AD-A246 310

NAVAL POSTGRADUATE SCHOOL
Monterey, California

DTIC
ELECTE
FEB 21 1992
S B D

THESIS

A TAXONOMICAL STRUCTURE FOR CLASSIFYING THE SERVICES PROCURED BY THE FEDERAL GOVERNMENT

by

Scott Thomas Allen

December, 1991

Thesis Advisor: David V. Lamm

Approved for public release; distribution is unlimited.

92-04357

92 2 19 039

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5 MONITORING ORGANIZATION REPORT NUMBER(S)
6a NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b OFFICE SYMBOL (If applicable) 036	7a NAME OF MONITORING ORGANIZATION Naval Postgraduate School
6c ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	7b ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000
8a NAME OF FUNDING/SPONSORING ORGANIZATION 8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c ADDRESS (City, State, and ZIP Code)	10 SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO / PROJECT NO / TASK NO / WORK UNIT ACCESSION NO

11 TITLE (Include Security Classification)
A TAXONOMICAL STRUCTURE FOR CLASSIFYING THE SERVICES PROCURED BY THE FEDERAL GOVERNMENT (UNCLASSIFIED)

12 PERSONAL AUTHOR(S)
Scott Thomas Allen

13a TYPE OF REPORT	13b TIME COVERED	14 DATE OF REPORT (Year Month Day)	15 PAGE COUNT
Master's Thesis	FROM _____ TO _____	December, 1991	232

16 SUPPLEMENTARY NOTATION
THE VIEWS EXPRESSED IN THIS THESIS ARE THOSE OF THE AUTHOR AND DO NOT REFLECT THE OFFICIAL POLICY OR POSITION OF THE DEPARTMENT OF DEFENSE OR THE U.S. GOVERNMENT.

17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Service Procurement Taxonomy

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

This thesis was an attempt to develop a taxonomical scheme that practitioners may employ in classifying services that are procured by the Federal Government along a continuum from procurements that are relatively simple to those that are strategically complex. A secondary research objective was to determine what characteristics are appropriate for classifying services on a strategic basis.

A literature review, expert interviews, and survey using 20 heterogeneous sample services were conducted to determine the relationship between characteristics and services. Cluster analysis was used to group services into categories with similar compositions of selected characteristics.

A taxonomical structure was developed for classifying services into five categories. Potential benefits may arise via application to staffing and directing of procurement functions and refinement of procurement policy. It is recommended that the taxonomical model resulting from this thesis be validated and refined through further use.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION Unclassified
22a NAME OF RESPONSIBLE INDIVIDUAL David V. Lamm	22b TELEPHONE (Include Area Code) (408) 646-2775 22c OFFICE SYMBOL AS/LT

DD Form 1473, JUN 86 Previous editions are obsolete SECURITY CLASSIFICATION OF THIS PAGE

S/N 0102-LF-014-6603

Approved for public release; distribution is unlimited.

A Taxonomical Structure for Classifying the Services Procured by the Federal Government

by

Scott Thomas Allen
Captain, United States Marine Corps
B.A., University of Virginia, 1984

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
December 1991

Author: _____
Scott Thomas Allen

Approved by: _____
David V. Lamm, Thesis Advisor

_____
Kishore Sengupta, Second Reader

_____
David R. Whipple, Chairman
Department of Administrative Sciences

ABSTRACT

This thesis was an attempt to develop a taxonomical scheme that practitioners may employ in classifying services that are procured by the Federal Government along a continuum from procurements that are relatively simple to those that are strategically complex. A secondary research objective was to determine what characteristics are appropriate for classifying services on a strategic basis.

A literature review, expert interviews, and survey using 20 heterogeneous sample services were conducted to determine the relationship between characteristics and services. Cluster analysis was used to group services into categories with similar compositions of selected characteristics.

A taxonomical structure was developed for classifying services into five categories. Potential benefits may arise via application to staffing and directing of procurement functions and refinement of procurement policy. It is recommended that the taxonomical model resulting from this research be validated and refined through further use.

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

TABLE OF CONTENTS

I. INTRODUCTION 1

A. BACKGROUND 1

B. OBJECTIVES 5

C. RESEARCH QUESTIONS 6

D. RESEARCH METHODOLOGY 7

E. SCOPE, LIMITATIONS, AND ASSUMPTIONS 10

F. REVIEW OF LITERATURE 12

G. ORGANIZATION OF STUDY 13

II. TAXONOMY BACKGROUND 16

A. INTRODUCTION 16

B. DEFINITION OF TERMS 18

C. WHY WE CLASSIFY 20

D. TWO GENERAL TYPES OF CLASSIFICATION SCHEMES . . 21

E. CLASSIFICATION PRINCIPLES 22

F. CLASSIFICATION SCHEMES WITHIN THE GOVERNMENT . 25

G. THE NEED FOR A STRATEGIC CLASSIFICATION SCHEME 30

H. SUMMARY 32

III. DEVELOPMENT OF A TAXONOMICAL MODEL 33

A. INTRODUCTION 33

B. MAIN OBJECTIVE AND CONCEPTUAL BASIS 34

C. DETERMINATION OF THE CHARACTERISTICS 35

D. OPERATIONALIZING THE SCHEME 61

E. PRE-TESTING THE MODEL 75

F. REVISION OF THE MODEL 85

G. SUMMARY 89

IV. DATA COLLECTION AND PRELIMINARY ANALYSIS 91

A. INTRODUCTION 91

B. DATA COLLECTION 91

C. AN "A PRIORI" CLASSIFICATION 100

D. PREPARATION FOR CLUSTER ANALYSIS 105

E. INITIAL CLUSTER ANALYSIS RESULTS 110

F. SUMMARY 136

V. SIMPLIFYING THE TAXONOMICAL MODEL 138

A. INTRODUCTION 138

B. ANALYSIS OF MODEL CHARACTERISTICS 138

C. STREAMLINING THE MODEL 156

D. RESULTANT TAXONOMICAL SCHEME 167

E. SUMMARY 174

VI. CONCLUSIONS AND RECOMMENDATIONS 176

A. INTRODUCTION 176

B. CONCLUSIONS 176

C. RECOMMENDATIONS 179

D. RESEARCH QUESTIONS 182

E. RECOMMENDATIONS FOR FURTHER RESEARCH 187

F. SUMMARY 188

LIST OF REFERENCES 189

APPENDIX A 193

APPENDIX B 195

APPENDIX C 200

APPENDIX D 209

INITIAL DISTRIBUTION LIST 220

LIST OF TABLES

2-1 MAJOR CATEGORIES IN USE FOR CLASSIFYING & MONITORING GOVERNMENT PROCURED SERVICES 27

2-2 MAJOR CATEGORIES OF SERVICES CITED IN THE FEDERAL ACQUISITION REGULATION 29

3-1 CANDIDATE CHARACTERISTICS: PRELIMINARY CHARACTERISTICS USED BY WENGER 38

3-2 CANDIDATE CHARACTERISTICS: CHARACTERISTICS SUGGESTED BY MULHERN 39

3-3 CANDIDATE CHARACTERISTICS: PROPOSED TO GAIN STRATEGIC MARKETING INSIGHTS CONCERNING PROVIDER-CUSTOMER INTERFACE 40

3-4 CANDIDATE CHARACTERISTICS: A TYPOLOGY OF SERVICE ORGANIZATIONS 41

3-5 CANDIDATE CHARACTERISTICS: SELECTIVE SUMMARY OF PREVIOUSLY PROPOSED SCHEMES FOR CLASSIFYING SERVICES 42

3-6 PRELIMINARY CHARACTERISTICS 46

3-7 ADDITIONAL PRELIMINARY CHARACTERISTICS RECOMMENDED BY PANEL 50

3-8 SUMMARY TITLES FOR PRELIMINARY CHARACTERISTICS . . 53

3-9 PRELIMINARY CHARACTERISTICS GROUPED BY DIMENSION . 56

3-10 PRE-TEST CHARACTERISTICS GROUPED BY DIMENSION . . 58

3-11 SAMPLE SERVICES & RESPECTIVE SIC CODES 78

4-1 SURVEY RESPONSES BY TYPE 98

4-2 "A PRIORI" CLASSIFICATION OF 20 SAMPLE GOVERNMENT PROCURED SERVICES 104

4-3 MEAN VALUE MATRIX 107

4-4 COMPARISON BETWEEN "A PRIORI" MODEL & TWO CLUSTERING METHODS AT FIVE CLUSTER LEVEL USING EXPERTISE & LABOR % OF COST CHARACTERISTICS 112

4-5 COMPARISON BETWEEN "A PRIORI" MODEL & TWO CLUSTERING METHODS AT FIVE CLUSTER LEVEL USING THE TWELVE DATA COLLECTION CHARACTERISTICS 120

4-6 WIDTH OF RANGE OF NORMALIZED AVERAGE DISTANCES BETWEEN CLUSTERS 130

4-7 COMPARISION BETWEEN TWO CLUSTERING METHODS AT THREE AND EIGHT CLUSTER LEVELS USING THE TWELVE DATA COLLECTION CHARACTERISTICS 132

5-1 STANDARD DEVIATION MATRIX 143

5-2 CLUSTERING THE CHARACTERISTICS BY STANDARD DEVIATIONS EQUALLING OR EXCEEDING 1.00 AND BY MEAN STANDARD DEVIATION 145

5-3 FREQUENCY OF TOP THREE CHARACTERISTIC RATINGS . 147

5-4 CLUSTERING THE CHARACTERISTICS BY WEIGHTED AND UNWEIGHTED TOP THREE MEASURES 149

5-5 MEAN OF CHARACTERISTIC VALUES PER CLUSTER . . . 154

5-6 SUMMARY OF THE PROCESS OF CHARACTERISTIC RETENTION/REMOVAL 161

5-7 MEAN VALUES FOR THE CATEGORIES AFTER REMOVAL OF CHARACTERISTICS 165

LIST OF FIGURES

3-1 Pre-test Classification Model 72

5-1 Individual Service Classification Scheme 172

ACKNOWLEDGEMENTS

I wish to express my deep appreciation to Dr. David V. Lamm and Dr. Kishore Sengupta of the Naval Postgraduate School for their comments and suggestions as advisor and second reader during the development of this thesis.

I also wish to acknowledge the assistance and help provided by all of the individuals interviewed for this study. Special thanks go to Ms. Elinor Sue Coates, consultant, San Francisco, CA; Janice C. Goodwin and Diva A. Heitz, Sacramento Air Logistics Center, Sacramento, CA; Emmet E. Hearn, consultant, Los Altos, CA; Sherry A. Mullen and Charles A. and Barbara W. Marshall, Sacramento Army Depot, Sacramento, CA, and; Stanley D. Zemanski, author, Paradise, CA.

Lastly, I would like to express my gratitude to my father, Thomas F. Allen, for emphasizing the importance of education during my childhood and inspiring me to join the United States Marine Corps.

I. INTRODUCTION

A. BACKGROUND

"What is that?" is a familiar question to anyone who has dealt with children. One reason for this identification query is obvious; it provides a common nomenclature with which to communicate. When providing an identification to a child, however, it may be difficult to explain what, for instance, is the difference between a "pot" and a "pan", or in the case of an adult, is the difference between an "industrial engineer" and a "systems engineer". The routine nature of identification thus obscures an implicit purpose of the process; the "...allocation or assignment of unidentified objects to the correct class, once such classes have been established by prior classification." (Sokal, 1974, p. 1116)

The establishment of classes via the process of classification is therefore inherent to any human endeavor. It has been further asserted that classification "...is an important aspect of most sciences" (Sokal, 1974, p. 1115) and that the description and classification of subject matter is a requirement for the establishment of a contracting science (Park, 1986, p. 90). Several classification studies have been conducted concerning Federal Government procurement. Contracting officer tasks (Fowler, 1987) (Page, 1989),

contracting literature (Sweeney, 1989), and goods procured by the Government (Wenger, 1990) have been the subject of recent inquiry.

The subject of services procured by the Government, however, is one where classification efforts appear to be narrow and uncoordinated. The few classification schemes provided for Government procurement organizations either reflect the mandate to "...rely on commercially available sources to provide commercial products and services" or resulted from an amalgamation of statutory and regulatory requirements, and therefore incorporate commercial practices and interests (OMB A-76, 1983, p. 2). While numerous studies have classified services from the strategic perspective of industry (Lovelock, 1983, p. 11), a classification scheme (taxonomy) from the strategic perspective of Government is lacking.

This study endeavors to develop a taxonomy by which Government procured services may be classified from a strategic perspective. It is limited, in its primary scope, to the study of service contracts as defined in paragraph 37.101 of the Federal Acquisition Regulation (FAR) (FAR, 1990, para 37.101). "Construction", as defined in FAR paragraph 36.102 is specifically excluded from this narrow scope since its output may be considered a "good", amenable to

classification with the taxonomy of goods developed by Brian L. Wenger (Wenger, 1990).

The taxonomy of Government procured goods by Wenger is not only an antecedent to this effort; the study specifically recommends research into the development of a classification scheme for Government procured services (Wenger, 1990, p. 101). Goods and services may be viewed as polar extremes, however, there are few "pure" goods or services (Lovelock, 1983, p. 11). Since goods and services and their hybrids (such as the various forms of construction services) constitute the entire procurement domain, the use of these taxonomies should, in fact, provide an overall scheme for the classification of Government procurements. Use of either classification scheme, alone or in tandem, should reveal how good-like or service-like an item such as construction really is. For example, construction of a ship may be contrasted with that of a building. If the goods (materials) are "complex" according to Wenger's classification scheme, as they may be for construction of a ship, the Government may be procuring both "complex" services (*labor*) and "complex" goods. If the goods (*materials*) used on a construction job are "simple" according to Wenger's classification scheme, as they may be for construction of a building, the Government may primarily be procuring a "complex" service (*labor*) (Interview-

Mulhern, 1991). Building construction may therefore be closer to the service extreme, and Government contracting may be modified, for example, to test strategies such as separating procurement of the "simple" material from that of the "complex" labor of a building construction firm.

In its broad scope, a goal of this study is to develop a taxonomy of Government procured services that will complement the taxonomy of Government procured goods. Combined use of classification schemes for goods and services is suggested as an area of future research, but use of a classification scheme for Government procured services, in and of itself, may provide strategic insight by categorizing services across a spectrum from the relatively simple to the complex. One major benefit would be that current Government classifications can be examined to measure their validity, and subsequently be improved or supplanted. A classification scheme could also be used:

a. For the purposes of determining appropriate competitive environment elements.

b. For the purposes of determining the appropriate contract instrument to use. The structure should allow for a better relationship between service and contract instrument. The fixed-price award-fee contract type, for instance, may have an optimal application to one classification of services, and there may be "borderline" classifications where different contract types should be considered.

c. For the purposes of developing and utilizing new methods of contract administration and organization. Categorization may provide a target "market" in which to test new methods. The Government way wish to restrict tests on the concept of competitive contracting offices, for example, to certain categories of services.

d. For the purposes of dividing omnibus, "umbrella" contracts for services into categories which may be properly grouped together for contract administration.

e. For the purposes of highlighting those categories of services which require less statutory and regulatory oversight during contract administration.

f. For the purposes of comparing the salient characteristics of present and prospective service procurements and estimating procurement costs, as in performance evaluation or comparable worth studies.

B. OBJECTIVES

The primary purpose of this research effort is to develop a model which may classify Government procured services on a strategic basis. This taxonomy should differentiate between various classes of services and identify categories within the extremes of the classification scheme.

In parallel with the taxonomy of Government procured goods, the objectives to be achieved include:

1. Determining the characteristics of services, other than their obvious functional differences, to use in classifying.

2. Refining procedures for comparing a sample group of services with the chosen characteristics.

3. Testing of the procedures by use in actual data collection.

4. Deriving a taxonomical model based on the data analysis. (Wenger, 1990, p. 3)

C. RESEARCH QUESTIONS

This research addressed the following questions, which, in their essential character, coincide with those of the taxonomy of Government procured goods:

Primary:

What would be the essential characteristics or features of a taxonomical structure that would classify the services procured by the Federal Government?

Subsidiary:

1. What steps or procedures are appropriate in developing a classification scheme for Government procured services?

2. What are some of the distinguishable characteristics of the services procured by the Federal Government?

3. Which characteristics of Government procured services are the most important for classification purposes?

4. What should be the decision criteria for classifying Government procured services?

5. What are the various homogenous categories of services procured by the Government?

6. In what areas of Government procurement will this classification scheme be most useful?

7. What would a taxonomical structure for classifying Government procured services consist of?

(Wenger, 1990, p. 4)

D. RESEARCH METHODOLOGY

The correlation of the classification scheme for services with Wenger's classification of goods, their potential partnership, and their proximity in time (one year) and location (the same) placed the researcher in a quandary: whether or not to use Wenger's methodology. Since, by its very nature, the objectives of this research correspond to those of the classification of goods, the theory and methodology used in that endeavor provide an appropriate and convenient procedural model for construction of a classification scheme for services. While it would be careless to follow Wenger's methodology without examining it closely, it would, in the researcher's opinion, also be pointless and egotistical to deviate from it for the sake of "originality". In the interests of efficiency and effectiveness, the researcher therefore often uses the research process and structure employed by Wenger in the classification of goods effort, deviating when it is necessitated by the peculiarities of classifying services or when refinements in techniques such as clustering are both appropriate and feasible.

Wenger's research was primarily qualitative and followed a six-step process: 1) a comprehensive literature review; 2) determination of the characteristics to use in the

classification effort; 3) development of the procedures to allow for comparison between products and their characteristics; 4) testing of the procedures by use in actual data collection; 5) cluster analysis of the data, and; 6) determination of a proposed taxonomical model (Wenger, 1990, p. 5).

A comprehensive review of available literature on the subjects of taxonomies, typologies, and service classification schemes was conducted. As examined in Chapter II, classification schemata have been used since ancient times for tangible items such as plant and animal forms. Social taxonomies are more applicable to a service than to tangible items, however, since it is a "...deed, act or performance" (Lovelock, 1983, p. 10), and procuring a service is essentially the act of renting, and thereby modifying or controlling, human behavior for a period of time (Interview-Mulhern, 1991). The major reference publications, therefore, either concern classification science itself, classification of human tasks, or classifications from the perspective of business such as schemata for organization and marketing of goods and services. Section II.G. details how Government classifications provided by the Commercial Activities program, Standard Industrial Classification system, and the Federal

Acquisition Regulation were examined and rejected as classification models for this research.

In order to determine the characteristics to use in the classification effort, the researcher derived several lists of candidate characteristics, predominantly from literature. A "filtering" process was then applied to develop a listing of candidate characteristics. This list, along with explanatory information, was submitted to experts with a broad range of experience in Government contracting. Detailed discussions of the candidate characteristics, as well as the context of research questions, provided feedback to produce a refined listing of preliminary characteristics. These characteristics were defined and differing degrees of application to services were quantified via the formulation of ordinal scales for each characteristic. Next, a matrix and accompanying instructions were designed to allow for a comparison of services with the preliminary characteristics. Finally, a list of 20 heterogeneous sample services was selected and used in a pre-test of the matrix, instructions, definitions, and scales that resulted in refinement of the model. This process is discussed in Chapter III.

The improved matrix, instructions, characteristic definitions, and scales were then used to collect data relative to the sample services by submitting this data

collection package to a population of 300 procurement professionals. Cluster analysis techniques were then used to analyze 85 responses in which the degree of presence of all of the characteristics was assigned a score for all of the sample services. These results were compared with a categorization that had been prepared beforehand in order to ensure that the results were valid on a strategic basis. Five categories of services were produced. Chapters IV and V recount the data collection and analysis process in detail.

The model was then simplified using cluster analysis techniques, measures of the variability of the data collection scores, and data collection feedback on the relative strategic importance of the various characteristics. Categories were labelled and ranges and boundaries were assigned in order to produce a functional taxonomical scheme that may be used in future classification efforts. Both areas are described in Chapter V.

E. SCOPE, LIMITATIONS, AND ASSUMPTIONS

The scope of this study is the development of a scheme for classifying the services procured by the Federal Government from a strategic perspective. In addition to the taxonomy of Government procured goods, classifications from commercial market research were utilized as a starting point for research

since this area is "...replete with classificational schemata" for different kinds of goods, firms, pricing policies, and marketing strategies (Hunt, 1983, p. 348). While this field did not furnish any specific models that directly apply to Government procured services, suitable theoretical and methodological information was available to provide a foundation on which to develop a model.

In common with the taxonomy of Government procured goods, the following assumptions apply:

1. Characteristics of Government procured services exist that lend themselves to ordinal scaling.
2. All Government procured services can be classified.
3. A model may be developed to allow for repetitive classification efforts.

The following limitations apply:

1. Due to time constraints, this thesis effort does not classify all Government procured services.
2. While the model may serve to highlight differences between goods and services as general categories, it will not be appropriate for classifying goods procured by the Government since their essential characteristics and categories are quite different.
3. Because of the diversity of Government procured services and the expertise necessary to classify them, the results of this research should be considered as an introductory services classification model.

