetrating :__:__:__:__:__:__:__: superficial
unsupporting :__:__:__:__:__:__:__: supporting
inconclusive :__:__:__:__:__:__:__: conclusive
general :__:__:__:__:__:__:__: specific
sound :__:__:__:__:__:__:__: unsound
convenient :__:__:__:__:__:__:__: inconvenient
illogical :__:__:__:__:__:__:__: logical
unsophisticated :__:__:__:__:__:__:__: sophisticated
obstructive :__:__:__:__:__:__:__: helpful
demanding :__:__:__:__:__:__:__: undemanding
indispensable :__:__:__:__:__:__:__: dispensable
subjective :__:__:__:__:__:__:__: objective
formal :__:__:__:__:__:__:__: informal
<table>
<thead>
<tr>
<th>Word</th>
<th>Antonym</th>
</tr>
</thead>
<tbody>
<tr>
<td>verifiable</td>
<td>unverifiable</td>
</tr>
<tr>
<td>disorderly</td>
<td>orderly</td>
</tr>
<tr>
<td>universal</td>
<td>particular</td>
</tr>
<tr>
<td>useless</td>
<td>useful</td>
</tr>
<tr>
<td>routine</td>
<td>varied</td>
</tr>
<tr>
<td>sufficient</td>
<td>insufficient</td>
</tr>
<tr>
<td>complete</td>
<td>incomplete</td>
</tr>
<tr>
<td>broad</td>
<td>narrow</td>
</tr>
<tr>
<td>unsystematic</td>
<td>systematic</td>
</tr>
<tr>
<td>rational</td>
<td>intuitive</td>
</tr>
<tr>
<td>boring</td>
<td>interesting</td>
</tr>
<tr>
<td>inconsistent</td>
<td>consistent</td>
</tr>
<tr>
<td>active</td>
<td>passive</td>
</tr>
<tr>
<td>multi-purposed</td>
<td>single-purposed</td>
</tr>
<tr>
<td>infallible</td>
<td>fallible</td>
</tr>
<tr>
<td>impractical</td>
<td>practical</td>
</tr>
<tr>
<td>monotonous</td>
<td>unmonotonous</td>
</tr>
<tr>
<td>significant</td>
<td>insignificant</td>
</tr>
<tr>
<td>simple</td>
<td>complex</td>
</tr>
<tr>
<td>ineffective</td>
<td>effective</td>
</tr>
<tr>
<td>direct</td>
<td>indirect</td>
</tr>
<tr>
<td>necessary</td>
<td>unnecessary</td>
</tr>
<tr>
<td>unimportant</td>
<td>important</td>
</tr>
<tr>
<td>concrete</td>
<td>abstract</td>
</tr>
<tr>
<td>optional</td>
<td>required</td>
</tr>
</tbody>
</table>
unexploratory  :___:___:___:___:___:___: exploratory
untedious     :___:___:___:___:___:___: tedious
unchallenging :___:___:___:___:___:___: challenging
productive    :___:___:___:___:___:___: unproductive
deliberate    :___:___:___:___:___:___: random
stimulating   :___:___:___:___:___:___: unstimulating
repetitive    :___:___:___:___:___:___: unrepetitive
APPENDIX F

FACTOR ANALYSIS AS A MODEL FOR SEMANTIC SPECIFICATION BEHAVIOR

Purpose

The purpose of this appendix is to indicate the basic nature of the factor model and point out two fundamental differences between it and the model of Chapter II. One of these differences is the fact that the variables of $g$ are not unique.

Statement of the Factor Model

For a given experiment $E(Q, P, C)$, the factor model can be stated:

\[ z = z^{(1)} + z^{(2)} \]

where:

a. $z^{(1)}$ -- a p x l random vector of standardized semantic variables (Chapter II, section 1.10).

b. $g$ -- a k x l random vector of semantic factor variables (Chapter II, section 1.7).

c. $z^{(1)} = Bg$

d. $\text{var}(g) = I_{k \times k}$
e. \( z^{(2)} = Us \), where \( s \) is a px1 random vector of "specific" factor variables.

f. \( \text{var}(z^{(2)}) = U^2 \), a diagonal, nonsingular matrix of order \( p \).

g. \( \text{cov}(g, z^{(2)}) = \phi_{kxp} \)

h. \( B = AT \)

i. \( A \) - a pxk matrix of rank \( k \)

j. \( T \) - a kxk transformation matrix defined as in Chapter II (section 3).

**Properties of the Factor Model**

The following properties of the model, given without proof, can readily be verified:

\[ \text{var}(z) = BB' + U^2 = R \]

\[ \text{cov}(z^{(1)}, z^{(2)}) = E(z^{(1)}z^{(2)'}) = \phi_{pxp} \]

\[ \text{cov}(z, g) = E(zg') = B \]

\[ \text{cov}(z, s) = E(zs') = U \]

\[ \text{trace(var} z^{(1)}) = \text{trace (BB')} \]

\[ \text{trace(var} z^{(2)}) = \text{trace (U^2)} \]

\[ \text{trace(var} z) = \text{trace (R)} = p \]
Discussion of the Model

From the statement and the properties of the model, it is seen that the factor model resembles the model given in Chapter II (section 3). There are, however, two fundamental differences:

(a) The variables of \( z^{(2)} \) for the factor model are defined to be statistically independent according to (1)f. That is,

\[
E(z_i^{(2)}z_j^{(2)}) = \begin{cases} 
0 & i \neq j \\
\mu_i^2 & i = j
\end{cases}
\]

where \( \mu_j^2 \) is the jth element of \( U^2 \). This is contrary to the model of Chapter II in which the variables of \( z^{(2)} \) are not in general statistically independent.