(Wenger, 1990, p. 7)

F. REVIEW OF LITERATURE

Organizational Systematics by Bill McKelvey was one of the principal sources for this effort. It provided detailed reviews, not only of organizational classifications, but of fundamental taxonomical theory as well. Taxonomies of Human Performance: The Description of Human Tasks, by Edwin A. Fleishman and Marylin K. Quaintance, also contributed a theoretical framework for the development of classification schemes, and its concentration on applications to social science are also specifically germane to service performance. Marketing Theory: The Philosophy of Marketing Science, by Shelby D. Hunt furnished a complementary classification framework in its description of the development of business classification schemata.

Direct application to classification of services, however, was limited to journal articles. Christopher H. Lovelock's "Classifying Services to Gain Strategic Marketing Insights" provided a valuable summary of commercial classifications of services. Lovelock also applied his synthesis of categories to the firm/client interface, an approach that provided a pertinent approach to Government interests (as a service client).

Two references were critical to the application of cluster analysis techniques to the research effort. H. Charles

Romesburg's Cluster Analysis for Researchers provided step-by-step examples for the application of cluster analysis techniques to classification schemes. The SAS User's Guide: Statistics, Version 5 Edition provided useful explanations concerning the actual performance of a number of different cluster techniques and data processing options.

G. ORGANIZATION OF STUDY

The organization of this study was focused on its primary purpose, to develop a scheme that may classify Government procured services in a useful, strategic structure. Toward that end, this introductory chapter has discussed specific research objectives, questions, methodology, scope, assumptions, limitations, and literature.

Chapter II, *Taxonomy Background*, provides definitions of terms and discusses classification purposes, schemes, and principles, as well as the Government's need for a strategic services classification scheme.

Chapter III, *Development of a Taxonomical Model*, discusses the main objective and conceptual basis of the proposed taxonomy, and the determination of its characteristics. The selection of a taxonomical approach, including a services versus characteristics matrix, is detailed. The characteristics are defined and scaled, and instructions are

formulated to orchestrate use of these elements with the use of the matrix into a complete data collection package. The chapter closes with a pre-test of the classification scheme and consequent revisions.

Chapter IV, *Data Collection and Preliminary Analysis*, discusses how the data collection package was used to collect data on the relationship between selected characteristics and 20 sample services. The analysis of these data via cluster analysis techniques is detailed, and results are compared with an "a priori" categorization of the sample services. The chapter concludes with a discussion on the validation of clustering results and the decision to use five categories in which to group the services.

Chapter V, *Simplifying the Taxonomical Model*, delineates the need to simplify the model, as well as measures of characteristic ascertainability and strategic importance that are used to gauge the relative contribution of characteristics to the model. Cluster analysis techniques are also applied to the characteristic removal process, and simplification is achieved by reducing the number of characteristics from 12 to eight. Category boundaries, ranges, and labels are also assigned. The chapter closes with a discussion of the use of the proposed classification scheme.

Chapter VI, *Conclusions and Recommendations*, summarizes conclusions and recommendations that are directly related to the area of the research effort. Recommendations are also provided concerning the pursuit of procurement classification efforts using other methods or areas of application.

II. TAXONOMY BACKGROUND

A. INTRODUCTION

Classifications have originated out of a need to order or group objects or phenomena that "...transcends all disciplinary boundaries." (Carper & Snizek, 1980, p. 65) "Regardless of whether behavior is learned or instinctive, organisms must be able to perceive similarities in stimuli for survival." (Sokal, 1974, p. 1115)

While the *process* of classification may be dated back to mankind's figurative infancy, the *science* of classification originates with the ancient Greeks. Aristotle applied classification to the study of biology (Margulis & Schwarz, 1982, p. 4), and the first modern type of clas ification, Linnaean taxonomy, is based upon Aristotelian logic (Fleishman & Quaintance, 1984, p. 26). The division of the "natural system" into a universal order is familiar to many people from their literal infancy, when a classification system of various kinds of genus, species, and subspecies was applied in school studies of biology (Sokal, 1974, p. 1115). The science of classification has received special attention in the realm of biology, in part, due to a greater need to elaborate classificatory theory. Biological "...phenomena presented

more of a classificatory problem", and biologists "...had more difficulty coming to an agreement about an acceptable theory than did the physicists, chemists, or mineralogists." (McKelvey, 1982, p. 35) The contribution of biology to classificatory science is such that "...at the heart of any theory of differences is an implicit species concept." (McKelvey, 1982, p. 169)

Classificatory science has been applied to the realms of clinical psychology, medicine, personality, environments, education, and organizational behavior, in order to render a more logical and valid understanding of relationships (Fleishman & Quaintance, 1984, pp. 386-418). Recognition of the need to arrange subject matter in a systematic manner has spread throughout the various social sciences. It has been asserted, for instance, in the case of constituting organizational study as a science, that "...it is relevant, if not crucial, to assess the current state of the discipline in terms of its classificatory schemas." (Carper & Snizek, 1980, p. 65)

Since contracting science falls "...within the category of a social science", it is relevant to apply classifications from other social science studies to classification of Government procured services (Park, 1989, p. 59). In particular, extant classifications of services by industry are

reviewed in section III.C., below. In the pursuit of this classificatory scheme, however, it is important to note that:

> Classificational schemata, no matter how elaborate or complex, are not by themselves theoretical, although most theoretical constructions will contain classificational schemata as components. (Hunt, 1983, p. 348)

Classification schemes simply serve to *organize* phenomena, and the classification scheme that results from this effort is a tool that must be exploited in the future by the *systematic investigation* and *development* of theories in the realm of contracting science.

B. DEFINITION OF TERMS

The following definitions are provided in order to preclude any possible confusion.

Classification is defined as "...the ordering or arrangement of objects into groups or sets on the basis of their relationships." Such relationships can be based on observable or inferred properties. (Sokal, 1974, p. 1116)

The common employment of the term "classification" as a noun will be avoided, since the result of classification, under that usage, could be classification. The terms classificatory system and classificatory scheme are thus defined as synonyms for "the end result of the classification process." (Sokal, 1974, p. 1116)

Taxon (plural: taxa) is defined as "a set of objects of any rank recognized as a group in a classificatory system." (Sokal, 1974, p. 1116)

The terms taxonomy and typology are defined as synonyms that encompass the process and the end product of the set of taxa, when applied in a theoretical study of systematic classifications. This usage includes the bases, principles, procedures, and rules of the classification. (Fleishman & Quaintance, 1984, p. 22)

Defining the term "service" is rather problematic, since this research effort deals specifically with classifying services along a spectrum that recognizes "mixed", as opposed to "pure", services. The definition of a service, for the purposes of this work, is "...a deed, act or performance whose results are mainly intangible and very perishable." (Lovelock, 1983, p. 10) Within the context of this effort, the definition of a good is "a tangible item purchased by the Government to satisfy a need or requirement." (Wenger, 1990, p. 12)

C. WHY WE CLASSIFY

In general, the paramount purpose of a classification scheme is

> to describe the structure and relationship of the constituent objects to each other and to similar objects, and to simplify these relationship's in such a way that general statements can be made about classes of objects.
> (Sokal, 1974, p. 1116)

Four subordinate objectives are: 1) economy of memory; 2) ease of manipulation; 3) ease of retrieval of information, and; 4) description of the structure and relationship of constituent objects. (Sokal, 1974, p. 1116)

1. Economy of memory is achieved by grouping many individual objects into a taxon. The description or definition of the taxon thereby subsumes the individual descriptions of the objects contained within it.

2. Ease of manipulation is produced since the objects are arranged in systems "...in which the several taxa can be easily named and related to each other." If the relationships are very complex, however, labeling or handling of the taxa will be quite difficult. (Sokal, 1974, p. 1116)

3. Ease of retrieval of information is therefore a consideration for classification schemes, since there may be a tradeoff between the descriptive accuracy of the scheme and the ease of accessing the appropriate taxa.

The paramount purpose of classification "...is to describe the structure and relationship of objects to each other and to similar objects." (Sokal p. 1116) Classification thereby equips us to surmise, test and revise hypotheses, policies, and decisions.

D. TWO GENERAL TYPES OF CLASSIFICATION SCHEMES

There are two general methods for generating classification schemes, *logical partitioning* and *grouping*. In logical partitioning, sometimes called "deductive classification", "a priori classification", or "classification from above", a researcher produces a classification scheme *prior to* analyzing any specific set of data, and imposes the scheme on the data. Logical partitioning starts with the specification of the objects or phenomena to be classified, and the properties or characteristics on which the categorizing is conducted. Finally, labels are given to the various categories, or taxa, that emerge from applying the properties or characteristics to the phenomena. (Hunt, 1983, pp. 349-353)

Grouping, also known as "inductive classification", "ex post classification", "classification from below", "numerical taxonomy", or "quantitative classification" generates the classification scheme *after* the analysis of data. Grouping

also begins with a specification of objects or phenomena and respective properties and characteristics, however, the researcher deduces the classification scheme from the data analysis. Basic models include factor analysis, multiple discriminant analysis, multidimensional scaling, and cluster analysis. Numerous computer programs support these models and their methodology. Grouping, in comparison to logical partitioning, requires "...substantially less a priori knowledge" of the properties or characteristics that are likely to be valuable for classifying objects or phenomena (Hunt, 1983, p. 355).

This research effort uses logical partitioning to develop a benchmark "a priori" model. The grouping procedure of cluster analysis is used to develop the proposed classification scheme. Discussion of the general function and advantages of cluster analysis, as well as specific methodologies, is provided in section IV.D.2..

E. CLASSIFICATION PRINCIPLES

How may one differentiate a good classification scheme from a poor one? Regardless of whether logical partitioning or grouping procedures are used, the following questions

embody principles that may be used to evaluate a classification scheme:

1. Does the classification scheme adequately specify the phenomenon to be classified?
2. Does the classification scheme adequately specify the properties or characteristics that will be used in classifying?
3. Does the classification scheme have categories that are mutually exclusive?
4. Does the classification scheme have categories that are collectively exhaustive?
5. Is the classification scheme useful?

(Hunt, 1983, p. 355)

If the answer to all of these questions is "yes", the classification scheme is fundamentally sound. Further explanation of these principles is provided below.

While the *adequate specification of phenomenon* required by the first principle may seem obvious, the phenomenon must be carefully appraised. For example, in classifying a product by its commercial life cycle, does the schema refer to an industry's product or to an individual company's product? The phenomena must be adequately specified.

The *adequate specification of characteristics* is similar to that of phenomenon, since they must be applied consistently to the phenomenon. However, to be appropriate for classificatory purposes, they must also differentiate the

objects or phenomena to be classified, be naturally associated with the objects or phenomena, be relevant to the end-use goal, ascertainable to intended users, and unchanged as long as the end-use goal is unchanged. (Sobczak, 1978, p. 9)

The principle of *mutual exclusivity* refers to a particular situation, where "...if an item fits one category or class, it will not fit any other class." (Hunt, 1983, p. 359) Hence, an item may not be classified into two different categories at a particular level of classification. For example, if the third level of a hierarchical classification split all automobiles into categories of those with two doors or those with four doors, an auto should not be able to be classified in both categories. Many classifications, however, do not meet the mutually exclusive criterion. In marketing, for example, "consumer goods", which are consumed by the ultimate consumer, are often differentiated from "industrial goods", which are used to produce other goods or services. Yet, "...relatively few goods are exclusively industrial goods." (Hunt, 1983, p. 359) While lack of exclusivity degrades the precision of a classification scheme, it is necessary to balance this attribute with the ability to achieve the designated end-use goal.

In order for a classification system to be *collectively exhaustive*, every item classified should belong to a category.

Classification systems may use the simple expedient of incorporating the catch-all category "Other" to conform with this principle. Using the automobile door classification described above, if an automobile had three doors, and this trait was relatively rare, it may be best to classify this auto in the "Other" category. Common classification in this category, however, could nullify the ability of the classification to aid memory, manipulation, and retrieval of information, and its use should be monitored carefully.

The ultimate measure of a classification scheme is its *usefulness*. As noted in the discussion on mutual exclusivity, this criteria is a "first among equals", and attainment of other criteria must not sacrifice the utility of the scheme. The utility of a classification scheme, however, is more difficult to gauge at the outset than conformance with the principles noted above. Early identification of potential uses, users, and benefits will increase the probability that a scheme will be useful.

F. CLASSIFICATION SCHEMES WITHIN THE GOVERNMENT

There are only two listings circulated throughout the bulk of the Federal Government which may be considered as classification schemes for services.

The most detailed classification is provided by the Office of Management and Budget (OMB) in order to monitor conformance with the Commercial Activities program. Table 2-1 is provided as a listing of broad categories established by OMB Circular A-76, which, in conjunction with subsidiary lists, are furnished as "examples" to aid in identifying commercial activities. Commercial activities are services that may be procured from the private sector. This list was issued with the proviso that it should not be considered as collectively exhaustive. (OMB A-76, 1983, pp. 7-10)

The agencies of the Federal Government must report their compliance with the Commercial Activities program to OMB. The lower half of Table 2-1 provides a listing of broad functional areas which the Department of Defense (DoD) uses to report to OMB and Congress. An asterisk (*) highlights duplicate nomenclature to that used in OMB Circular A-76, and it should be noted that most of the nomenclatures differ. This difference in categorization at even the first level of a hierarchy used for the same program demonstrates the limits of a functional approach to classification. While DoD monitored 360 total categories of services, their nomenclatures and interpretation are subject to the whims of fashion, changing technology, and usage by the people administering the program

TABLE 2-1

MAJOR CATEGORIES IN USE FOR CLASSIFYING & MONITORING GOVERNMENT PROCURED SERVICES

Sources: OMB Circular A-76, 1983, pp. 7-10 and DoD 4100.33-INV, 1988.

Office of Management and Budget (OMB)

These categories and subsidiary lists (excluded) are furnished by OMB Circular A-76 as "examples" to aid in identifying commercial activities, with the proviso that the list should not be considered exhaustive.

Audiovisual Products and Services
Automatic Data Processing
Food Services
Health Services
Industrial Shops and Services
Maintenance, Overhaul, Repair, and Testing
Management Support Activities
Manufacturing, Fabrication, Processing, Testing, & Packaging
Office and Administrative Services
Other Services
Printing and Reproduction
Real Property (design, construction, landscaping, dredging)
Security
Special Studies and Analyses
Systems Engineering, Installation, Operation, Maintenance, and Testing
Transportation

Department of Defense (DoD)

These categories are used by DoD for a summary of total man-years by "functional area" to report the status of commercial activities to OMB and Congress. An asterisk (*) highlights duplicate nomenclature to that used in OMB Circular A-76, above.

Social Services
Health Services *
Intermediate Maintenance
Depot Maintenance
Base Maintenance/Multifunction Contracts
Research, Development, Test, Evaluation
Installation Services
Other Nonmanufacturing Operations
Education and Training
Automatic Data Processing *
Product Manufactured/Fabricated In-House
Real Property Maintenance

(DoD 4100.33-INV, 1988). The functional approach of this listing, while comprehensive, does not adequately describe the characteristics used in classifying or utilize mutually exclusive categories. The utility of application of this classification scheme by the Federal Government has also been quite limited.

The Federal Acquisition Regulation (FAR) provides another listing that may be viewed as a classification scheme for services. The structure of the FAR itself, while not intended as a listing of the attributes of services, may be used as a measure of the division of services currently applied by Federal statute and regulation. Table 2-2 provides a listing of FAR categories that are intended to group services by their regulatory structure. While the FAR does not have jurisdiction over all aspects of Government contracting, it refers to the other authoritative regulations, as applicable, in the discussions of each category. These categories have been organized by experts in the field of Government contracting and are updated to reflect statutory and regulatory changes, as well as the concerns of Government buyers and other users. This categorization therefore reflects the current division of Government regulation. (FAR, 1990, Parts 31-39)

TABLE 2-2

MAJOR CATEGORIES OF SERVICES CITED IN THE FEDERAL ACQUISITION REGULATION

Source: FAR, 1990, Parts 31-39.

The FAR uses these categories and sub-categories for organizational purposes. While the FAR does not have jurisdiction over all aspects of Government contracting, it refers to the other authoritative regulations, as applicable, in the discussions of each category. These categories have been organized by experts in the field of Government contracting and are updated to reflect statutory and regulatory changes, as well as the concerns of Government buyers and other users. This categorization therefore reflects the current division of Government regulation.

Category	Defined in FAR Part/para
Research and Development Contracting	
Applied Research	35.001
Basic Research	35.001
Development	35.001
Federally Funded Research and Development Centers	
Independent Research and Development	31.205-18
R&D for Major Weapons Acquisition	34
Construction and Architect-Engineer Contracts	
Architect-Engineer Services	36.6
Service Contracting	
General	37.101
Advisory and Assistance Services	37.2
Dismantling, Demolition, or Removal of Improvements	37.3
Nonpersonal health care services	37.4
Acquisition of Information Resources	39
Transportation and related services	47

This classification, while comprehensive, does not adequately describe the characteristics used in classifying or

utilize mutually exclusive categories. It reflects the hodgepodge approach of statute and regulation, and its potential utility is also quite limited.

The Federal Government produces listings for the classification of service occupations, goods, and manufacturers. They are, respectively, the Service Contract Act Directory of Occupations (U.S. Department of Labor, 1986),the Federal Supply Classification (DoD, 1989), and the Standard Industrial Classification (SIC) Manual (OMB, 1987). The Service Contract Act Directory of Occupations, as well as other Department of Labor listings, are functional in their approach and share the same shortcomings as the functional classification scheme used by the OMB Commercial Activities listing. The SIC Manual classifies and defines activities by industry categories, and is also based on functional categorization and titles (FAR, 1990, para. 19.102). While the SIC is intended to cover the entire field of economic activities, none of these listings incorporate all of the strategic differences between the various services procured by the Federal Government into a classification scheme.

G. THE NEED FOR A STRATEGIC CLASSIFICATION SCHEME

While each of the classifications described above serve a purpose, they do little to reveal the best strategic approach

to procuring services for the Government. In the researcher's opinion, it would be more useful to focus on the various tangible and intangible characteristics of services that have a substantive impact on the acquisition process. A classification on this basis would have a more lasting validity, and could be tested and modified in a systematic manner. The haphazard nature of the functional approach must be overcome in order to provide a systematic body of knowledge that will allow the study of contracting to evolve into a science.

The Federal Government is also extremely dependent on private sector production of services. Approximately $87 billion was spent on service contracting in 1990, a total that increased more than 30 percent in real terms (Welsh, 1991, pp. 14 & 47). Increased spending on maintenance, repair, and upgrades may also be expected during the tight fiscal environment expected in the next decade. DoD, for example, may be expected to use service life extension programs in lieu of purchasing new weapons systems. The Administrator, Office of Federal Procurement Policy has stated that the Government must "...do a better job up front of defining what it wants and of structuring the approach to focus on results." (Welsh, 1990, p. 47) A valid classification scheme would provide Government managers and policy makers with a tool that

accounts for all of the strategic aspects of the procurement of a particular service. Several specific applications that demonstrate the potential value of a classification of Government procured services are listed in section I.C., above.

H. SUMMARY

This chapter provided a general introduction to the science of classification, and discussed the classification schemes currently used by the Government in the procurement of services. By developing a taxonomy which classifies services on the basis of their strategic differences, several benefits are possible in the areas of policy formulation and execution.

The next chapter examines the process of developing a taxonomical model for Government procured services. This process begins by examining the conceptual basis for this classification effort and determining appropriate characteristics. These characteristics are then operationalized into a model, which is pre-tested and revised.

III. DEVELOPMENT OF A TAXONOMICAL MODEL

A. INTRODUCTION

This chapter discusses the process used to develop a taxonomical model for the purpose of classifying Government procured services. Since an ancillary objective of this research is to provide a complementary classification to that which has been completed on Government purchased *goods*, its development process will parallel that used by Wenger for the classification of goods (Wenger, 1990, p. 24). Deviations from that method will therefore be either refinements of the process or necessitated by the peculiarities of classifying services.

Steps delineated in Taxonomies of Human Performance by Fleishman and Quaintance were used as a guideline for development of the model. Since this work dealt chiefly with *task* classification schemes, the necessary procedures were even more appropriate for the classification of services than they were for the classification of goods. The steps identified were:

1. Determining the main objective for the classification effort.
2. Identifying the conceptual basis for the classification.

3. Deciding on the descriptors or characteristics.

4. Operationalizing the scheme.
(Fleishman & Quaintance, 1984, pp. 64-65)

The remaining sections of this chapter will address each of these areas, as well as the methodology of pre-testing the model, pre-test results, and consequent changes. It is important to note that the procedure described in Taxonomies of Human Performance uses an a priori approach: it imposes preconceptions on the classification effort. It also requires that a researcher exercise judgment at every stage of the effort. The researcher has attempted to minimize the imposition of his own preconceptions by maximizing the use of relevant literature and the opinions of experts in the field of Government contracting. Stages where the researcher relied on his own judgment are labeled for the reader by citing the work as resulting from the "researcher's analysis". The thesis also describes, at a minimum, the judgmental issue(s) at-hand and the subsequent decision(s) made by the researcher.

B. MAIN OBJECTIVE AND CONCEPTUAL BASIS

The main objective of the classification effort, as noted in section I.B. above, is to classify services in a way that produces the most strategic insight to the science and practice of Government contracting. The conceptual basis of

this effort is linked to this objective, since the scheme should be employed to classify services procured by the Government in a way that will provide the most information for the formulation of contracting policies and methods.

The conceptual basis for the classification effort is therefore to focus on classifying in a way that offers the most strategic insight. Since current Government schemes do not classify on a strategic basis, this model development takes a fresh look at services by gathering a comprehensive listing of candidate characteristics. A sound methodology, in consonance with the conceptual basis and classification principles (as described in section II.E., above), will cull these characteristics to create a useful classification scheme.

C. DETERMINATION OF THE CHARACTERISTICS

The determination of characteristics is probably the most problematic stage of this research methodology (Wenger, 1990, p. 25). The characteristics that are used, as well as their application, largely determine the relationships that group the objects in the classification. The conceptual basis of the scheme requires the selection of characteristics of Government procured services that have the greatest influence on the acquisition process.

A five-step procedure used by Wenger was evaluated in terms of applicability to the classification of services. While many of the characteristics selected were inappropriate for application to services, the strategy outlined by Fleishman and Quaintance was applicable to the selection of characteristics of human performance, including the service act (Fleishman & Quaintance, 1984, p. 65).

1. Preliminary Listing of Characteristics

The first step was the formulation of a preliminary listing of characteristics. The objective was to formulate a listing of adequate, "candidate" characteristics that would foster creative thinking among a panel of experts. Then, through the interview process with the experts, the candidate characteristics could be modified as necessary to arrive at a group to use in the classification effort.

It is important to note at this point that a search for "the list" of characteristics is, in the words of one author, "...a pipe dream." (McKelvey, 1982, p. 353) Nothing can model a good or other tangible item perfectly, except for the good itself. Since a service is intangible and a product of human behavior, even an "optimal" list will fall well short of perfection.

The researcher first consolidated candidate characteristics from the literature review (see section I.F.,

above,) and an exploratory interview. They are provided in Tables 3-1 through 3-5, below.

Table 3-1 provides Wenger's listing of preliminary characteristics. This list provided an appropriate consolidation of many of the tangible characteristics of services. Since most services involve the use of materiel, that component of Government procured services was represented by many of the preliminary characteristics of goods. Use of Wenger's intermediate and final characteristics, however, was rejected as introducing a premature aggregation, and goods-based bias, into the process.

Tables 3-2 through 3-5 include detailed explanations of the source of their candidate characteristics. The characteristics suggested by Dr. Mulhern in Table 3-2 were solicited in an exploratory interview in order to provide a listing of candidate characteristics from the perspective of the procurer, the Government. The characteristics in Tables 3-3, 3-4, and 3-5, while varied and engaging, were derived from publications concerned with the behavior and interests of industry. The Mulhern list, combined with Wenger's, helped to offset the producer bias of those characteristics.

Since the domain of characteristics provided by this search was quite sizable, it was necessary to reduce the

TABLE 3-1

CANDIDATE CHARACTERISTICS:
PRELIMINARY CHARACTERISTICS USED BY WENGER

Source: Wenger, 1990, p. 27. Italicized term "*service*" substituted for term "good" as appropriate. Term "maintenance" substituted for term "service" in characteristic number 8 to clarify meaning.

Characteristic		Reject Code/ Accept Code*
1.	Unit value.	5
2.	Significance of each individual purchase to the Government.	4
3.	Time and effort spent purchasing by the buyer.	G
4.	Rate of technological change.	K
5.	Technical complexity.	H
6.	Need for maintenance (before, during, or after sale).	5
7.	Frequency of purchase.	D
8.	Rapidity of consumption.	D
9.	Extent of usage (number and variety of users and variety of ways in which the *service* provides utility).	I,J
10.	Amount of price negotiation.	L
11.	Alternate sources availability.	B
12.	Degree of contractor financing required.	P
13.	Amount of product homogeneity.	F
14.	Factors considered by the buyer (price, quality, availability, and technology).	5
15.	What determines price.	4
16.	Amount of choice available to the buyer.	B
17.	Stability of requirements.	C
18.	Amount of short-range versus long-range planning involved.	N
19.	Usage - planned and useful consumption or acquired as "insurance" (i.e., major weapons systems).	3
20.	Extent to which *services* are customized.	F
21.	Extent to which buyer exercises judgment in meeting needs of requiring activity.	M
22.	What is the nature of the demand for the good relative to supply.	B

* see discussion in section III.C.1.

TABLE 3-2

CANDIDATE CHARACTERISTICS:
CHARACTERISTICS SUGGESTED BY MULHERN

Source: Interview-Mulhern, 1991; was an exploratory "brainstorming" session with the researcher to explore possible characteristics and use appropriate nomenclature.

Characteristic	Reject Code/ Accept Code*
1. Ease of formulating Performance Standards	E
2. Differentiation of skills required for contractor personnel	X
3. Professional certification/educational level required for contractor personnel	H
4. Security clearance required of contractor personnel	H,DD
5. Number of users (people)	J
6. Duration of contract	D
7. Geographic limitation of service	V
8. Necessity for information systems	H,S
9. Degree tasks change/are unpredictable over time	K,O
10. Size of firm desired	4

* see discussion in section III.C.1.

number of candidate characteristics to a number that would be manageable for review with a panel of experts. Throughout this effort the researcher had to be alert to the fact that:

> ...construction of classificatory systems, like other data simplifications, involves throwing away information, and it seems advisable to throw away information in as gradual and controlled a manner as possible.
> (Jardine & Sibson, 1971, p. xiii)

The researcher therefore used a filtering model to provide an objective basis to reduce the number of candidate

TABLE 3-3

CANDIDATE CHARACTERISTICS: PROPOSED TO GAIN STRATEGIC MARKETING INSIGHTS CONCERNING PROVIDER-CUSTOMER INTERFACE

Source: Lovelock, 1983, p. 12 - p. 18. Dimensions numbered one through four originally presented as 2 x 2 category matrices, number five originally presented as a 3 x 2 matrix. Italics indicate scale for differentiation with characteristic.

Dimension/Characteristics	Reject Code/ Accept Code*
1. **Understanding the Nature of the Service Act**	
a. What is the Nature of the Service Act? *Tangible versus Intangible actions*	Q
b. Who or What is the Direct Recipient of the Service? *People versus Things*	I,J
2. **Relationships with Customers**	
a. Nature of Service Delivery. *Continuous delivery versus Discrete transactions*	U
b. Type of Relationship between the Service Organization and Its Customers. *"Membership" versus No Formal relationship*	3
3. **Customization and Judgement in Service Delivery**	
a. Extent to Which Customer Contact Personnel Exercise Judgment in Meeting Individual Customer Needs. *High versus Low*	O
b. Extent to Which Service Characteristics Are Customized. *High versus Low*	F
4. **What is the Nature of Demand for the Service Relative to Supply?**	
a. Extent to Which Supply is Constrained. *Peak Demand Can Usually Be Met without a Major Delay versus Peak Demand Regularly Exceeds Capacity*	B
b. Extent of Demand Fluctuations over Time. *Wide versus Narrow*	C
5. **Method of Service Delivery**	
a. Nature of Interaction between Customer and Service Organization. *Customer Goes to Service Organization versus Service Organization Comes to Customer versus transact at Arm's Length*	W
b. Availability of Service Outlets. *Single versus Multiple site*	V

* see discussion in section III.C.1.

TABLE 3-4

CANDIDATE CHARACTERISTICS:
A TYPOLOGY OF SERVICE ORGANIZATIONS

Source: Mills and Margulies, 1980, p. 262. "Dimensions" and their subset characteristics were used to classify service organizations in a resultant taxonomy as being either *Maintenance*, *Task*, or *Personal Interactive*.

Dimension/Characteristics	Scale Provided	Reject Code/ Accept Code*
Information		
Information quantity	*Low/Moderate/High*	Y
Information quality	*Low/Moderate/High*	Z
Confidentiality	*Low/Moderate/High*	DD
Decision		
Employee decisions	*Simple/Complex*	O
Importance	*Low/Moderate/High*	O
Feedback (client to employee)	*Slow/Immediate*	EE
Time		
Interface duration	*Brief/Moderate/High*	BB
Total time in direct contact	*Moderate/High*	CC
Problem Awareness		
Client knowledge about problem	*Low/Moderate/High*	3
Client ability to evaluate services	*Low/Moderate/High*	E
Client expectations vs. service capabilities	*Low/Moderate/High*	3
Transferability		
Substitutability of employee	*Low/Moderate/High*	H,X
Power		
Perceived power of employee with respect to client	*Low/Moderate/High*	H,X
Employee status to client	*Low/High*	H,X
Employee authority with client	*Low/High*	H,X
Attachment		
Employee identification with client	*Low/Moderate/High*	D
Conflict potential	*Low/Moderate/High*	FF

* see discussion in section III.C.1.

TABLE 3-5

CANDIDATE CHARACTERISTICS: SELECTIVE SUMMARY OF PREVIOUSLY PROPOSED SCHEMES FOR CLASSIFYING SERVICES

Source: Provided below as cited by Lovelock, 1983, p. 11. Dimensions are excluded for the sake of economy if used by a previous source or in Lovelock's 1983 scheme (Table 3-3).

Source	Dimensions	Reject Code/ Accept Code*
Judd (1964)	(1) Rented goods services (right to own and use a good for a defined time period)	GG
	(2) Owned goods services (custom creation, repair or improvement of goods owned by the customer)	GG
	(3) Nongoods services (personal experiences or "experiential possession")	Q
Rathmell (1974)	(1) Type of seller	4
	(2) Type of buyer	3
	(3) Buying motives	4
	(4) Buying practice	G,M
	(5) Degree of regulation	A
Hill (1977)	(1) Permanent vs. temporary effects of the service	R
	(2) Reversibility vs. nonreversibility of these effects	R
	(3) Physical vs. mental effects	Q
	(4) Individual vs. collective services	HH
Thomas (1978)	(1) Primarily equipment based	S,T
	(a) automated (e.g., car wash)	
	(b) monitored by unskilled operators (e.g., movie theater)	
	(c) operated by skilled personnel (e.g., airline)	
	(2) Primarily people-based	H,T
	(a) unskilled labor (e.g., lawn care)	
	(b) skilled labor (e.g., repair work)	
	(c) professional staff (e.g., lawyers, dentists)	

* see discussion in section III.C.1.

characteristics. Characteristics were rejected if they did not possess desirable traits. These traits include:

Reject Code/Required Trait

1. Differentiation - the characteristic should have the ability to segregate services into at least two different classes.

2. Concomitance - the characteristics should accompany or be naturally associated with services.

3. Relevance - each characteristic should be valid and support the end-use goal(s).

4. Ascertainability - each characteristic should allow the user of the classification scheme to precisely determine the presence of the characteristic and the degree of that presence.

5. Permanence - the characteristic should be present and definable.

6. Consistency - the application of the characteristic should be the same for various types of services.

(Sobczack, 1978, p. 9)

In order to display the use of these traits in the analysis, the researcher assigned "Reject Codes" to candidate characteristics that were rejected. The numeric designators one through six, used above to list the desirable traits, were annotated to rejected characteristics in Tables 3-1 through 3-5 based on the respective reason for rejection. For example, the first candidate characteristic in Table 3-1, "unit value", was annotated with Reject Code 5 since the researcher did not believe that a unit value could be permanently defined (trait

5) for a service. The number 5 is provided as a Reject Code for that candidate characteristic on the right-hand side of Table 3-1.

After the use of this filtering process, the researcher still needed to reduce the number of candidate characteristics to a quantity that would be manageable for review with a panel of experts. The researcher therefore analysed candidate characteristics to consolidate redundant characteristics into preliminary characteristics that could be expressed in a few words. It has been stated, however, that it is "...hard to conceive of any general procedure for the elimination of redundancy in selections of attributes." (Jardine & Sibson, 1971, p. 28) A characteristic may be excluded as redundant if it "...is a logical property of another", as are hemoglobin and redness of blood in a medical classification (Sokal & Sneath, 1963, p. 66). Similarly, characteristics may be rejected as redundant if they are perfectly correlated, statistically. In many cases, however, "...the dependence of one character upon another is not total but only partial." (Sokal & Sneath, 1963, p. 67) For instance, when a given characteristic **A** depends in part upon another characteristic **B**, the decision of whether to use both should depend on the nature of the factors, other than **A**, that affect **B**. Since the researcher did not have any statistical

evidence concerning the correlation of the strategic significance of any of the characteristics at this stage, the analysis had to depend on the researcher's own experience. The researcher attempted to group characteristics which describe the same phenomena or depend on the same factors.

In order to display the consolidation of redundant candidate characteristics, the researcher assigned "Accept Codes" to all of the candidate characteristics that had not been heretofore rejected. These were designated alphabetically in order to distinguish them from numeric Reject Codes, and were assigned as the researcher reviewed the candidate characteristics for redundancy. An Accept Code indicates the preliminary characteristic, as listed alphabetically in Table 3-6, that the candidate characteristic was adopted into. For example, the first candidate characteristic in Table 3-2, "*Ease of formulating Performance Standards*" was, in the opinion of the researcher, redundant with the "*Client ability to evaluate services*" listed in Table 3-4. They were therefore both assigned to preliminary characteristic "E" in Table 3-6, as being approximately the same as "*Ease of measuring performance*". The researcher thereby condensed the candidate characteristics into the list of 34 preliminary characteristics provided in Table 3-6.

TABLE 3-6

PRELIMINARY CHARACTERISTICS

Source: Researcher's Analysis. Order of characteristics does not coincide with any presumed order of importance, nor is the grouping necessarily logical. Alphabetical designator for characteristics is source for Accept Codes indicated on other tables. *The list is continued on the next page.*

Preliminary Characteristic

A. Degree of regulation
B. Availability of alternate sources
C. Stability of requirements
D. Duration of contract
E. Ease of measuring performance
F. Degree to which service is customized
G. Time and effort spent purchasing by the buyer
H. Professional certification/experience level required of contractor personnel
I. Degree to which things are the direct recipient of the service
J. Degree to which people are the direct recipient of the service
K. Rate of technological change
L. Amount of price negotiation
M. Degree of buyer judgment exercised
N. Amount of short-range versus long-range planning
O. Degree to which contractor personnel exercise judgment
P. Degree of contractor financing required
Q. Degree to which benefit of service is physical versus mental
R. Permanence of effects of the service
S. Cost of material and equipment used in service production
T. Relative cost of labor versus cost of material and equipment
U. Extent delivery is discrete versus continous
V. Geographic extent of service delivery/availability
W. Whether service is delivered, user goes to service site, or delivery is at arms length
X. Differentiation of skills required for contractor personnel

---

TABLE 3-6

PRELIMINARY CHARACTERISTICS
(CONTINUED)

Preliminary Characteristic
Y. Quantity of information exchanged between Government and contractor personnel
Z. Quality of information exchanged between Government and contractor personnel
AA. Total price of contract
BB. Duration of typical contractor/user direct contact
CC. Total time in direct contact
DD. Confidentiality of service task and information
EE. Degree of feedback from Government to contractor
FF. Conflict potential between Government and contractor
GG. Degree to which Government furnishes materials and equipment
HH. Number of contractor personnel employed on contract

2. Expert Panel Selection and Interviews

The second step was to assemble and interview an expert panel, which eventually included fifteen members. Seven members who had assisted Wenger in his classification of goods were selected due to their recent involvement with a taxonomic research effort (Wenger, 1990, pp. 105 & 137). These members were either academics or consultants who, however, had not recently procured Government services themselves, and the researcher sought to balance their expertise with members who were currently employed in Government acquisition. Acquisition officers and civilians from the U.S. Army, Navy, and Air Force, as well as a

professor at the Defense Systems Management College, were added to the panel. The names of panel members, as well as brief descriptions of their qualifications, are provided in Appendix A.

The researcher first contacted experts by telephone to determine whether they were willing to be panel members, and to provide a brief description of the research effort. Panel members were then mailed a background package that consisted of an introductory letter and enclosures describing: (1) the classification scheme's objective, conceptual basis, uses, and principles; (2) questions to determine appropriate characteristics for the scheme, and; (3) attributes each characteristic must possess, and the list of preliminary characteristics. A copy of the contents of this package, except for enclosure (3), is provided in Appendix A. The contents of enclosure (3) of the package are provided in the text of this effort as the Required Traits and Table 3-6 in section 3.C.1., and are excluded in the interest of economy. The questions provided in enclosure (2) of the package were intended to orient panel members to issues involved in the taxonomic effort. The main focus of the interviews was a review of the validity of the various preliminary characteristics.

Nine interviews were conducted by telephone, but the remainder were conducted in-person by the researcher. In-person interviews are denoted in Appendix A by the use of asterisks next to the description of member qualifications. Of special interest were four group interviews conducted with at least two people in the member's contracting organization. They proved very useful when the groups would become involved in their own mini-discussions, questioning each other's answers and generally providing more descriptive responses. The group at the Sacramento Army Depot, for instance, included the administrator and supervisor for an omnibus base support contract. The group was able to discuss the efficacy of using a taxonomy to split such service contracts into strategically coherent subcategories and subcontracts, and was receptive to such an end-use. The interviews primarily provided additional characteristics for the classification scheme, a consensus of expert opinion on the superior strategic relevance of some preliminary characterisics, as well as a basis for rejecting others as less relevant to the end-use goal.

3. Analysis of Preliminary Characteristics

The third step was the analysis of preliminary characteristics. Several panel members recommended additional preliminary characteristics, along with basic definitions. These characteristics are listed in Table 3-7.

TABLE 3-7

ADDITIONAL PRELIMINARY CHARACTERISTICS RECOMMENDED BY PANEL
Source: Researcher's Analysis

1.	Management complexity of service production
2.	Vulnerability to externalities
3.	Risk to the Government
4.	Small/disadvantaged business development by the Government
5.	Documentation by the Government
6.	Oversight by the Government

The original preliminary characteristics (Table 3-6), as well as the additional preliminary characteristics (Table 3-7) that had been suggested to-date, were examined by the panel members. The researcher recorded member responses concerning the strategic relevance of characteristics, and their ability to be used as discriminators between perceived service types. It was difficult to precisely gauge the conviction of panel members concerning the relative validity of all of the characteristics, so the researcher categorized the recording into three general types for analytical purposes. Either:

1. The panel member thought the characteristic was a poor discriminator;

2. The response was conditional, (that the characteristic was useful subject to certain conditions), or;

3. The member thought the characteristic was a strategically valid discriminator.

This breakdown proved useful in the analytical effort to determine the panel consensus concerning each preliminary characteristics. It should be noted, however, that this method was highly dependent upon the accuracy of the researcher in recording and categorizing responses of panel members, and served as a simple aid to the researcher's judgment in measuring the intensity of member responses. It is not presented as a thoroughly objective method. The researcher assigned a point scale to each of the three responses: zero to a negative response (response one, above); one to the conditional response (response two, above); and three to the positive response (response 3, above). The researcher then entered the responses of panel members to each preliminary characteristic in a database program, and obtained average responses for each preliminary characteristic. Of these, ten were clearly preferred. These were preliminary characteristics B, C, E, F, H, K, L, T, DD, from Table 3-6, and additional preliminary characteristics from Table 3-7.

Up to this point, the preliminary characteristics had been listed using a few words to provide a descriptive, but non-restrictive, definition that could be discussed with panel members. Summary titles, however, are desirable for the purpose of simplicity of presentation and communication. One-to-three word summary titles for characteristics, as well as

general definitions, were discussed with panel members. The researcher depended mainly, however, on his own discernment to create titles for each characteristic. They are listed in Table 3-8, opposite the alphabetical or numeric character used for their original titles as preliminary characteristics in Tables 3-6 and 3-7. For instance, preliminary characteristic "A" from Table 3-6, "*Degree of regulation*", is listed as preliminary characteristic "A" on Table 3-8 as **Regulation**.

4. Selection of Dimensions Upon Which to Classify

The fourth step used by Wenger was to classify preliminary characteristics by a broad category or "dimension" in order to help analyze characteristic attributes and identify any repetition or overlap. Wenger's dimensions were *characteristics of the goods per se*, *characteristics of the buyer's effort*, and *characteristics of the environment* of the particular procurement, and were selected based on his own perceptions (Wenger, 1990, p. 30). In the opinion of the researcher, these three basic dimensions were valid for services, although the researcher replaced the word "buyer" with "Government" since more than one Government organization may be involved with characteristics of service procurement, such as **Documentation**.

TABLE 3-8

SUMMARY TITLES FOR PRELIMINARY CHARACTERISTICS
Source: Researcher's Analysis. Alphabetical and numeric designators coincide with those used for preliminary characteristics in Tables 3-6 and 3-7.