That is,

\[
\text{var}(z^{(2)}) = M^{(2)} \neq \text{a diagonal matrix}.
\]

(b) The semantic factor variables of \( g \) are not defined as transformations of the first k standardized principal component variables of \( x^{(1)} \) as they are in the model of Chapter II. In fact, as we shall see, they are not uniquely definable at all.

The first difference implies a difference in interpretation. Whereas for the model of Chapter II, the variables of \( g \) are viewed as the "sources" of the total variance of the variables of \( z \), for the factor model, they
are considered as the sources of "common" variance of the variables of $z$. This is clearly seen by noting that

$$E(z_i'z_j') = E(z_i^{(1)}z_j^{(1)'}), \quad i \neq j$$

$$= E(z_i^{(1)}z_j^{(1)'}), + E(z_i^{(2)}z_j^{(2)'}) \quad i = j$$

From (2), we see that the intercorrelation of the $i$th and $j$th variables involves only the variables of $z^{(1)}$, which are, in turn, defined as linear functions of the variables of $g$. On the other hand, the total variance of the $i$th (or $j$th) variable, according to the factor model, is the sum of the "common" variance, $E(z_i^{(1)}z_i^{(1)'})$, and $E(z_i^{(2)}z_i^{(2)'})$, which is not common, or which is "specific" to the $i$th variable.

The second difference relating to the lack of uniqueness of the semantic factors of $g$ is demonstrated in the following theorem. In the theorem, the variables of $g$ are broken into two parts, one of which is completely arbitrary, and it is shown that such partially arbitrary variables satisfy the model as stated in (1)a-j.
A Theorem Demonstrating Lack of Uniqueness of Factor Variables

Statement of Theorem

Let $F$ be a $k \times k$ matrix such that

$$FF' = I - B'R^{-1}B$$

Let $v$ be a $k \times 1$ random vector of arbitrary random variables such that

(a) $\text{var}(v) = I_{k \times k}$

(b) $E(v) = \phi_{k \times 1}$

(c) $\text{cov}(z, v) = \phi_{p \times k}$

Then it can be shown that

(4) 
$$g = g_d + g_i = B'R^{-1}z + Fv$$

and

(5) 
$$z^{(2)} = z_d^{(2)} + z_1^{(2)} = U^2R^{-1}z - BFv$$

satisfy the factor model as stated in (1)a-j. Specifically, it can be shown that $g$ and $z^{(2)}$ as defined in (4) and (5), satisfy

(i) $\text{var}(g) = I_{k \times k}$

(ii) $\text{var}(z^{(2)}) = U^2$

---

1This theorem (and proof) was first developed by Guttman (1955). The particular form of the theorem and proof used here was suggested by Professor Schoneman of the Psychology Department at Ohio State University.
Proof of Theorem

(i) \( \text{var}(g) = E(g g') \)

\[ = E[(B' R^{-1} z + Fv)(B' R^{-1} z + Fv)'] \]

\[ = E[B' R^{-1} z z' R^{-1} B + Fv F'] \]

\[ = B' R^{-1} B + FF' \]

\[ = B' R^{-1} B + (I - B' R^{-1} B) \]

\[ = I_{kxk} \]

(ii) \( \text{var}(z^{(2)}) = E(z^{(2)} z^{(2)'}) \)

\[ = E[(z_d^{(2)} + z_1^{(2)})(z_d^{(2)} + z_1^{(2)'})'] \]

\[ = E[(U^2 R^{-1} z - BFv)(U^2 R^{-1} z - BFv)'] \]

\[ = U^2 R^{-1} E(z z' R^{-1} U^2 + BF(Evv') F'B' \]

\[ = U^2 R^{-1} U^2 + BFBB' \]

\[ = U^2 R^{-1} U^2 + B(I - B' R^{-1} B)B' \]

\[ = U^2 R^{-1} U^2 + R - U^2 - [(R - U^2) R^{-1} (R - U^2)] \]

\[ = U^2 R^{-1} U^2 + R - U^2 - (I - U^2 R^{-1})(R - U^2) \]

\[ = U^2 R^{-1} U^2 + R - U^2 - R + U^2 + U^2 - U^2 R^{-1} U^2 \]

\[ = U^2 \]
(iii) $\text{cov}(g, z^{(2)}) = E(gz^{(2)})$  

\[
= E[(B'R^{-1}z + Fv)(U^2 R^{-1}z - BFv)'] \\
= B'R^{-1}E(zz')R^{-1}U^2 - FE(vv')F'B' \\
= B'R^{-1}U^2 - FF'B' \\
= B'R^{-1}U^2 - (I - B'R^{-1}B)B' \\
= B'R^{-1}U^2 - B' + B'R^{-1}BB' \\
= B'R^{-1}(U^2 + BB') - B' \\
= B'R^{-1}(I) - B' \\
= B' - B' \\
= \phi_kxk
\]

(iv) $z = Bg + z^{(2)}$  

\[
= B(B'R^{-1}z + Fv) + (U^2 R^{-1}z - BFv) \\
= BB'R^{-1}z + BFv + U^2 R^{-1}z - BFv \\
= (BB' + U^2)R^{-1}z \\
= (R)R^{-1}z \\
= z \\
= z^{(1)} + z^{(2)}
\]
The theorem demonstrates that $g$ which is part indeterminate as in (4) will satisfy the factor model of (1). Consequently, the variables of $g$ are not uniquely definable and cannot be measured or estimated. And further, any entity or concept which the variables of $g$ are interpreted to represent are likewise not uniquely definable, not measurable and not estimable. This includes the case in which the variables of $g$ are designated semantic factors and interpreted as representing semantic dimensions of specification.
BIBLIOGRAPHY

Books


Articles


Other Publications