Preliminary Characteristic	Preliminary Characteristic
A. Regulation	U. Consumption Continuity
B. Competition	V. Geographic Availability
C. Stability	W. Delivery Requirement
D. Duration of Contract	X. Skill Differentiation
E. Measurability	Y. Quantity of Data Exchanged
F. Customization	Z. Quality of Data Exchanged
G. Buyer Attention	AA. Task Price
H. Expertise	BB. Typical User Contact Duration
I. Object Recipient	CC. Total User Contact Duration
J. Personal Recipient	DD. Confidentiality
K. Complexity	EE. Feedback
L. Negotiation	FF. Disagreement Potential
M. Buyer Judgment	GG. Government Material Employed
N. Planning	HH. Total Labor Hours
O. Judgment (contractor)	1. Management Complexity
P. Financing	2. Vulnerability to Externalities
Q. Tangibility	3. Risk to the Government
R. Perishability	4. Small/Disadvantaged Business Development
S. Total Material Cost	
T. Labor % of Cost	5. Documentation
	6. Oversight

Characteristics of the goods per se had also been further divided by Wenger into *characteristics inherent to the good* and *external to the good*. *Inherent* characteristics were defined as "...those that could be directly identified to the good and would not depend on outside influence" to determine their presence or absence (Wenger, 1990, p. 31). The researcher believed that, due mainly to the intangible nature

of services, that a characteristic could not be directly identified to a service, per se. Such a mental construct is, at best, awkward. For example, **Expertise** is not *inherent* to a service per se. The researcher believed that a more lucid concept would be to directly identify characteristics as *internal* to production of a service as a deed, act, or performance. The researcher therefore replaced the word "inherent" with the word "internal" in the *per se* dichotomy. Wenger defined *external* characteristics as "...those that remain, to a large extent, related to the good but require some outside influence to recognize if the characteristic is present or not." (Wenger, 1990, p. 31) The researcher regarded that definition as useful, for example, in classifying **Measurability** as a characteristic of a service per se that is, nonetheless, external to the production of the service, since it is a function of an observer. The researcher therefore accepted Wenger's definition of "external" characteristics as the opposing category to "internal" characteristics in the *per se* dichotomy. The only further modification was to replace the word "good(s)" with "service(s)" for each dimension.

The researcher, based on his own judgment, then used the resultant dimensions to group preliminary characteristics. This grouping is displayed in Table 3-9. Of the ten

characteristics that had been preferred by consensus of expert panel members, seven were characteristics of a service per se, one was a characteristic of the Government's effort, and two were characteristics of the environment. The distribution selected by Wenger at this stage consisted of eight characteristics of goods per se, one was a characteristic of the buyer's effort, and three were characteristics of the environment for a total of twelve (Wenger, 1990, p. 33). Since the distribution of dimensions was very similar, the researcher believed that the services classification effort was proceeding on a sound basis to provide a complementary study. As has been noted, however, the researcher wanted to "throw away" information, or characteristics, in as gradual and controlled a manner as possible. Since twelve characteristics had proved manageable in the Wenger study, the researcher reexamined preliminary characteristics which had a mixed preference by the expert panel in order to keep two more characteristics for a pre-test.

In particular, the preliminary characteristics of **Buyer Attention**, **Buyer Judgment**, **Documentation**, and **Oversight** had received mixed reviews. The breakdown of preliminary characteristics provided in Table 3-9, however, highlighted this group as belonging to one dimension of classification.

TABLE 3-9

PRELIMINARY CHARACTERISTICS GROUPED BY DIMENSION
Source: Researcher's Analysis. Characteristics suggested during interviews are *italicized*

CHARACTERISTICS OF THE SERVICE PER SE

INTERNAL TO THE SERVICE	EXTERNAL TO THE SERVICE
Customization	Measurability
Expertise	Personal Recipient
Complexity	Object Recipient
Judgment	Perishability
Financing	Consumption Continuity
Tangibility	Geographic Availability
Total Material Cost	Quantity of Data Exchanged
Labor % of Cost	Quality of Data Exchanged
Delivery Requirement	Task Price
Skill Differentiation	Typical User Contact Duration
Total Labor Hours	Total User Contact Duration
Management Complexity	Confidentiality
Vulnerability to Externalities	Feedback
	Disagreement Potential
	Government Material Employed
	Risk to the Government

CHARACTERISTICS OF THE GOVERNMENT'S EFFORT

Buyer Attention
Negotiation
Buyer Judgment
Planning
Small/Disadvantaged Business Development
Documentation
Oversight

CHARACTERISTICS OF THE ENVIRONMENT

Regulation
Competition
Stability
Duration of Contract

Since they were all characteristics of the Government's effort, the researcher decided that the validity of these characteristics might be further tested by keeping the characteristic **Buyer Attention**. This characteristic seemed to have the greatest degree of overlap with the others, and yet be both strategically significant and ascertainable to a user. **Perishability** and **Total Labor Hours** were the preliminary characteristics that were the next most preferred by the expert panel. Some panel members observed that **Total Labor Hours** was more a function of contract price than the service per se, and the researcher agreed with that assessment. Since the definition of a service used in this effort was that of a deed, act, or performance that is *very perishable*, **Perishability** was selected by the researcher to see if it might be a valuable discriminator between pure and mixed services. The final list of characteristics that were used for a pre-test of the taxonomic model is provided, by dimension, in Table 3-10. It must be stressed that the decisions to maintain **Buyer Attention** and **Perishability** as characteristics were particularly subject to the imposition of judgment by the researcher.

5. **Preliminary Characteristic Definitions**

The fifth step in the determination of characteristics was to define each pre-test characteristic. The researcher

TABLE 3-10

PRE-TEST CHARACTERISTICS GROUPED BY DIMENSION
Source: Researcher's Analysis. Characteristics suggested during interviews are *italicized*

CHARACTERISTICS OF THE SERVICE PER SE	
INTERNAL TO THE SERVICE	EXTERNAL TO THE SERVICE
Customization	Measurability
Expertise	Perishability
Labor % of Cost	Confidentiality
Complexity	*Risk to the Government*
CHARACTERISTICS OF THE GOVERNMENT'S EFFORT	
Buyer Attention	
Negotiation	
CHARACTERISTICS OF THE ENVIRONMENT	
Competition	
Stability	

developed brief definitions for each characteristic, including strategic relevance to a buyer, based on his own experience and interview comments by panel members. They would be used for a pre-test, and some would later be modified according to input received from the pre-test panel. The pre-test definitions are listed below.

1. Customization is the degree to which the production of a service is modified from standard provider practice to conform with the buyer's unique specifications. All services are modified to some degree to reflect circumstances unique to each customer, but they will differ on the magnitude to which their procedures, or the entire service process, are customized exclusively for a buyer. In general, a greater degree of customization will increase the amount of buyer attention, and contract cost, necessary to ensure successful service performance.

2. Expertise is the degree of professional certification, skill, and experience required of the principal service production personnel to produce a service at an acceptable quality level. Higher levels of required expertise will usually increase the difficulty of evaluating service performance, as well as the extent to which the buyer should validate service provider qualifications.

3. Complexity is the degree of technical complexity of equipment and techniques used in the scope of service production. Typically, a high degree of technical complexity will require that the buyer devote substantial attention to evaluating the skill level or equipment required to produce a service, as well as evaluating potential providers for those capabilities.

4. Labor Percentage of Cost is the degree to which total service cost is expended on provider labor (as opposed to material and equipment). Buyer validation of provider qualifications, especially in the realm of financing, should be affected by the proportion of labor and material and equipment required to perform a service.

5. Measurability is the degree of effort necessary to describe and measure acceptable service performance. While performance of some services is obvious and readily measured, others may necessitate extensive description and detailed review by the buyer to determine if service performance satisfies buyer requirements.

6. Confidentiality is the degree to which release of information produced by, or required to produce, a service may be detrimental to either the buyer or service provider. The magnitude of potential damage, whether it be financial, competitive, related to reputation, or to national security, from a release of service information determines the level of service confidentiality. A high grade of confidentiality should necessitate extensive buyer validation of provider qualifications for controlling confidential information.

7. Risk to the Government is the likelihood and magnitude of potential harm to the Government that would result if a service is not completed in accordance with cost, schedule, or performance specifications. Buyer attention should increase throughout the entire acquisition process as the degree of risk to the Government escalates.

8. Government Attention is the degree of time, effort, and judgment that buyer personnel typically dedicate to acquisition of a service. Personnel allocation, work assignments, and other buyer organization plans and policies should vary with the distinctive degree of buyer attention customarily required by different types of services.

9. Negotiation is the degree to which price, schedule, and performance criteria are discussed and adjusted by the buyer and potential service providers during the service acquisition process. More negotiation will generally require a longer and more detailed acquisition effort.

10. Competition is the degree to which multiple, autonomous providers are willing and able to produce a service. Typically, the intensity of competition will influence buyer selection of contract type, as well as the extent to which price is the dominant source-selection factor.

11. Stability is the degree to which the schedule and performance criteria of a service remain the same over a period of time. A more stable service will typically require less attention on the part of the buyer.

12. Perishability is the length of time that the product of service performance is beneficial to, or consumed by, the buyer organization. A service with a relatively high degree of perishability will be consumed almost instantaneously, while the product of other services may provide benefits for many years.

D. OPERATIONALIZING THE SCHEME

Once the characteristics to be used for the pre-test were generated and defined, the next part of model development was to select an appropriate method for using the characteristics to classify services.

1. Matrix Approach Versus Decision Tree Approach

Wenger considered use of a decision tree, but decided upon a matrix, approach to classifying Government purchased goods (Wenger, 1990, p. 37). The researcher reviewed the matrix approach in the exploratory phase of the research effort, and found it to be visually uncomplicated and intuitively appealing. These surface impressions were reasons cited by Wenger, but he also added that a matrix is itself a superior tool for data collection and observing relationships. Such relationships, if quantified, may in turn be analyzed using cluster analysis techniques "...to determine the resultant 'clusters' or categories of objects." (Wenger, 1990, p. 38)

Since the researcher desired, if possible, to produce a complementary classification to Wenger's classification of

goods, a decision tree approach would need to be superior for a classification of services. The decision tree approach produces a hierarchical classification scheme, and hierarchies are the most powerful method of achieving economy of memory in a classification (Sokal & Sneath, 1963, p. 171). A hierarchy that is familiar to many people is used in the field of biological science: Kingdom, Phylum, Subphylum, Class, Order, Family, Genus, and Species are commonly used as tiers in biological classifications (Margulis & Schwarz, 1982, p. 3). At every tier of a hierarchy the taxonomist must decide which category that a subject belongs to. Prior to the theory of evolution, biologists found that, empirically, a hierarchy gave the most satisfactory and "natural" arrangement of the data. Such a system could also be constructed using a few characteristics:

> The art of the practice lay in finding suitable characters, to prevent the classification from creating strange bedfellows, clearly incongruous as judged by their great differences in other characteristics.
> (Sokal & Sneath, 1963, p. 17)

Several characteristics had been generated at this stage in model development, and the researcher attempted to apply them to a decision tree. Some characteristics could be scaled in a numeric manner that would lend themselves to clear-cut decisions concerning which category a service

belonged to. For example, **Competition** and **Perishability** could be scaled with categories encompassing, respectively, different numbers of competitors or different periods of productivity. Such characteristics would eventually have to be tiered in descending order according to their relative strategic significance in order to produce a decision tree. In the opinion of the researcher, however, characteristics such as **Confidentiality** and **Risk**, however, would not lend themselves to clear-cut decisions. To reject characteristics on the basis that they could not contribute to a decision tree approach would be to prematurely discard potentially useful information. The taxonomist, if not careful, may also produce "strange bedfellows" if characteristics were prematurely selected. Use of a matrix approach, conversely, would allow for clustering and thereby preclude the grouping of services into incongruent categories with great differences in characteristics. Use of a matrix approach may, furthermore, provide knowledge necessary for generation of a decision tree by enriching our understanding of the relative strategic importance of various service characteristics.

2. Scaling the Characteristics

In order to provide for clustering, a matrix must allow classifiers to quantitatively judge the presence of service characteristics (Romesburg, 1984, p. 33). The

researcher defined scales for each pre-test characteristic based on their established definitions and his own expertise. A five-point scale was selected since it had proved manageable in the Wenger effort, and was simple to use. In the opinion of the researcher, a ten-point scale is a more common scale in American society, but differentiation between scaling levels would be difficult to define and score. Characteristics such as **Competition** may lend themselves to numeric definition, for example, by scaling a "1" as having only one competitor, and a "10" as having ten competitors. Characteristics such as **Customization**, however, do not lend themselves to such clear-cut scaling. Scorers would, in turn, have greater difficulty deciding whether **Customization** is an "8" or a "9" on a ten point scale than in differentiating between a "3" or a "4" on a five point scale. The researcher therefore selected a five point scale.

In choosing a scale and defining it, there is also a danger of being too prescriptive. To use the **Competition** characteristic example, if the researcher defined a "3" on a five point scale as being three-to-five competitors, scorers would essentially have to follow the researcher's idea of what constitutes a moderate amount of competition. The *researcher* would impose his own preconceptions on an integral part of the effort. Characteristics that lend themselves to numeric

scaling would also seem superior to more conceptual characteristics, since the variance between scores would be reduced due to the imposed ascertainability to the user. While ascertainability is important, it may be tested by calculating variances between scores at a later stage. A neutral, consistent approach to defining scales would allow *scorers* to determine the desirability of using characteristics in the model. The researcher therefore composed scale definitions that were descriptive, but not prescriptive. Common adjectives were selected to represent the degrees of each scale.

The researcher also sought to define the scales so that ascending values would coincide with greater strategic complexity, on a range from simple-to-complex. This range was also used by Wenger (Wenger, 1990, p. 39). A danger of this format was that some scales may appear to be counter-intuitive. For example, a scoring of "5" for the characteristic **Measurability** may intuitively denote to many people that a service is *very measurable*. A very measurable service, however, may typically be *very simple*. In order to produce a scale that makes a "5" typical of a *very complex* service, the scale may have to be counter-intuitive. There

are two methods that may be used to counteract the danger of counter-intuitive characteristics, either:

1. Scale counter-intuitive characteristics according to a presumably intuitive order, and reverse their values (from a "1" to a "5", from a "2" to a "4", etc.) when computing mean values on a range of simple-to-complex, or;

2. Warn scorers in their instructions that scales may appear to be counter-intuitive, and clearly label the scales as such.

While the first option would appear to be simpler, arranging scales in their intuitive order would still be confusing for scorers since a typically complex service that had been on the upper end of the numeric scale would suddenly be on the low end. The scorer would still have to consult the instructions and scales to quell any confusion and render the desired scaling. The second option would directly label the instructions and scales, without the need for reversing the scale values during computation. The researcher therefore deemed the second option to be the preferable course of action. It should be stressed, however, that labeling scales as being counter-intuitive was mainly a precautionary measure on the part of the researcher. Deciding that a scale was counter-intuitive was a presumption on the part of the researcher.

The researcher drafted descriptive scales and labeled those that may appear to be counter-intuitive, based on pre-test definitions and his own experience. Their utility would be checked later during a pre-test, and some would be modified according to input received from the pre-test panel. Pre-test scales are listed below.

1. **Customization:**
   Scale
   1 - No customization
   2 - Customization does not substantively alter service production
   3 - Customization substantively alters a few important elements of service production
   4 - Customization substantively alters the bulk of important elements of service production
   5 - The service is produced exclusively for the Government

2. **Expertise:**
   Scale
   1 - No expertise needed by principal service production personnel
   2 - Expertise needed requires brief or inexpensive training/qualification
   3 - Expertise needed requires lengthy or expensive training/qualification
   4 - Expertise needed requires very lengthy or very expensive training/qualification
   5 - Expertise needed requires extremely lengthy or extremely costly training/qualification

3. **Complexity:**
   Scale
   1 - Technical complexity is rudimentary
   2 - Technical complexity is modest
   3 - Technical complexity is sophisticated
   4 - Technical complexity is advanced
   5 - Technical complexity is on the frontier of human knowledge and capabilities

4. **Labor Percentage of Cost:**
   Scale
   1 - A modest amount of total service cost is expended on labor
   2 - A moderate amount of total service cost is expended on labor
   3 - The bulk of total service cost is expended on labor
   4 - The vast preponderance of total service cost is expended on labor
   5 - Almost all of total service cost is expended on labor

5. **Measurability:**
   Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
   1 - Description and measurement of acceptable service performance is obvious and almost effortless
   2 - Description and measurement of acceptable service performance is uncomplicated
   3 - Description and measurement of acceptable service performance is difficult
   4 - Description and measurement of acceptable service performance is cryptic and laborious
   5 - Description and measurement of acceptable service performance is profoundly perplexing and intricate

6. **Confidentiality:**
   Scale
   1 - Release of service production information is not at all potentially detrimental to the provider or Government
   2 - Release of service production information would potentially cause inconsequential damage to the provider or Government
   3 - Release of service production information would potentially cause notable damage to the provider or Government
   4 - Release of service production information would potentially cause extensive damage to the provider or Government
   5 - Release of service production information would potentially cause enormous damage to the provider or Government

7. **Risk to the Government:**
   Scale
   1 - The likelihood and magnitude of potential harm to the Government due to service performance failure is negligible
   2 - The likelihood and magnitude of potential harm to the Government due to service performance failure is conspicuous
   3 - The likelihood and magnitude of potential harm to the Government due to service performance failure is considerable
   4 - The likelihood and magnitude of potential harm to the Government due to service performance failure is great
   5 - The likelihood and magnitude of potential harm to the Government due to service performance failure is enormous

8. **Buyer Attention:**
   Scale
   1 - Service procurement requires inconsequential time and effort from buyer personnel
   2 - Service procurement requires minor time and effort from buyer personnel
   3 - Service procurement requires moderate time and effort from buyer personnel
   4 - Service procurement requires considerable time and effort from buyer personnel
   5 - Service procurement requires extraordinary time and effort from buyer personnel

9. **Negotiation:**
   Scale
   1 - There is no negotiation between buyer and potential providers during the service procurement process
   2 - Negotiation is insignificant between buyer and potential providers during the service procurement process
   3 - Negotiation is meaningful between buyer and potential providers during the service procurement process
   4 - Negotiation is extensive between buyer and potential providers during the service procurement process
   5 - Negotiation is critical and comprehensive between buyer and potential providers during the service procurement process

10. **Competition:**
Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
1 - Numerous autonomous providers are willing and able to produce the service and are very aggressive in their willingness to do so
2 - It is quite easy to find several providers who are willing and able to produce the service
3 - It is uncomplicated to find a few autonomous providers who are willing and able to produce the service
4 - It is difficult to find a few autonomous providers who are willing and able to produce the service
5 - It is extremely difficult to find a provider willing and able to produce the service

11. **Stability:**
Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
1 - Any alteration to schedule or performance criteria is, at most, trivial for extremely lengthy periods of time
2 - Important schedule or performance criteria seldom undergo significant alteration
3 - Important schedule or performance criteria infrequently undergo significant alteration
4 - Important schedule or performance criteria frequently undergo significant alteration
5 - Important schedule or performance criteria almost constantly undergo significant alteration

12. **Perishability:**
Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
1 - The period of benefit/consumption is abrupt
2 - The period of benefit/consumption is brief
3 - The period of benefit/consumption is moderate
4 - The period of benefit/consumption is lengthy
5 - The period of benefit/consumption is extremely lengthy

3. Preliminary Taxonomical Model

Once the taxonomical approach and scaling were selected, it was necessary to construct a specific format for data collection. The researcher examined the data collection method used by Wenger in his classification of goods due to

its convenience and the fundamental affinity of purpose, taxonomical approach, and scaling. During the follow-up process, Wenger had discovered that his major difficulties in data collection were due to the fact that survey participants felt they either lacked the necessary knowledge or the necessary time to score the model (Wenger, 1990, p. 54). In the opinion of the researcher, neither of these difficulties was a fault of the model format, per se. Instead, selection of the survey population size and content would have to be modified to counteract these data collection difficulties. These issues are addressed further in section IV.B.1.. The model format selected by Wenger had been tested and proved adequate for data collection, and the researcher did not find a compelling reason to differ from its basic structure when constructing a pre-test model for classification of Government procured services.

The pre-test matrix was therefore patterned on that used by Wenger, and is provided in Figure 3-1. The grid compares services, listed in the left-hand column, to the twelve pre-test characteristics labeled by row at the top of the matrix. The researcher drafted pre-test instructions, and

PLEASE GRADE EACH SERVICE 1, 2, 3, 4, OR 5 USING SCALES & DEFINITIONS PROVIDED FOR EACH OF THE 12 CHARACTERISTICS

CHARACTERISTICS

SERVICE	1. CONSULTING	2. INFORMATION	3. MAINTENANCE	4. CUSTODIAL	5. LABOR	6. TRANSPORTATION	7. CONSTRUCTION	8. SUPPLY	9. LEGAL	10. ADMINISTRATIVE	11. PROFESSIONAL	12. TECHNICAL	13. SUPERVISORY	14. MANAGERIAL	15.	TOP THREE Characteristics in order of Preference	Comments
1.Biological Research	2															2, 3, 12	
2.Dining Facility Ops	2																
3.Weapons Engineering	5																
4. Linen Supply	2																
5. Roofing Repair	2																
6. Garbage Collection	2																
7. Non-local Trucking	2																
8. Indoor Painting	2																
9. Legal Consultation	1																
10. Automotive Repair	1																
11. Printing/Copying	1																
12. Furniture Repair	1																
13. Architect Design	2																
14. Grounds Maint.	2																
15. Computer Maint.	2																
16. Television Repair	1																
17. Guard Services	2																
18. Dentistry Clinics	1																
19. ADPE Programming	2																
20. Packing & Crating	2																

FIGURE 3-1

Pre-test Classification Model
Source: Researcher's Analysis and Wenger, 1990, p. 42

the full text is provided in Appendix B. They primarily directed scorers to:

1. Read the pre-test definition for a characteristic;
2. Grade each service (1-5) using the pertinent pre-test scale. Scorers were urged to read the scales with care since some may be counter-intuitive.
3. Provide comments or suggestions if the scorer believed that any pre-test definitions, scales, or service titles required modification, either by annotating the pre-test packet directly or writing them separately.
4. Repeat steps (1) through (3) above for each of the twelve pre-test characteristics provided in columns 1 through 12.

This procedure provided for detailed scorer feedback concerning the utility of pre-test instructions, definitions, scales, and service titles. A hypothetical scorer, for example, would fill out the pre-test matrix by reading the definition of **Customization**, read the scale and assign an appropriate numeric value. If the scorer assigned a "2" for the amount of **Customization** for the service **Biological Research**, as depicted in Figure 3-1, that value would signify that the scorer felt that production of **Biological Research** is not substantively altered for the Government. By grading twenty different services, the scorer would, presumably, refer repeatedly to the definition and different gradients of the scale. The scorer could conveniently annotate the definition

and scale with modifications to correct perceived deficiencies, and return them to the researcher.

The utility of characteristics was also checked a final time in preparation for the survey test. Preliminary characteristics, other than the twelve assigned to columns for the pre-test, were listed by dimension at the bottom of the pre-test matrix sheet. If a scorer believed that one of these characteristics had been prematurely rejected, or that a new characteristic should be considered, columns 13, 14, and 15 were available for pre-test. The scorer was instructed to write the nomenclature of the overlooked characteristic at the top of these columns, provide a definition and suggested scale, and grade the services accordingly.

As a final measure, all scorers were instructed to write their "top three characteristics in order of preference" for each service on the right-hand side of the pre-test matrix sheet. This was requested after all characteristics had been used to value the services in the manner depicted for **Customization** (column 1) of Figure 3-1. A sample Top Three ranking is also provided in the Figure, and the scorers would indicate which characteristics were <u>most</u> meaningful by providing rankings for a wide range of services. These rankings, conversely, would be used to examine the <u>least</u> meaningful characteristics as those which were rarely ranked

in the Top Three. This procedure is described further in section III.E.3.G., but the end-purpose of this ranking was that least-meaningful characteristics would be candidates for replacement by characteristics suggested by the scorers.

E. PRE-TESTING THE MODEL

Pre-test of the model, like the use of a prototype, was a measure that would serve to check the components of the model and its performance as a whole. Pre-test feedback would, at minimum, provide for modification of model features and ensure that all essential features had been included in the version used for the survey test.

1. Selection of the Services

Pre-testing and, eventually, survey data collection required the use of selected services. The researcher decided to select service titles from the Standard Industrial Classification (SIC) codes, as listed in Part 19 of the Federal Acquisition Regulation (FAR, 1990, p. 19-7). As discussed in section II.F., functional titles for services may be obtained from a number of Government publications, such as those relating to the Commercial Activities program. The researcher opted to use the FAR SIC titles since the FAR is the most predominant document in Government procurement, and both pre-test and survey participants would consequently be

comfortable referring to it if they questioned the use of a service title. The SIC codes are also categorized in a hierarchy with Divisions and Major Groups as tiers. This categorization would assist the researcher in providing a wide variety of services for pre-test and survey data collection.

The researcher's primary concern, however, was to provide services with generally recognizable and self-explanatory titles. As noted in section III.4., Wenger had discovered that perceived lack of knowledge and time on the part of survey participants produced difficulties during the data collection phase (Wenger, 1990, p. 54). If services were not generally recognizable, participants would likely feel less knowledgeable about their procurement. Obscure service titles would also require extra definitions, and additional participant time to read and contemplate them. The researcher therefore proceeded through the FAR SIC titles, highlighting those that, based on his experience, were generally recognizable and self-explanatory. If, in the opinion of the researcher, these titles were broad in scope, a specific service activity that is an appropriate subset of the SIC was furnished for the model.

It was desirable, however, to test the characteristics with a wide variety of functional service types to ensure that characteristic attributes such as differentiation and

concomitance were adequately tested. The researcher therefore attempted to select services from a wide variety of Major Groups. The researcher also referred to the regulatory division of services as reflected by the FAR (see Table 2-2). Services from each FAR Part were included to ensure that a wide distribution was selected. It should be noted, however, that the selection of services was not random, but was actively determined by the researcher. The service titles that were selected for both pre-test and survey data collection are provided in Table 3-11, along with their respective SIC and SIC title. The researcher also attempted to distribute these sample services in an irregular sequence, as provided in the left-hand column of Table 3-11 and Figure 3-1, but this sequence is not represented as being random in a thoroughly scientific sense. The researcher's main goal in sequencing was to ensure that a scorer would not readily perceive any grouping in the sequence, and would therefore be more attentive in looking for differences when scoring the successive services down the column.

2. The Pre-Test Panel

The pre-test panel of scorers consisted of seven of the original expert panel members and four Naval Postgraduate School professors, for a total of 11 members. The expert

TABLE 3-11

SAMPLE SERVICES & RESPECTIVE SIC CODES
Source: Federal Acquisition Regulation, 1990, Part 19.

Title Used	SIC	SIC Title
Biological Research	8731	Commercial, Physical and Biological Research
Dining Facility Ops	5812	Food Services
Weapons Engineering	8711	Engineering Services: Military Weapons
Linen Supply	7213	Linen Supply
Roofing Repair	1761	Roofing and Sheet Metal Work
Garbage Collection	4953	Refuse Systems
Non-local Trucking	4213	Trucking, Except Local
Indoor Painting	1721	Painting, Paper Hanging, and Decorating
Legal Consultation	8111	Legal Services
Automotive Repair	7538	General Automotive Repair Shops
Printing/Copying	7334	Photocopying and Duplicating Services
Furniture Repair	7641	Reupholstery and Furniture Repair
Architect Design	8712	Architectural Services (Other Than Naval)
Grounds Maint.	8744	Base Maintenance
Computer Maint.	7378	Computer Maintenance and Repair
Television Repair	7622	Radio and Television Repair Shops
Guard Services	7381	Detective, Guard and Armored Car Services
Dentistry Clinics	8021	Offices and Clinics of Dentists
ADPE Programming	7371	Custom Computer Programming Services
Packing & Crating	4783	Packing and Crating

panel members were those who, during the interview process, had been most receptive to the research effort and had indicated a willingness to pre-test the model. The professors chosen were from the acquisition and contracting faculty and

were intimately familiar with the peculiarities of Government procurement. Their inclusion would conceivably serve to provide fresh input to the classification scheme.

3. Pre-Test Results

The pre-test panel provided detailed feedback on the validity of components, as well as model performance as a whole, for the model's end-purpose. Criticisms and suggested modifications are summarized below.

a. Feedback on Overall Model Performance

Two members stated that the model required too much time to complete, and one of them asserted that "...only the very interested will complete it....The entire package is rather complicated." That member continued by suggesting that:

> ...it might be better to conduct it in person in groups of buyers, i.e. have the interviewer explain the characteristics, put them on a slide, transparency, and then have them complete it. Then discuss the results. A type of focus group. They might be able to suggest an easier and faster way to do it.

b. Feedback on Instructions

The instructions asked scorers to write the Top Three Characteristics for each service in order of preference (see Appendix B). One member argued that the word "preference", as used in the instructions and in the Top Three column of the model matrix, was not specific enough in

directing scorers. The words "...most important... characteristics that provide the most distinguishing characteristics" were suggested as a replacement.

c. Feedback on Service Titles

The only direct criticism of a service title, per se, was that the acronmym "ADPE" may not be familiar to scorers. ADPE, an acronym for Automated Data Processing Equipment, is the statutory term used in the Brooks Act to, amongst other things, describe computer hardware (P.L. 89-306, 1982, para 759). While the General Services Administration (GSA) has recently used Federal Information Processing (FIP) as an umbrella term for anything that the Brooks Act includes, ADPE is still a valid term for use in Federal procurement (GSA FIRMR, 1989, para 201-4.001). The service title "ADPE Programming" was therefore meant to describe custom computer programming services.

One member stated that "...a major service you have omitted is custodial." The member evidently thought that a service was meant to be, or should be, "major" to be graded by scorers. Another provided the comment "...make this also engineering (A/E)" in the matrix row for the service title Architect Design.

d. Feedback on Characteristic Definitions

While one mispelling was corrected, no modifications were suggested to alter the substantive content of the pre-test characteristic definitions as provided in section III.C.5..

e. Feedback on Characteristic Scales

Three pre-test characteristic scales were criticized (see section III.D.2.). According to one member, the scale for the characteristic **Expertise** was "weak", and another stated:

> ...it does not indicate boundaries for "lengthy". Many of our common services do benefit from having highly experienced managers. Therefore I am further defining brief as six months - one year, lengthy as one - four years, very lengthy as four - six years, and extremely as above six years.

The scale for the characteristic **Measurability** was criticized for providing inadequate descriptive differentiation along the scale from grade three to grade five. The member suggested that grade three be modified by the use of the word "moderate" or "somewhat" before the word "difficult", and that the language in grade four be replaced entirely with words that provide an appropriate midpoint between a modified grade three and the pre-test version of grade five.

The scale for the characteristic **Risk to the Government** was similarly criticized for providing inadequate descriptive differentiation from grades one to four. In particular, the difference between the adjectives "conspicuous" (grade two), "considerable" (grade three), and "great" (grade four) seemed to be, at best, modest to the member.

f. Proposals for Additional Characteristics

One member proposed that the preliminary characteristic *Small/Disadvantaged Business Development* be included in the model. The suggested scale was that grade one be "No opportunity" and grade five be "Very much opportunity", with no scale suggested between these extremes. The member scored the services with this scale, but did not provide a suggested definition. Another member suggested that the model use preliminary characteristics *Buyer Judgment*, *Tangibility*, and *Documentation*. The services were scored for these characteristics, however, definitions and scales were not suggested. In opposition to any additions, one member stated that the "Other Candidate Characteristics are really included in your twelve." Moreover, since no new characteristics were suggested as additions, this feedback served at minimum to validate the comprehensive nature of the list of preliminary characteristics (see Table 3-10).

g. *Proposals for Deletion of Characteristics*

One member commented that pre-test characteristics "...4, 5, 6, 7, 8, 9, & 12 seem to add little to the model. I think they could be deleted." While no other members proposed specific deletions, the researcher used the scoring of "top three characteristics in order of preference" for each service to summarize panel feedback on the relative strategic importance of the pre-test characteristics. One member did not provide a different Top Three for each service, but provided one Top Three for the all twenty services. He stated that "I do not feel respondents will be able to <u>change</u> their Top Three choices by type of service and I'm not so sure this is important. Why not elimiminate it?"

Ten members scored different Top Three rankings for each service, and these were totalled to see which characteristics were <u>not</u> frequently included in the Top Three. Specifically, **Negotiation**, **Perishability**, and **Risk to the Government** were respectively, the *least preferred* pre-test characteristics in terms of times they were excluded from the Top Three for <u>each service</u>. For example, **Negotiation** was the *least preferred* of all pre-test characteristics since it was only in a Top Three on 12 occasions, while Perishability was the next *least preferred* since it was in a Top Three on 17 occasions. By way of comparison, a characteristic that had

been in every Top Three received from the panel would have been included on 220 occasions (20 services X 11 panel members).

The researcher also tracked the total number of members that used a pre-test characteristic in any of their Top Three rankings. For example, **Confidentiality** and **Negotiation** were each used by five different members in their Top Three's, while **Labor % of Cost**, **Risk to the Government**, and **Perishability** were used by six of the 11 members. By that all member measure, **Confidentiality** and **Negotiation** were equally *least preferred*.

The researcher used these measures of total Top Three frequency for each service and for all members to decide which characteristic was the *least preferred*. Since **Negotiation** was a *least preferred* pre-test characteristic for either measure, the researcher decided that **Negotiation** should be replaced if an additional characteristic was selected for the model.

The researcher did not use the weighting provided by the "order of preference" of Top Three rankings. The intent in requesting an ordinal ranking was to intimate that the Top Three may, at minimum, change preference order from service-to-service, thereby spurring scorers to consider that

other characteristics may enter and leave the Top Three from service-to-service.

F. REVISION OF THE MODEL

Several changes were made to the model based on the feedback from the pre-test panel, and they are summarized below. The revised model is provided in Appendix C.

1. Changes to Model Matrix Sheet and Instructions

In order to reduce the "complicated" appearance of the package, the researcher decided to delete the list of other preliminary (candidate) characteristics from the bottom of the model matrix sheet and delete related instructions. The model matrix sheet and instructions were also revised by asking scorers to "write the Top Three Characteristics in order of *strategic importance*" instead of "order of *preference*". Finally, the instructions were streamlined so that they were only one page long to reduce both the "complicated" appearance of the model and time required to complete it.

2. Changes to Service Titles

No changes were made to service titles. The researcher believed that the acronym ADPE would be familiar to most procurement professionals. Furthermore, the researcher presumed that it would not be desirable to have persons score the matrix who were <u>not</u> familiar with this common term. In

using a survey method, the researcher was ultimately relying on scorers to evaluate whether their own experience was sufficient to score the matrix. The researcher decided that use of the term would mainly have a positive result as a screening device, since survey respondents would either omit scoring of the service "ADPE Programming" or note that they were unfamiliar with the term when they completed the matrix.

Custodial services were not added to the list of service titles since it was not intended to be a comprehensive list of "major" services. Addition of the word "engineering" to the title Architect Design was also rejected, since, in the opinion of the researcher, it would serve to make the service title less recognizable and self-explanatory.

3. Changes to Characteristic Definitions

The characteristic definitions were not changed from those provided in section III.C.5. since no substantive modifications were suggested.

4. Changes to Characteristic Scales

Three pre-test characteristic scales were changed in response to criticism from those provided in section III.D.2. to those provided in Appendix C. The researcher depended on his own judgment and experience to provide improved descriptions that differentiated between grades without imposing the reseacher's own a priori views.

The researcher reviewed the scale for the characteristic **Expertise**, and rejected the use of specific time frames to describe periods of "length" as being too prescriptive. Instead, the researcher decided that grade three of the scale should be changed to provide for greater differentiation between it and grade four. The adjective "moderately" was added to the original description of grade three, and is italicized in the result: "Expertise needed requires *moderately* lengthy or *moderately* expensive training/ qualification".

The scale for the characteristic **Measurability** was also revised to provide greater descriptive differentiation from grade three to grade five. The adjective "moderately" was added to the original description of grade three, and is italicized in the result: "Description and measurement of acceptable service performance is *moderately* difficult". The words "cryptic and laborious" were deleted from the original description of grade four and replaced with the words "quite complex". Grade four of the **Measurability** scale thus read that "Description and measurement of acceptable service performance is *quite complex*".

The scale for the characteristic **Risk to the Government** was similarly revised to provide adequate descriptive differentiation from grades one to four. The

adjectives *negligable* (grade one), *conspicuous* (grade two), *considerable* (grade three), and *great* (grade four) were respectively replaced with the adjectives *insignificant*, *slight*, *modest*, and *substantial*. For example, grade four thus read that "The likelihood and magnitude of potential harm to the Government due to service performance failure is *substantial*".

5. Addition and Deletion of Characteristics

The researcher decided to neither add nor delete characterics from the model. All of the suggested additions were selected from the list of preliminary characteristics (Table 3-10), and had thus been reviewed by the expert panel. One panel member specifically stated that no additions were necessary, and of the 11 pre-test panel members, only two submitted additions. In the opinion of the researcher, this indicated that there were, at minimum, no gross omissions. Additionally, none of the additions were suggested by more than one member. For example, *Small/Disadvantaged Business Development* was suggested by only one member, despite the fact that it was provided to all panel members on the list of "Other Candidate Characteristics" at the bottom of the pre-test matrix sheet. Since the pre-test panel did not provide compelling feedback for any particular additions, the

researcher decided to review the case of the *least preferred* pre-test characteristic to decide if it should be replaced.

The *least preferred* characteristic, **Negotiation**, was not used frequently, however, it *was* used by five different members in their Top Three's. Just less than half (45.4%) of the eleven pre-test members therefore included it in a Top Three for a service. One member may not have excluded it since he selected the same Top Three for all services and rejected the idea of reviewing them in detail. In sum, suggested additions had already been rejected in favor of **Negotiation** during the expert panel interview process, suggested additions were only supported by one pre-test member, and **Negotiation** was only excluded by a slim majority pre-test panel members. The researcher therefore decided that the preponderance of panel opinion indicated that **Negotiation** should be maintained on the list of characteristics for the final data collection survey.

G. SUMMARY

This chapter has focused on how specific service characteristics were selected, defined, and scaled with the use of a literature review, expert opinion, and researcher analysis. These characteristics were then incorporated into a classification model that utilized a matrix to compare

characteristics with services. A sample of heterogeneous services was selected then selected for use in a pre-test that included the matrix, definitions, scales, and accompanying instructions. Expert panel input and researcher analysis were used to refine the elements of the model into a functional scheme.

The next chapter will describe the use of the model in a data collection survey, and the subsequent analysis of that data. The results provide a categorization of the twenty sample services into homogenous groups.

IV. DATA COLLECTION AND PRELIMINARY ANALYSIS

A. INTRODUCTION

This chapter will describe how the data collection package, consisting of the instructions, matrix, and characteristic definitions and scales (see Appendix C) was used to collect data. Next, production of a benchmark "a priori" classification of the sample services, for comparison with cluster results, is delineated. Preparation for cluster analysis is outlined, along with a description of two clustering methods. The chapter concludes with determination of the number of categories with which to group the services, and selection of a preferred clustering method for use in additional iterations.

B. DATA COLLECTION

The data collection model was equivalent, in its fundamental components, to that used by Wenger. As noted in Chapter III, the researcher was aware that a perceived lack of knowledge and lack of time on the part of the survey population had delayed that data collection process (Wenger, 1991, p. 54). The researcher had simplified both the content and appearance of the model to avoid unnecessary intimidation of scorers. While computer labeling was used for envelopes, the instructions were addressed and signed by hand, and a pre-

addressed and stamped return envelope was mailed with the data collection model. Other measures, explained later in this chapter, were taken to maximize collection of appropriate data.

1. Survey Method

The model package was mailed to 300 procurement professionals. This number was selected since the researcher desired to have, at a minimum, 50 useful responses. Regardless of any statistical inferences this number has, it is believed that it would be sufficient to test the functional qualities of the data collection model. The researcher used letters as a follow-up method, and they were sent to the survey population approximately one week after the mailing of the data collection model. The letter asked scorers to call if they had <u>not</u> received this survey or had any questions, but its main purpose was to serve as a reminder to busy individuals. If the researcher achieved a modest 20% positive response rate by use of this method, 60 useful surveys would be received.

2. Selection of Survey Population

The researcher decided to use three different populations of procurement professionals. One hundred fifty scorers were selected from the directory of Certified Professional Contracts Manager (CPCM). The CPCM designation

"...represents the highest level of qualification in the contracts profession" and requires extensive contracting experience and successful completion of a rigorous certification exam (Contract Management, 1991, p. 55). The researcher endeavored to select CPCM's who provided addresses at Government agencies external to the Department of Defense. Additionally, at least one member was selected from each Chapter of the National Contract Management Association in order to provide for the widest geographic dispersion.

One hundred prospective scorers were selected from the Department of Defense Competition Advocate Listing (DoD, 1990). Competition Advocates review the acquisition and contract management programs at their activities to ensure that competition is maximized and consistent with current legislation. They typically participate in the formulation of procurement strategies, and the designated Competition Advocate is often actually the top official in the activity's procurement division. The researcher selected the Competition Advocates for those activities which, in his opinion, were most likely to procure a wide variety of services.

Fifty prospective scorers were selected from a list of attendees at a March, 1991 DoD Procurement Conference (DoD, 1991). The researcher endeavored to select individuals from Government agencies which had not been included in the

previous population selections. Failing that, the researcher endeavored to select individuals from activities which, in his opinion, were most likely to procure a wide variety of services.

3. Use of Instructions to Maximize Qualified Responses

While use of these populations ensured that scorers were survey professionals, it did not ensure that they had "enough" expertise with procurement of services. The issue of surveying "qualified" individuals was rather problematic, since the matter of how much expertise is "enough" would need to be decided. Additionally, a detailed and time-consuming screening process would be necessary to ensure that scorers had "enough" expertise. An *efficient* method of screening prospective scorers would be to select an organization that procured a wide variety of services, and use the applicable procurement professionals as a survey population. This method, though, may not be *effective* for the scheme's end-purpose since the population base would be narrow and might be remote from using a Government-wide approach. Selection of several organizations could be orchestrated, but this would require a vast amount of time for screening that would delay model development.

The researcher decided that, at this stage in model development, the self-knowledge of the selected survey

participants was a sufficient screening mechanism. As procurement professionals, prospective scorers could judge for themselves if they had enough expertise to complete the matrix. If participants believed they lacked the knowledge necessary to score the model, taxonomy development would be best-served if they, in fact, did not score the model. Hasty completion of the model would also be detrimental to development of a useful taxonomy, and the researcher endeavored to minimize these outcomes. The researcher believed that prospective scorers would be likely to use lack of familiarity as a justification (to themselves) to avoid expending the time necessary to complete the matrix. This opinion was reinforced by the fact that Wenger had noted that lack of familiarity with the sample items "...was the most frequently cited reason for not completing the matrix." (Wenger, 1990, p. 54) The data collection model instructions therefore did not directly address respondent qualifications and instead asked that individuals complete it "based on your expertise in Government procurement" (see Appendix C). Additionally, the instructions asked that if scorers did not have the time to complete the matrix but knew "a procurement professional who does" to "please forward this package to that person" (see Appendix C). Prospective scorers who used this provision would presumably need to justify the transfer of the

package (to themselves and the other person) by giving it to professionals who were equally or better qualified to complete it. The possible outcomes of this particular data collection package would be:

1. Scorers who felt comfortable completing the matrix and had the time would do so, or;
2. Scorers who were uncomfortable with the matrix or lacked sufficient time would:
   A. Pass it on to a qualified individual with sufficient time;
   B. Not respond, or;
   C. Respond negatively.

The cumulative effect of the instructions, in the opinion of the researcher, was to provide scorers with an outcome that suited their self-judgment and circumstances. Uncomfortable or busy individuals could pass the matrix on and avoid the need to provide a careless or hasty response. Completed matrices would, therefore, most likely be the product of careful, deliberative, and professional, judgment, and should provide valid data for cluster analysis.

4. Survey Response Statistics

One hundred ten of the 300 surveys were returned. Sixty six were "positive" responses, where all of the sample services were scored using the twelve characteristics on the returned matrix and the Top Three characteristics were

selected for each service. Nineteen were "partially positive", where all of the sample services were scored using the twelve characteristics, but the Top Three characteristics were not selected for each service. "Partially positive" responses would be useful for clustering, but would not be useful for determining the relative strategic importance of characteristics.

Twenty-five of the responses were "negative", meaning only *some* of the services were scored and a lack of *knowledge* was cited; *none* of the services were scored and a lack of *knowledge* was cited; some characteristics were not scored numerically due to a lack of knowledge, or; *none* of the services were scored and a lack of *time* was cited.

Several phone calls were received where the procurement professional questioned the researcher about qualifications necessary to complete the matrix. The researcher described the methodology and purpose of the data collection effort, and left the onus on callers to determine (based on their own perceptions) whether they were expert "enough" to complete the matrix. Many of the telephone calls were prompted by receipt of the follow-up letter. In seven cases the caller had not received the original package and data collection models were immediately mailed to them. Telephone conversations of any sort were not considered to be

a "response". The data collection package, and its follow-up letter, were returned by the postal service for one addressee due to an incorrect address. This also was not considered to be a "response". Table 4-1 provides a synopsis of survey responses by type.

The synopsis reveals that the most common reason for a "negative" response was a lack of knowledge, where either some or all of the services were not scored due to a perceived lack of knowledge on the part of the scorer, or some characteristics were not scored numerically. However, 190 of 300 (63.33%) surveys were not returned. Since only one data

TABLE 4-1

SURVEY RESPONSES BY TYPE
Source: Researcher's Analysis

Response Type	Number of Responses	% of Total Responses(110)	% of Survey Population(300)
Positive	66	60.00%	22.00%
Partially Positive	19	17.27%	6.33%
Sub-total (useful for clustering)	*85*	*77.27%*	*28.33%*
Negative-knowledge	24	21.82%	8.00%
Negative-time	1	.91%	.33%
Total responses	*110*	*100.00%*	*36.67%*

collection package and follow-up letter were returned for incorrect addresses, it is likely that these letters were opened by someone at the intended address, and that person either lacked the time or knowledge to complete the matrix. The researcher cannot presume, however, to decide what was the reason (time, knowledge, or something else) for the non-responses. The information is inconclusive.

Of the positive and partially positive responses, 24 (28.24%) were returned by persons other than those to whom the data collection package and follow-up letter were originally sent. As discussed in section IV.B.3. above, the researcher presumed that the transfer of the data collection package by the original procurement professional would be justified (to themselves and the other person) by giving it to professionals who were equally or better qualified to complete it. Of the negative responses, 8 (32.00%) were returned by persons other than those to whom the data collection package and follow-up letter were originally sent. In the opinion of the researcher, the reasons for 3.76% differential in transfers of negative and positive responses may only be presumed and are inconclusive. The provision for tranfer of the package, however, was utilized enough that inclusion of such a proviso in the instructions can significantly increase the response rates. The total response rate for this survey was 36.67%,

10.67% of which were transfers. The cumulative effect of the instructions and follow-up letter, in the opinion of the researcher, succeeded in providing responses that were the product of careful, deliberative, and professional, judgment, and provided valid data for cluster analysis.

C. AN "A PRIORI" CLASSIFICATION

In preparation for cluster analysis, the researcher sought to construct an "a priori" classification of the sample services. An "a priori" classification would serve as an objective benchmark for comparison with the results of the cluster analysis. The results of the cluster analysis need not be equivalent to the "a priori" classification, and differences, in fact, may be desirable for the production of strategic insights concerning the classification of Government procured services. If clustering results were fundamentally different from a sound benchmark, however, the rational basis of the clustering would need to be closely examined. The benchmark should be produced "a priori", beforehand, as opposed to an after-the-fact comparison to avoid any "a posteriori" rationalization of the results. (Romesburg, 1984, p. 258)

1. Desirable "A Priori" Scheme Attributes

The researcher decided that the most important qualities for an "a priori" classification would be that it use strategically significant characteristics, i.e. those among the twelve used in the data collection model, and that it produce between five and seven categories. The range of five to seven categories was selected since the researcher desired to keep the scheme as simple as possible, but have it produce potentially informative results. According to a pathbreaking psychological study, the mind cannot reliably hold in short-term memory more than five to seven separate items. Specifically, it was found that there

> ...is a span of absolute judgment [immediate memory] that can distinguish about seven categories and that there is a span of attention that will encompass about six objects at a glance. (Miller, 1956, p. 90)

Furthermore, "...retrieval from long-term memory is limited to about five items at a time." (Eysenck, 1977, p. 113) The production of eight or more categories may be too many for a user to manipulate mentally, while four or fewer categories may not provide sufficient refinement of the strategic differentiation between services.

2. Use of a Government "A Priori" Scheme

As discussed in section II.F., no scheme currently exists to classify Government procured services on a strategic

basis. The researcher first examined the regulatory structure provided by the Federal Acquisition Regulation (Table 2-2) as an "a priori" scheme. As a primary disadvantage, this scheme would predominantly reflect the characteristic **Regulation**, which had not been selected with the final twelve strategically significant characteristics. The researcher also attempted to classify the twenty sample services with categories that are specifically defined by the FAR or left in a category of general services. This review produced eight categories (that are either in italic or bold print on Table 2-2), a second disadvantage since the researcher considered this number to be excessive for mental manipulation by potential users. Finally, this effort produced seven categories with **two or fewer** sample services in them, and an "other", eighth category of general services with **twelve** sample services in it. To illustrate, "Biological Research" and "Weapons Engineering" were the only sample services that the researcher assessed as subject to the category of *Research and Development Contracting* (FAR, 1990, Part 35). At the other extreme, "Dining Facility Ops" was assessed as pertaining to *Service Contracts-General* (FAR, 1990, Part 37.101), along with eleven other sample services. This disproportionate grouping was understandable since special laws are drafted to manage exceptionally unique services,

however, the researcher decided that this statutory grouping was deficient from a strategic viewpoint. While asymmetry does not necessarily indicate that a classification is deficient, the researcher believed that statutorial classification largely overlooked the differences between many of the more common services.

3. Use of a Non-Government "A Priori" Scheme

The researcher therefore considered the non-Government classification schemes provided in Chapter II and concluded that the scheme suggested by Thomas (see Table 3-5) was the most valuable for use in an "a priori" classification. The Thomas scheme is production-oriented, relying on characteristics of service production per se. It first divides services into two bases: as either being primarily equipment based or primarily people-based, a division that corresponds to the characteristic **Labor % of Cost**. It next divides those types by the relative skill of labor used in their production, a division that corresponds to the characteristic of **Expertise** (Thomas, 1978, p. 159). The researcher decided to alternate between the two bases (labor and equipment-based) in sequence from relatively unskilled-to-skilled types of labor until labor-based, highly skilled services was the most complex service category. This method

produced the five categories of services that are provided in Table 4-2.

The researcher relied on examples provided in the Thomas article and his own experience to classify the twenty

TABLE 4-2

"A PRIORI" CLASSIFICATION OF 20 SAMPLE GOVERNMENT PROCURED SERVICES
Source: Thomas, 1978 & Researcher's Analysis

Category Number	Primary Production Basis	Labor Skill Level	Sample Services Classified in Category
I	Labor	Unskilled	Roofing Repair, Garbage Collection, Indoor Painting, Grounds Maint., Packing & Crating
II	Equipment	Unskilled	Linen Supply, Printing/Copying
III	Labor	Skilled	Dining Facility Ops, Furniture Repair, Computer Maint., Television Repair, Guard Services, ADPE Programming
IV	Equipment	Skilled	Non-local Trucking, Automotive Repair
V	Labor	Professional	Biological Research, Weapons Engineering, Legal Consultation, Architectural Design, Dentistry Clinics

sample services using this scheme. These categories are listed in ascending sequence (I to V) based on the increase in complexity perceived by the researcher, the sequence of services within each category matched that used on the data collection matrix and is not intended to be in simple-to-complex sequence.

D. PREPARATION FOR CLUSTER ANALYSIS

The positive and partially positive responses provided scoring of all twenty sample services with all twelve strategic characteristics. The data in these 85 matrices required consolidation and conversion for use in cluster analysis.

1. Data Preparation for Cluster Analysis

The data representation of objects to be clustered can take many forms. A common form is a coordinate matrix, "...in which the rows are observations and the columns are variables." (SAS, 1985, p. 46) The data collection survey had collected observations (scores) in rows for each service in columns for each variable (characteristic). In order to summarize this data for cluster analysis, the researcher needed to calculate mean values of the 85 observations in order to have a single observation value for each variable.

The researcher first assigned numbers 1-85 to the observations for record keeping purposes, since the identities of the individual respondents were immaterial to this stage of the research effort. The researcher then used a computer spreadsheet program to convert the data into 20 separate matrices, one for each service. Each respondent code was entered on the vertical axis, and the twelve characteristics were represented on the horizontal access with codes that corresponded in sequence (1-12) to that used on the data collection matrices. Next, the score assigned by each respondent for each of the characteristics, as they related to the pertinent service, was entered into the matrix. A mean value for each characteristic was then computed by averaging the individual cell scores. Finally, the mean scores were recombined into a single matrix, which depicted the mean values of each characteristic for each service used in the data collection effort. The mean value matrix is displayed in Table 4-3.

2. **Cluster Analysis Strategy**

As explained in section II.D., cluster analysis is an example of grouping, where the researcher deduces the classification scheme from the data analysis. Grouping requires less "a priori" knowledge of the characteristics that are likely to be valuable for the classification effort, and

TABLE 4-3

MEAN VALUE MATRIX
Source: Researcher's Analysis

SERVICE	C.1	C.2	C.3	C.4	C.5	C.6	C.7	C.8	C.9	C.10	C.11	C.12
BioRsch.	3.35	4.48	4.54	3.53	4.45	4.07	3.81	4.01	4.05	3.73	3.72	4.34
DinFacOps.	2.36	2.12	1.48	3.51	2.01	1.25	2.61	2.82	2.53	1.84	1.93	1.49
Wpns.Engi	4.66	4.46	4.39	3.41	4.34	4.72	4.55	4.62	4.73	3.73	4.18	4.52
Linen Sup	1.56	1.62	1.29	3.08	1.64	1.2	1.95	2.02	1.92	1.56	1.49	1.53
Roof Repa	1.64	2.24	1.69	2.96	2	1.18	2.45	2.31	1.98	1.36	1.76	3.11
GarbColle	1.41	1.41	1.24	3.39	1.38	1.14	2.15	2.18	1.91	1.72	1.51	1.52
NonlocTrk.	1.74	2.04	1.55	2.89	1.71	1.28	2.27	2.15	1.99	1.54	1.8	1.62
IndoPaint	1.71	1.87	1.34	3.35	1.73	1.11	1.75	2.06	1.84	1.26	1.61	2.61
LegalCons	2.85	4.18	3.36	4.52	3.52	3.48	3.61	3.21	3.25	2.65	2.84	2.89
AutoRepai	1.64	2.71	2.33	2.89	2.06	1.35	2.33	2.25	2.06	1.55	1.82	2.32
PrintCopy	1.82	2.05	1.78	2.89	1.59	1.54	2.06	1.91	1.81	1.47	1.73	1.56
FurnitRep	1.47	2.31	1.78	3.36	1.84	1.16	1.64	1.88	1.8	1.65	1.53	2.41
ArchDesig	2.87	3.99	3.44	4.22	3.59	2.44	3.21	3.45	3.52	2.67	3.18	3.84
Grnds.Mai	1.62	1.67	1.29	3.68	1.66	1.15	1.75	2.27	2.12	1.55	1.73	1.8
CompuMain	2.12	3.22	2.99	3.44	2.54	2.15	3.26	3	2.88	2.19	2.32	2.49
TV Repair	1.38	2.68	2.42	3.22	2.02	1.34	1.86	2.07	1.95	1.66	1.65	2.35
GuardSvcs.	2.53	2.04	1.42	4.4	1.85	2.04	3.48	2.58	2.48	2.01	1.86	1.75
DentClini	2	4.04	3.36	3.54	3.06	1.52	2.55	2.98	3.02	3.06	2.33	2.79
ADPE Prog	3.21	3.72	3.65	4.13	3.46	3.34	3.69	3.67	3.61	2.76	3.41	3.45
Pack&Crat	2.34	1.94	1.36	3.19	1.71	1.29	2.29	2.11	2.05	1.8	1.67	1.78

All numbers are rounded to nearest 1/100th. Clustering performed with numbers extended to 1/1,000,000.

KEY:

C.1 = Customization	C.5 = Measurability	C.9 = Negotiation
C.2 = Expertise	C.6 = Confidentiality	C.10 = Competition
C.3 = Complexity	C.7 = Risk to Govt	C.11 = Stability
C.4 = Labor % of Cost	C.8 = Govt Attention	C.12 = Perishability

it also reduces the ability of the researcher to impose his own views on the construction of categories. Grouping procedures are also "...better equipped to handle large numbers of categorical terms or properties." (Hunt, 1983, p. 353)

The data provided in the mean value matrix reflects the consolidated "a priori" knowledge of research literature, the interview panel, pre-test panel, respondents, and the researcher. Cluster analysis investigates the systematic (or latent) structure of a data matrix. The term "structure" is used to mean "...the orderly groupings of data points in the data matrix....A major contribution of cluster analysis is its ability to reveal such natural groupings." (Fleishman & Quaintance, 1984, p. 78) Clustering techniques have been used in such diverse areas as studies of manufacturing firms, supervisors (McKelvey, 1982, pp. 46-47), neurology, psychology, and thesauri (Fleishman & Quaintance, 1984, pp. 80 & 197). In addition to providing structure to a classification scheme, clustering may be used to confirm, refine, or revise a pre-existing scheme. Cluster models are numerous and diverse, but they share:

> ...the common property of separating phenomena into groups that maximize both the degree of "likeness" *within* each group and the degree of differences *between* groups according to some objective function. (Hunt, 1983, p. 354)

In the common hierarchical techniques, agglomerative methods start by assigning each constituent Operational Taxonomic Unit (OTU) to its own taxon. In this case, each service would belong to its own separate, single member cluster. The two OTU's having the highest similarity are then grouped together into categories (taxa). Similarity, or "distance", is calculated according to the clustering method. The distance among the remaining single member clusters and the two OTU cluster is then compared, with the number of groups being reduced by one in an iterative process until, ultimately, they may be combined into a single cluster. (Dunn & Everitt, 1982, p. 77)

3. **Cluster Analysis Technique**

The researcher employed two hierarchical clustering methods, average linkage and Ward's minimum variance. While "...no single method is best in every situation", group average clustering and Ward's method have been specifically identified as providing sound methodologies (Dunn & Everitt, 1982, p. 87). Comparison of these methods, along with the "a priori" scheme, provided a basis for verifying the validity of resultant taxonomic clusters.

Both methodologies attempt to maximize the internal homogeneity of taxa, and the Statistical Analysis System (SAS) program was used to execute their specific procedures. The

average linkage method defines the distance between clusters as the arithmetic average between the combined OTU's in one cluster and the OTU's in another (SAS, 1985, p. 263). It combines OTU's iteratively as outlined in section IV.D.2., above, and may be ceased at the desired number of clusters. Ward's method seeks to minimize the distance (as a sum of squares of each character) within a cluster. Given a desired number (k) of taxa, Ward's method will partition the OTU's into k clusters which have the lowest within-cluster sums. This procedure can only guarantee a local, vice global, minimum within-cluster sum (Dunn & Everitt, 1982, p. 88), Methods based on the least-squares criterion, such as Ward's, "...tend to find the clusters with roughly the same number of observations in each cluster", while average linkage "is somewhat biased toward finding clusters of equal variance." (SAS, 1985, p. 48)

E. INITIAL CLUSTER ANALYSIS RESULTS

The researcher's initial objective was to verify the general validity of cluster analysis output. Comparison with the "a priori" benchmark provided a means to verify whether the statistical output generally conformed with strategic "common sense". The researcher then evaluated the number of clusters/categories that would be desirable for the

classification scheme. Finally, the researcher selected a preferred clustering method for grouping services and for simplification of the classification scheme.

1. Comparison Between Clustering Methods and the "A Priori" Model using Two Characteristics

As discussed in section IV.C.3., the "a priori" scheme was based on two characteristics, suggested by Thomas, that are elements *internal to production* of a service. Since the data collection model used twelve characteristics, including those that are elements of a service per se but are *external to production*, characteristics of the *Government's effort*, or of the procurement *environment* (see Table 3-9), the researcher did not expect the results to be equivalent. The researcher therefore first compared the "a priori" scheme with clustering output using only **Labor % of Cost** and **Expertise** as input variables. This would allow the researcher to focus his "a posteriori" analysis on the production factors of the labor/equipment ratio and the relative skill level of labor. The researcher used each procedure to produce a desired number (k) of five taxa, since five taxa would allow for complete comparison with the "a priori" scheme. The resultant taxonomical clusters at the five cluster level are provided in Table 4-4.

TABLE 4-4

COMPARISON BETWEEN "A PRIORI" MODEL & TWO CLUSTERING METHODS AT FIVE CLUSTER LEVEL USING EXPERTISE & LABOR % OF COST CHARACTERISTICS

Source: Researcher's Analysis

"A PRIORI"	WARD'S MIN. VAR.	AVERAGE LINKAGE
Category I	Cluster One - 2.48	Cluster One - 2.58
Roofing Repair	-*Linen Supply*	=**Dining Facility Ops**
Garbage Collection	Garbage Collection	-*Linen Supply*
Indoor Painting	Grounds Maint.	Roofing Repair
Grounds Maint.		Garbage Collection
Packing & Crating	Cluster Two - 2.70	*Non-local Trucking
	-*Dining Facility Ops*	Indoor Painting
Category II	-*Roofing Repair*	-*Printing/Copying*
Linen Supply	=**Non-local Trucking**	=**Furniture Repair**
Printing/Copying	-*Indoor Painting*	Grounds Maint.
	Printing/Copying	Packing & Crating
Category III	-*Furniture Repair*	
Dining Facility Ops	-*Guard Services*	Cluster Two - 3.03
Furniture Repair	-*Packing & Crating*	=**Automotive Repair**
Computer Maint.		-*Computer Maint.*
Television Repair	Cluster Three - 3.03	-*Television Repair*
Guard Services	-*Automotive Repair*	
ADPE Programming	Computer Maint.	Cluster Three-3.22
	Television Repair	Guard Services
Category IV		
Non-local Trucking	Cluster Four - 3.91	Cluster Four - 3.91
Automotive Repair	-*Biological Research*	-*Biological Research*
	-*Weapons Engineering*	-*Weapons Engineering*
Category V	-*Dentistry Clinics*	-*Dentistry Clinics*
Biological Research		
Weapons Engineering	Cluster Five - 4.13	Cluster Five - 4.13
Legal Consultation	Legal Consultation	Legal Consultation
Architectural Design	Architectural Design	Architectural Design
Dentistry Clinics	=**ADPE Programming**	**ADPE Programming**

TOTAL NUMBER OF SERVICES DIFFERING FROM "A PRIORI":

-*By One Category/Cluster*:	9	8
=**By Two Categories/Clusters**:	2	3
*By Three Categories/Clusters:	0	1
Total:	11	12

The clusters were assigned numbers one through five based on the combined mean value of their constituent services for the characteristics of **Labor % of Cost** and **Expertise.** This mean value increases from cluster one to cluster five, and is indicated next to the title of each cluster. For example, the constituent services of the first cluster of Ward's minimum variance are "linen supply", "garbage collection", and "grounds maintenance". The combined mean value of these services for the characteristics of **Labor % of Cost** and **Expertise** was calculated to be 2.48 from their corresponding values in the mean value matrix (provided in Table 4-3). Since this value is lower than that which was calculated for the other clusters, this cluster was designated as cluster one.

The sequence of the services within each cluster, however, was the order that was used for the data collection model. This is a convenient order for scanning the data, and is not intended to be a respresentation of the simple-to-complex ranking <u>within</u> each cluster. The *relative ranking* of the services by each characteristic is already known from the data collection process. The purpose of this stage in the classification effort is to verify the validity of cluster analysis methods, given the data collection input, to reveal "natural groupings" that correspond to "common sense" (as

represented by the "a priori" benchmark). The groupings that were eventually selected would eventually serve to identify the *mean values and boundaries* of each category.

To aid in comparison, the constituent cluster services in Table 4-4 are printed in different styles and the data are summarized at the bottom. Services that were in the equivalent cluster to their "a priori" category are printed in normal typeface. Services that differed by one cluster from their corresponding "a priori" category are typed in italic and marked with a "-" (dash) on the left-hand side. Services that differed by two clusters from their corresponding "a priori" category are typed in boldface and marked with a "=" (equal sign) on the left-hand side. Services that differed by three clusters from their corresponding "a priori" category are typed in boldface and marked with an "*" (asterix) on the left-hand side.

The "a priori" scheme and cluster analysis methods produced different results. Review of Table 4-4 reveals that in 11 cases Ward's minimum variance produced equivalent categories, while the average linkage method produced eight equivalents. In nine cases Ward's minimum variance grouped services one cluster away from their corresponding "a priori" category, while the average linkage produced eight such differences. In the opinion of the researcher, such

deviations were not critical since the mean values of adjoining clusters were not radically different and a moderate change in the boundaries of the clusters would have produced equivalent groupings. The researcher, however, reviewed deviations of more than one cluster in detail.

Ward's minimum variance grouped two services, "ADPE programming" and "non-local trucking", two clusters away from their corresponding "a priori" categories. "ADPE programming" was grouped "a priori" by the researcher into category III of the scheme as a *labor-based, skilled* labor service. The groupings would have been equivalent if the researcher had designated "ADPE programming" as a *labor-based, professional* service. The researcher therefore had to assess "a posteriori" whether "ADPE programming" requires a similar skill level to that represented by the services listed in category V of Table 4-4. The researcher believed that, of the labor-based services listed in category III of Table 4-4, "ADPE programming" required the most similar labor skill level to that required by category V services. The researcher therefore considered that this mismatch did not violate "common sense".

"Non-local trucking" was also clustered one skill level different from its "a priori" category. The researcher had designated it as being an *equipment-based, skilled labor*

service. If it had been designated as an *unskilled* labor service, the category would have matched the cluster. The researcher reviewed the two services listed in category IV of Table 4-4 and believed that "non-local trucking" required less skill than automotive repair. The researcher therefore concluded that this mismatch did not violate "common sense".

The average linkage method grouped five services two clusters away from their corresponding "a priori" categories. "ADPE programming" was mismatched in an equivalent manner as that discussed above in the case of Ward's minimum variance. The researcher had concluded that this mismatch did not violate "common sense". "Dining facility operations" was clustered one skill level different from its "a priori" category. If the researcher had designated it as using *unskilled* labor (instead of *skilled*) the category would have matched the clustering. Of all the services listed in category III, the researcher believed that "dining facility operations" either required analogous or lower skill levels than the other category members. The researcher therefore concluded that this mismatch did not violate "common sense".

"Furniture repair" was also clustered one skill level different from its "a priori" category. If the researcher had designated it as as using *unskilled* labor (instead of *skilled*) the category would have matched the cluster. The researcher

believed that furniture repair could have been naturally grouped with "dining facility operations" at the low-skill range of category III and similarly mismatched into cluster one of the average link method. The researcher therefore concluded that this mistmatch did not violate "common sense".

"Automotive repair" was also clustered one skill level different from its "a priori" category. The researcher had designated it as being an *equipment-based, skilled labor* service. If it had been designated as an *unskilled* service, the category would have matched the cluster. The researcher reviewed the two services listed in category IV of Table 4-4 and believed that "non-local trucking" required less skill than automotive repair. "Non-local trucking" was grouped in a lower cluster than "automotive repair", so in that sense the grouping was "natural". The researcher, however, believed that "automotive repair" is <u>not</u> an unskilled labor service, and therefore concluded that this mismatch **may** violate "common sense".

Finally, "non-local trucking" was grouped three clusters away from its corresponding "a priori" category. In the opinion of the researcher, its grouping in a lower cluster than the other service of "a priori" category IV, "automotive repair" was "natural", but extreme. The researcher concluded that this mismatch **may** violate "common sense".

In comparison with Ward's minimum variance, the average linkage method produced a larger total number of mismatches (12), a larger number of two grouping mismatches (3), and the only three grouping mismatch. It should be noted, however, that many of these mismatches were caused by the size of cluster one. Cluster one contained all of the constituent services of category I of the "a priori" scheme, plus two of its one grouping mismatches, two of its two grouping mismatches, and its three grouping mismatch. A smaller category may have corrected this deficiency. As noted in section IV.D.3., the average linkage method is biased toward finding clusters of equal variance while Ward's tend's to find clusters with roughly the same number of observations.

The researcher concluded that, at this stage, Ward's was the superior clustering method since it produced only two mismatches of note, and these conformed with his "a postereori" perceptions. The researcher decided to test the average linkage method with all twelve characteristics at the five-cluster level, since it may have merely been inappropriate for use with two characteristics.

2. Comparison Between Clustering Methods and the "A Priori" Model using Twelve Characteristics

The researcher analysed the mean values of all twelve data collection characteristics using Ward's minimum variance

and average linkage methods. The results at the five cluster level are displayed along with the "a priori" scheme categories in Table 4-5. The clusters and services are arranged and marked in the same manner as they were for for Table 4-4.

Once again, the "a priori" scheme and cluster analysis methods produced different results. Review of Table 4-5 reveals that both Ward's minimum variance and the average linkage method produced eight service groupings equivalent to the "a priori" scheme. In seven cases Ward's minimum variance grouped services one cluster away from their corresponding "a priori" categories, while the average linkage produced four such differences. In the opinion of the researcher, such deviations were not critical since the mean values of adjoining clusters were not radically different and a moderate change in the boundaries of the clusters would have produced equivalent groupings. As in the case of the two characteristic comparison, the researcher reviewed deviations of more than one cluster in detail.

Ward's minimum variance grouped three services two clusters away from their corresponding "a priori" categories. "Furniture repair" and "television repair" were grouped "a priori" by the researcher into category III of the scheme as *labor-based, skilled* labor services. The groupings would have

TABLE 4-5

COMPARISON BETWEEN "A PRIORI" MODEL & TWO CLUSTERING METHODS AT FIVE CLUSTER LEVEL USING THE TWELVE DATA COLLECTION CHARACTERISTICS

Source: Researcher's Analysis

"A PRIORI"	WARD'S MIN. VAR.	AVERAGE LINKAGE
Category I	Cluster One - 1.76	Cluster One - 1.78
Roofing Repair	*-Linen Supply*	**=Dining Facility Ops**
Garbage Collection	Roofing Repair	*-Linen Supply*
Indoor Painting	Garbage Collection	Roofing Repair
Grounds Maint.	***Non-local trucking**	Garbage Collection
Packing & Crating	Indoor Painting	***Non-local Trucking**
	***Automotive Repair**	Indoor Painting
Category II	*-Printing/Copying*	***Automotive Repair**
Linen Supply	**=Furniture Repair**	*-Printing/Copying*
Printing/Copying	Grounds Maint.	**=Furniture Repair**
	=Television Repair	Grounds Maintenance
Category III	Packing & Crating	**=Television Repair**
Dining Facility Ops		Packing & Crating
Furniture Repair	Cluster Two - 2.09	
Computer Maint.	*-Dining Facility Ops*	Cluster Two - 2.19
Television Repair	*-Guard Services*	*-Guard Services*
Guard Services		
ADPE Programming	Cluster Three - 2.57	Cluster Three-2.92
	Computer Maint.	**=Legal Consultation**
Category IV	**=Dentistry Clinics**	**=Architectural Design**
Non-local Trucking		Computer Maint.
Automotive Repair	Cluster Four - 3.15	**=Dentistry Clinics**
	-Legal Consultation	ADPE Programming
Category V	*-Architectural Design*	
Biological Research	*-ADPE Programming*	Cluster Four - 3.70
Weapons Engineering		*-Biological Research*
Legal Consultation	Cluster Five - 3.86	
Architectural Design	Biological Research	Cluster Five - 4.02
Dentistry Clinics	Weapons Engineering	Weapons Engineering

TOTAL NUMBER OF SERVICES DIFFERING FROM "A PRIORI":

-By One Category/Cluster:	7	4
=By Two Categories/Clusters:	3	6
***By Three Categories/Clusters:**	2	2
Total:	12	12

been equivalent if the researcher had designated them as *labor-based, unskilled* services. The researcher therefore had to assess "a posteriori" whether these services require a similar skill level to that represented by the services listed in category I of Table 4-5. The researcher noted that almost all of the services listed in categories I, II, III, and IV were combined in clusters one and two of Ward's minimum variance. The only exceptions were "computer maintenance" and "ADPE programming". The researcher believed that this different treatment of computer-based services by the clustering of the data collection model was consistent with the end-use goal of providing a strategically useful classification scheme. While the *production* of computer services, as suggested by the Thomas "a priori" scheme, may be similar to the other services in category III, their *procurement* may be quite different. The researcher therefore concluded that the clustering was "natural" considering the end-use goal. The two services that were mismatched by three groupings, "non-local trucking" and "automotive repair", were also a result of this combination of categories I, II, III, and IV into clusters one and two. The researcher concluded that these mismatches did not violate "common sense".

"Dentistry clinics" was the final service that was grouped two clusters away from its corresponding "a priori"

category by the Ward's minimum variance method. The researcher was surprised by the lower grouping of this service, however, the fact that this service is not unique to the Government may account for its lower complexity vis-a-vis other services. It is possible that many other "a postereori" rationalizations may be formulated to explain this grouping, since twelve disparate characteristics were inputs in this clustering. The researcher concluded that this grouping may violate "common sense", but that a stronger conclusion could not be formulated.

Comparison with the average linkage results supported the relative validity of Ward's minimum variance method. Both methods mismatched 12 services in comparison to the "a priori" scheme, 11 of which were identical. Furthermore, the major difference between the two methods was that the average linkage method had mismatches that were more extreme in comparison to the "a priori" scheme. It grouped six services two clusters away from their corresponding "a priori" category. Mismatches of "furniture repair", "television repair", and "dentistry clinics" were identical to those resulting from Ward's method. Mismatches three clusters away from their corresponding "a priori" categories were also identical for the services "non-local trucking" and "automotive repair" to those resulting from Ward's method. As

discussed above, the researcher concluded that "dentistry clinics" mismatch was the only one that may have violated "common sense". However, the clustering results supported this grouping since both Ward's minimum variance and the average linkage method grouped "dentistry clinics" in cluster three.

The average linkage method also grouped "legal consultation" and "architectural design" two clusters away from their corresponding "a priori" categories. With the exception of "computer maintenance" and "ADPE programming", clusters three, four, and five, of both the average linkage and Ward's minimum variance method correspond with category V of the "a priori" scheme. The average linkage method, however, produced a larger, distant third cluster that is more of a mismatch when compared to the constituent members of category V of the "a priori" scheme. A two grouping mismatch of "dining facility operations" was also more extreme than a one grouping mismatch for Ward's method. While combination of the constituent services of clusters one and two would produce a virtually identical cluster, Ward's method seems more natural in comparison to the researcher's "a priori" perceptions.

The researcher concluded from this analysis that cluster analysis methods produced "natural" groupings from the data

collection input of all twelve characteristics at the five cluster level. Additionally, Ward's method appeared to be superior to average linkage for classifying Government procured services.

3. Determining the Number of Clusters

The researcher listed five classification principles in section II.E.. The first and second principles are, respectively, that the classification scheme should adequately specify the phenomenon to be classified and that the scheme specify properties or characteristics that will be used in classifying. The researcher had concluded that the cluster analysis input adequately specified the phenomenon to be classified, incorporated specific salient characteristics, and produced output that was valid for the end-use goal. The third and fourth classification principles are, respectively, that the categories be mutually exclusive and collectively exhaustive. The researcher had determined that the data collection scheme was collectively exhaustive, since a disparate range of services were classified using the twelve strategic characteristics. As demonstrated during the cluster analysis, the scheme was used to group these services.

The groupings, however, were not mutually exclusive, where, if an item fits one category, it will not fit any other category. As demonstrated during the cluster analysis,

services gravitated to different groups when different characteristics and methods were used. As demonstrated by the Thomas "a priori" scheme, however, the difference between many salient service characteristics is one of degree, and the relative difference between such designations as "skilled" and "unskilled" labor can be highly subjective. Cluster analysis provides a consistent, reproducible statistical basis for classifying a phenomenon. It also provides a means for consolidating the subjective evaluation of a large number of people. In this data collection effort, the evaluations of 85 professionals were consolidated into a form that may consistently use the same statistical techniques to produce identical results. The effect of changing the type and number of characteristics may also be consistently observed, reproduced, and manipulated. The different clusters that resulted from changing the input characteristics from two to 12, for instance, was observed in section IV.E.2.. Selection of the appropriate number (k) of clusters, however, may have the greatest effect on the property of exclusivity.

As described in section IV.D., the cluster analysis methods used in this effort start by assigning each constituent service to its own separate, single member cluster. The two services having the highest similarity are then grouped together into categories until, ultimately, they

may be combined into a single cluster. Only the last single, all-service cluster may be said to be truly mutually exclusive, since, without logical partitioning, only an all member cluster may satisfy the condition that a service not be able to potentially belong to any other category (cluster). A single, all-service cluster, however, is useless since it does not differentiate the services. The fifth classification principle is that the classification scheme should be useful. As discussed in section II.E., this criterion is the "first among equals". The researcher therefore considered both the relative exclusivity of clusters and their differentiation of services in selecting the appropriate number of categories (clusters).

At the two cluster level both methods produced clusters with 13 and 7 constituent services. In the researcher's opinion, these two categories did not provide useful differentiation of the services. At the 11 cluster level both methods still had seven services in single-member clusters. In the researcher's opinion, these categories were not exclusive enough, since slight changes in the values of the services would have altered the clustering. The researcher therefore decided to vary the number of clusters (k) between three and ten for all analyses for determining the appropriate number of clusters.

a. Application of the Kth-Nearest-Neighbor Method

There is no orthodox method that is used as a standard for determining the appropriate number of clusters. While "...there are no satisfactory methods for determining the number of population clusters for *any* type of cluster analysis..." [italics added], perhaps the "...best approach to the number of clusters problem" is the *k*th-nearest-neighbor method (SAS, 1985, pp. 65 & 67). The researcher first used this method to evaluate the appropriate number of categories.

The *k*th-nearest-neighbor method requires weak assumptions, namely that the observations are sampled independently and that each cluster corresponds to a mode of the population density. The 85 observations had been sampled independently. The mode is the most frequently occuring value in a series of observations, and for this application a modal cluster would have "...at least *n* members" and "...have a maximum density greater than the fusion density" for a cluster to be designated as a modal cluster (SAS, 1985, p. 260). The estimated density is essentially the number of observations within a sphere centered at mode *x* with a radius that is a function of *k* and *x*, divided by the volume of that sphere (SAS, 1985, p. 264). The method involves varying the number of clusters (*k*) and estimating the number of modal clusters. "If the estimated number of modal clusters is constant for a

wide range of k values, there is strong evidence of at least that many modes in the population." (SAS, 1985, p. 67)

The researcher analysed the mean value matrix data (Table 4-3) using different values of k and n. As discussed above, the researcher varied the value of k between three and ten for all analyses. During the first series of iterations, the value of n was equal to the quantity (k - 1) since "...the use of the kth-nearest-neighbor method limits the resolution that can be obtained for clusters with fewer than k members." (SAS, 1985, p. 260) This analysis produced two modal clusters at the three cluster level, but only one modal cluster in the four-to-ten cluster range.

Since the estimated number of modal clusters was not constant for a wide range of k values, this analysis was inconclusive. The researcher then set the value of n equal to one (1.00), even though this extremely low value would limit the precision of the identification of modes. This analysis produced three modal clusters for $k = 3$, two modal clusters at $k = 4$, and one modal cluster in the five-to-ten cluster range. The estimated number of modal clusters was not constant for this range of k values, either, despite the fact that n was extremely low relative to k. A contributing factor was that twenty services had been sampled in the data collection phase. A higher total number of services would have increased the

likelihood that a modal cluster would have at least *n* members, since more potential members would be available. The researcher decided that this method was inconclusive for the current input data, but may prove advantageous in the future if a large number of services are sampled.

b. Application of the Width of Range Method

Another method that may be used is to compare the width of the range for which the number of clusters remains constant, since "...a wide range indicates that clusters are well separated in the attribute space." (Romesburg, 1984, p. 213) The width of range is the difference between the average distances when one cluster is formed and the next cluster is formed. To illustrate this concept, the normalized average distances computed by the average linkage method, and their differences, are provided in numeric form in Table 4-6. These numbers were calculated to the nearest 1/1,000,000 but are displayed to the nearest 1/100 in the table.

The point of change in the number of clusters is commonly identified as a "cut". The width of range of the first cut is infinite (or undefined) since the difference between the normalized average distance of the previous zero, no-member cluster and that of the first, all member cluster is infinity minus 1.523. The width of range of the second cut,

TABLE 4-6

WIDTH OF RANGE OF NORMALIZED AVERAGE DISTANCES BETWEEN CLUSTERS
Source: Researcher's Analysis

CUT	Cluster Level *From*	Normalized Average Distance	Cluster Level *To*	Normalized Average Distance	Width of Range (Difference Between Clusters)	Rank of Width
1	Zero	Infinite	One	1.523	Infinite	1st
2	One	1.523	Two	1.142	.381	3rd
3	Two	1.142	Three	.730	.412	2nd
4	Three	.730	Four	.681	.049	5th
5	Four	.681	Five	.533	.148	4th
6	Five	.533	Six	.463	.071(rounded)	6th
7	Six	.4626	Seven	.4491	.0135	11th
8	Seven	.449	Eight	.419	.030	8th
9	Eight	.419	Nine	.390	.029	9th
10	Nine	.390	Ten	.359	.031	7th
11	Ten	.3593	Eleven	.3446	.0147	10th

the difference between the normalized average distance of cluster one and cluster two, is .381 (1.523 minus 1.142). Generally speaking, as each cluster is formed this value decreases, and the clusters iteratively get closer together and become less distinguishable. A relatively high range of width is therefore desirable since it is more difficult for a service to switch from one cluster to another. The categories are more mutually exclusive and are less "...sensitive to error" when the width of range is relatively large (Romesburg, 1984, p. 213).

Review of the "Rank of Width" column on the right-hand side of the table reveals the relative rank of the range of width of the first 11 cuts. While the range of width generally increases, on occasion it decreases as a new cluster is formed. Cut three, for example, has a smaller width of range than Cut two. After examining the relative rank of the range of width of the first 11 cuts, the researcher concluded that Cuts three, five, and eight were relatively less sensitive to error than the cuts that preceded and followed them. The clusters that were formed at these cuts were clusters three, five, and eight. The researcher therefore concluded that examination of cluster analysis results should be reviewed in detail at these levels to determine which level was most appropriate for achieving a balance between the goals of category exclusivity and service differentiation.

c. Selection of the "Optimal" Cluster Level

The researcher examined the results of the Ward's minimum variance and average linkage methods at the three, five, and eight cluster levels. The results at the five cluster level are displayed in Table 4-5, above. The results of both methods were identical at the three cluster level, while they differed at the eight cluster level, and these are displayed in Table 4-7, below.

TABLE 4-7

COMPARISON BETWEEN TWO CLUSTERING METHODS AT THREE AND EIGHT CLUSTER LEVELS USING THE TWELVE DATA COLLECTION CHARACTERISTICS

Source: Researcher's Analysis

THREE CLUSTER LEVEL	*EIGHT CLUSTER LEVEL*	
Both Methods	**Ward's Min. Var.**	**Average Linkage**
Cluster One - 1.18	Cluster One - 1.70	Cluster One - 1.70
Dining Facility Ops	Linen Supply	Linen Supply
Linen Supply	Garbage Collection	Garbage Collection
Roofing Repair	Non-local Trucking	Non-local Trucking
Garbage Collection	Printing/Copying	Printing/Copying
Non-local Trucking	Grounds Maint.	Grounds Maint.
Indoor Painting	Packing & Crating	Packing & Crating
Automotive Repair		
Printing Copying	Cluster Two - 1.84	Cluster Two - 1.84
Furniture Repair	Roofing Repair	Roofing Repair
Grounds Maint.	Indoor Painting	Indoor Painting
Television Repair	Automotive Repair	Automotive Repair
Guard Services	Furniture Repair	Furniture Repair
Packing & Crating	Television Repair	Television Repair
	Cluster Three -2.09	Cluster Three- 2.00
	Dining Facility Ops	Dining Facility Ops
	Guard Services	
Cluster Two - 2.92		Cluster Four - 2.19
Legal Consultation	Cluster Four - 2.51	Guard Services
Architectural Design	Computer Maint.	
Computer Maint.		Cluster Five - 2.57
Dentistry Clinics	Cluster Five - 2.63	Computer Maint.
ADPE Programming	Dentistry Clinics	Dentistry Clinics
	Cluster Six - 3.15	Cluster Six - 3.15
	Legal Consultation	Legal Consultation
	Architectural Design	Architectural Design
Cluster Three 3.86	ADPE Programming	ADPE Programming
Biological Research		
Weapons Engineering	Cluster Seven - 3.70	Cluster Seven -3.70
	Biological Research	Biological Research
	Cluster Eight - 4.02	Cluster Eight - 4.02
	Weapons Engineering	Weapons Engineering

The width of range at the three cluster level, as indicated in Table 4-6, was 257.43% (3.81 divided by .148) of the width at the five cluster level. This width of range would clearly make it less likely that services would shift between membership in different categories, thereby providing superior exclusivity. In the opinion of the researcher, however, the three cluster level did not differentiate enough between services to provide any strategic insight. The clusters were too large to provide for any implementation or review of policies, programs, or practices at a particular service "type". More differentiation would be required to yield an "optimal" number of categories.

The width of range at the five cluster level, as indicated in Table 4-6, was 493.33% (.148 divided by .030) larger than the width at the eight cluster level. The difference in exclusivity was substantial. The researcher, however, reviewed the difference between the differentiation of the five and eight cluster levels in detail. The first observation made by the researcher was that the grouping that resulted from Ward's minimum variance at the eight cluster level was, once again, intuitively superior to that of the average linkage method. Ward's method separated the sample services "computer maintenance" and "dentistry clinics" into their own unique clusters while it continued to group "dining

facility operations" and "guard services" together. The average linkage method, conversely, continued to group "computer maintenance" and "dentistry clinics" and separated "dining facility operations" and "guard services" into their own unique clusters. The researcher's "common sense" expectation was that procurement of "dining facility operations" and "guard services" is not as distinctive as that of "computer maintenance" or "dentistry clinics". While the methods otherwise produced identical results, the researcher concluded that Ward's was the superior method at the eight cluster level.

The researcher deemed that the major improvement in differentiation was that Cluster One of the five cluster level, a grouping of 11 services, was divided into two clusters of six and five services at the eight cluster level. Such differentiation could prove useful for implementation or review of policies, programs, or practices of these services. A major problem, however, was that the differentiation between these newly formed clusters was very slight. The difference between the mean scores for all twelve characteristics of Clusters One and Two, as displayed next to their titles in Table 4-7, was .14 (1.84 minus 1.70). An 8.24% (.14 divided by 1.70) increase in scoring of services with values greater than or equal to the mean of Cluster One would make them

migrate to Cluster Two. The composition of new clusters/ categories would be highly uncertain. At the five cluster level, the smallest difference between any two clusters was .33 (2.09 at Cluster One minus 1.76 at Cluster Two with Ward's method). An 18.75% (.33 divided by 1.76) increase in scoring of services in Cluster One with values greater than or equal to its mean of Cluster One would be required to make them migrate to Cluster Two. The composition of Clusters One and Two was more than two-times as stable at the five cluster level. This two-fold difference in exclusivity between the closest clusters, combined with the 493.33% larger total range of width noted above, clearly indicated that grouping at the five cluster level was substantially less sensitive to error than grouping at the eight cluster level.

The researcher compared the differentiation between services at the five cluster level and, as noted in section IV.E.2., had found that the results of Ward's method conformed with the researcher's "a priori" and "a posteriori" perceptions of "common sense". The differentiation between sample services, especially for the more "complex" clusters (those with higher means), was also deemed to be adequate for the implementation or review of procurement policies, programs, or practices. Moreover, as discussed in section IV.C.1., the researcher had concluded that a selection of

between five and seven categories would accommodate the limitations of the human attention span and retrieveal from long-term. The researcher therefore concluded that his measurement of an "optimal" number of categories - that which balanced the goals of exclusivity and differentiation - was accomplished at the five cluster level.

4. Selection of a Preferred Clustering Method

The average linkage and Ward's minimum variance methods produced identical groupings through the first four cluster levels. As discussed in section IV.E.2., however, their results differed markedly at the five cluster level and the researcher concluded that the results of Ward's method conformed with "a priori" and "a posteriori" perceptions of "common sense". As discussed in section IV.E.3.c., the researcher also deemed that Ward's is the superior method at the eight cluster level. The researcher therefore elected to use the results of Ward's method for the categorization of Government procured services. Additionally, the researcher opted to use Ward's method exclusively in an effort to streamline the data collection model.

F. SUMMARY

In this chapter the researcher delineated how the data collection model was used to survey professional evaluations

of the applicability of strategic characteristics to various sample services. Survey response rates, and the consolidation of this input in a mean-value matrix, and production of an "a priori" benchmark classification are described in preparation for cluster analysis. Next, cluster analysis methods are outlined and initial results compared to "a priori" expectations and desirable classification attributes. Methods for selecting the appropriate number of categories, as well as the choice of the five cluster level, are then recounted. Finally, the researcher explains the adoption of the Ward's minimum variance method as a source of specific categories at the five cluster level and for use in future clustering iterations.

Data analysis continues in the next chapter as the researcher uses the Ward's minimum variance method and other measures in an attempt to simplify the classification scheme.

V. SIMPLIFYING THE TAXONOMICAL MODEL

A. INTRODUCTION

This chapter describes the further refinement of the classification model into a functional scheme. The need to simplify the data collection model is substantiated, based on survey feedback. The characteristics are then evaluated according to the relative consistency with which their meanings, as defined and scaled by the researcher, may be ascertained and applied. Next, the characteristics in the data collection model are analyzed with regard to their relative strategic importance. Finally, the characteristics are evaluated according their potential ability to distinguish between services, and the issues of characteristic concordance and redundancy are addressed.

The researcher then uses these measures of characteristic desirability to retain or remove characteristics from the classification model. The model is functionally completed by the development of average and boundary characteristic values, as well as appropriate descriptions, for each category.

B. ANALYSIS OF MODEL CHARACTERISTICS

As discussed in section II.A., classification schemes serve to organize phenomena for future use in systematic investigation and development of theories. A "useful"

classification scheme must organize on a valid basis and be functional for future use. A scheme is not useful if it is too complex or time-consuming.

1. The Need For Model Simplification

Feedback from the data collection phase established that the model required further simplification. As discussed in section IV.B.4., 190 (63.33%) of the surveys were not returned. While the researcher did not presume to ascribe a reason for these non-responses, seven of the 110 respondents (6.36%) provided specific comments about the length of time required to complete the model. The one "negative" respondent who cited a lack of time stated that

> This is a good example of surveying gone berzerk - if this is a test of whether respondees are willing to waste any amount of time on anything sent to them, I just failed the test.

This respondent may have been speaking for many non-respondents who did not return the survey due to its apparent length and complexity. Furthermore, six "positive" respondents who completed the matrix stated that the effort required longer than the 45 minute period cited on the data collection survey instructions. In one case a respondent stated that completing the matrix required three hours. Another respondent made a summary comment that "I will be surprised if you get a lot of response from this complex

questionnaire. Remember the KISS concept." The acronym "KISS" is military jargon for Keep It Simple, Stupid. Although the respondent politely decided to cross out the last "S", the point was well-taken. If possible, the model should be simplified to reduce its complexity.

Reduction of the number of characteristics would simplify the model, since the number of definitions and scales that must be read and the number of scores that must be provided would be reduced. Reduction of the number of characteristics, however, may also reduce the validity of the model if a desirable characteristic is discarded. The researcher therefore used different measures to identify relatively undesirable characteristics and discern the impact of their removal.

2. **The Ascertainability of Characteristics**

Ascertainability is one of several desirable characteristic traits that were identified in section III.C.1.. A high degree of ascertainability allows a user to precisely determine the presence of the characteristic and the degree of that presence. The standard deviation of the data collection scores, as applied by each scorer to the twenty sample services, provided a measure of the ascertainability of the characteristics. A large standard deviation indicates that the degree of the presence of the characteristic, as

defined and scaled, is difficult to determine and the characteristic is relatively "nonascertainable".

The researcher constructed a standard deviation matrix using the same process described in section IV.D.1. for the mean value matrix. The standard deviation matrix is provided in Table 5-1. The researcher used two standard deviation measures to evaluate the ascertainability of the characteristics.

a. Standard Deviations Equalling or Exceeding 1.00

The researcher first sought a method to evaluate the ascertainability of characteristics as independently applied to each sample service. The researcher decided that a standard deviation equalling or exceeding 1.00 would indicate that a characteristic is relatively "unascertainabile" for a particular sample service. Since the characteristics were scaled from 1.00 to 5.00, a standard deviation equalling or exceeding 1.00 would indicate that the characteristic had a variability in application of at least 20% for a sample service. The row at the bottom of Table 5-1 labelled "# S.Dev>1" provides the total number of such services for each characteristic. For example the characteristic **Measurability** had a standard deviation of 1.01 when it was applied to the service "legal consultation", and

the row "# S.Dev>1" has a value of "1" in the column pertaining to **Measurability**.

While the presence of a characteristic may be quite ascertainable for a majority of services, a high standard deviation for several individual services would indicate at a minimum that its definition and scales should be reviewed for possible improvement. In some cases a characteristic may be, by nature, problematic (e.g., counterintuitive or difficult to scale), and therefore should be removed from the model since this deficiency nullifies any potential strategic insight.

b. Mean Standard Deviations for All Services

The ascertainability of characteristics was also measured by calculating the mean of the 21 sample service standard deviations. While measurement of standard deviations equalling or exceeding 1.00 gauged the ascertainability of a characteristic on an exception basis, the mean standard deviation would provide an evaluation at the other extreme by treating all services with equal weight. For example, while the characteristic **Customization** was applied with standard deviations equalling or exceeding 1.00 for five services, the mean of the standard deviations of all service scores was 0.84 for that characteristic. The mean standard deviation for each characteristic is provided in the row labelled "Avg S.Dev" in Table 5 - 1.

TABLE 5-1

STANDARD DEVIATION MATRIX
Source: Researcher's Analysis

SERVICE	C.1	C.2	C.3	C.4	C.5	C.6	C.7	C.8	C.9	C.10	C.11	C.12
BioRsch.	1.23	0.78	0.73	1.09	0.82	1.12	1.08	0.96	1	1	1	0.93
DinFacOps.	0.88	0.58	0.57	1.12	0.77	0.53	1.21	0.86	0.99	0.88	0.78	0.95
Wpns.Engi	0.78	0.73	0.69	0.97	0.71	0.63	0.71	0.65	0.56	0.93	0.83	0.66
Linen Sup	0.66	0.65	0.59	1.2	0.84	0.61	0.93	0.87	0.86	0.77	0.75	0.85
Roof Rep	0.7	0.68	0.61	0.9	0.74	0.46	0.95	0.8	0.74	0.59	0.64	1.48
GarbColle	0.56	0.52	0.45	1.15	0.78	0.46	1.09	0.98	0.9	0.85	0.63	1
NonlocTrk.	0.72	0.62	0.66	1.05	0.72	0.61	0.86	0.82	0.76	0.64	0.84	0.92
IndoPaint	0.72	0.61	0.52	0.98	0.73	0.41	0.78	0.8	0.75	0.51	0.7	1.39
LegalCons	1.17	0.84	0.88	0.79	1.01	1.12	0.92	0.95	0.84	0.93	0.96	1.23
AutoRepai	0.73	0.63	0.71	0.74	0.69	0.65	0.94	0.87	0.82	0.69	0.65	1.17
PrintCopy	0.83	0.59	0.76	0.91	0.66	0.89	0.85	0.82	0.8	0.61	0.93	0.8
FurnitRep	0.7	0.65	0.6	0.97	0.68	0.46	0.75	0.76	0.68	0.76	0.62	1.19
ArchDesig	1.1	0.77	0.74	0.91	0.91	1.01	0.95	0.85	0.98	0.93	0.94	1.12
Grnds.Mai	0.72	0.6	0.48	1.07	0.79	0.45	0.82	0.96	0.99	0.79	0.77	0.9
CompuMain	0.82	0.76	0.76	0.83	0.85	0.93	0.96	0.88	0.98	0.96	0.84	0.99
TV Repair	0.57	0.67	0.74	0.87	0.72	0.64	0.84	0.84	0.77	0.7	0.7	1.07
GuardSvcs.	0.95	0.56	0.54	0.91	0.8	1.19	1.15	0.89	0.89	0.91	0.78	1.17
DentClini	0.91	1.08	0.85	0.94	0.92	0.75	1.11	0.92	0.88	0.87	0.83	1.1
ADPE Prog	1.07	0.83	0.82	0.86	0.89	1.07	0.97	0.86	0.92	0.89	0.99	1.01
Pack&Crat	1.04	0.6	0.55	1.05	0.68	0.63	1.06	0.85	0.87	0.87	0.74	0.99
# S.Dev>1	5	1	0	7	1	5	6	0	1	0	1	8
Avg S.Dev	0.84	0.69	0.66	0.97	0.79	0.73	0.95	0.86	0.85	0.8	0.8	1.05

All numbers are rounded to nearest 1/100th.

KEY:

C.1 = Customization	C.5 = Measurability	C.9 = Negotiation
C.2 = Expertise	C.6 = Confidentiality	C.10 = Competition
C.3 = Complexity	C.7 = Risk to Govt	C.11 = Stability
C.4 = Labor % of Cost	C.8 = Govt Attention	C.12 = Perishability

c. Analysis of Standard Deviation Measures

The researcher's goal at this stage of the data analysis was to guage which characteristics have relatively superior, average, or inferior ascertainability. Formulation of ordinal rankings was considered, whereby characteristics would be sorted 1-12 in one ranking by the number of services with standard deviations equalling or exceeding 1.00 and in another by the mean standard deviations of all service scores. A simple ordinal ranking is a conventional method of data presentation that has been used in this effort. In this case, however, the researcher wished to combine evaluations of extreme measures into one evaluation of ascertainability, and do so in a way that accounted for the similarity between measurement values. The researcher therefore used cluster analysis to categorize the results of the standard deviation measures. As in the case of grouping services by diverse characteristics, cluster analysis could be used to group the characteristics themselves by different measures of their ascertainability. The researcher decided to classify the characteristics into three levels of ascertainability: "superior", "average", and "inferior". With an ordinal ranking, the researcher would have divided the number of characteristics by three and produced groupings with four (12 divided by three) members each, i.e. four "superior"

characteristics. Use of clustering, however, would incorporate statistically sound measures of similarity into the classification of characteristics. The researcher therefore used the values provided in Table 5-1 for "# S.Dev>1" and "Avg S.Dev" to cluster the characteristics.

While the researcher had concluded that Ward's minimum variance method was superior for the classification of services, these methods had not been compared for the classification of characteristics. The researcher therefore used Ward's minimum variance and the average linkage method, but no selection was necessary since the results were identical at the three cluster level. These groupings are displayed in Table 5-2. The sequence of characteristics within each cluster is that used for the data collection model

TABLE 5-2

CLUSTERING THE CHARACTERISTICS BY STANDARD DEVIATIONS EQUALLING OR EXCEEDING 1.00 AND BY MEAN STANDARD DEVIATION
Source: Researcher's Analysis

Cluster One: "Superior"	Cluster Two: "Average"	Cluster Three "Inferior"
Ascertainability	Ascertainability	Ascertainability
Expertise	Customization	Labor % of Cost
Complexity	Confidentiality	Risk to Government
Measurability		Perishability
Government Attention		
Negotiation		
Competition		
Stability		

and is not intended to portray the relative ascertainability within a cluster. This grouping revealed that five characteristics were either moderately or extremely nonascertainable. In the researcher's opinion, the definitions and scales for these "average" and "inferior" characteristics may need to be reviewed in the future in order to provide more consistent and accurate classification. These groupings would also be used by the researcher to evaluate the overall contribution of a characteristic to the classification model.

3. The Strategic Importance of Characteristics

The end-use goal of the classification scheme is to classify services on a strategic basis, so it is critical that model characteristics have strategic relevance. The relative strategic importance of characteristics is difficult to measure "a priori", however, since the presence of such characteristics has not been monitored or analyzed in a scientific manner. The best approximation available is to utilize expert opinion, and all twelve data collection model characteristics had been reviewed by expert panels to ensure that they were strategically significant. In the researcher's opinion, another review would ensure that "a priori" perceptions were fully utilized. A column for ranking Top Three characteristics in order of strategic importance was

therefore included in the data collection matrix (see Appendix C).

a. Formulation of an Unweighted Priority Ranking

The Top Three rankings provided specific feedback on the perceived importance of the various characteristics. Sixty-six procurement professionals ranked them in fully "positive" responses, and the results are provided in Table 5-3. The researcher used two extreme measures to analyze the Top Three survey feedback. One measure was to weigh the importance of each ordinal ranking equally, without regard to whether the characteristic was ranked first, second, or third. All first, second, and third ratings were multiplied by an

TABLE 5-3

FREQUENCY OF TOP THREE CHARACTERISTIC RATINGS
Source: Researcher's Analysis

Characteristic	Rated First	Rated Second	Rated Third	Unweighted Total	Weighted Total
Labor % of Cost	318	258	140	716	2504
Expertise	364	184	102	650	2474
Competition	184	139	174	497	1511
Measurability	43	137	167	347	793
Customization	48	70	71	289	1021
Risk to Govt	76	92	109	277	765
Complexity	49	123	91	263	705
Stability	58	78	121	257	645
Govt Attention	24	111	106	241	559
Negotiation	31	52	84	167	395
Perishability	11	36	119	166	282
Confidentiality	14	40	36	90	226

equal weight of 1 (unweighted), and totals using this measure are provided in the "Unweighted Total" column of Table 5-3. In the case of the characteristic **Labor % of Cost**, for example, the numbers 318, 258, and 140 were summed using normal addition to produce an unweighted total of 716.

b. Formulation of a Weighted Priority Ranking

The second measure was to weight the ratings unequally, with an extreme bias to the "higher" ratings. The researcher assigned weights of 5, 3, and 1, respectively, to the first, second, and third ratings. Totals using this method are provided in the "Weighted Total" column of Table 5-3. In the case of the characteristic **Labor % of Cost**, for example, the number 318 was multiplied by 5, 258 was multiplied by 3, 140 was multiplied by 1, and the resultant numbers were summed to provide a weighted total of 2504.

c. Analysis of Strategic Importance Measures

The researcher's goal at this stage of the data analysis was to guage which characteristics have relatively major, moderate, or minor strategic importance. The researcher decided to use cluster analysis to combine the evaluations of these extreme measures, for essentially the same reasons that are cited in section V.B.2.c. for application of cluster analysis to measurement of ascertainability. The researcher decided to classify the

characteristics into three levels of strategic importance: "major", "moderate", and "minor". The researcher used the values provided in the "Unweighted Total" and "Weighted Total" columns of Table 5-3 to cluster the characteristics.

The researcher used Ward's minimum variance and the average linkage method, but no selection of a preferred method was necessary since the results were identical at the three cluster level. These groupings are displayed in Table 5-4. The order of characteristics within each cluster is that used for the data collection matrix and is not intended to portray relative ascertainability within a cluster. The groupings revealed a remarkable differentiation at the relatively "high" end of the range of strategic importance. Only two

TABLE 5-4

CLUSTERING THE CHARACTERISTICS BY WEIGHTED AND UNWEIGHTED TOP THREE MEASURES
Source: Researcher's Analysis

Cluster One: **"Major"** *Strategic Importance*	Cluster Two: **"Moderate"** *Strategic Importance*	Cluster Three: **"Minor"** *Strategic Importance*
Expertise	Competition	Customization
Labor % of Cost		Complexity
		Measurability
		Confidentiality
		Risk to Government
		Govt Attention
		Negotiation
		Stability
		Perishability

characteristics were clearly of "major" strategic importance, and the researcher deemed that these characteristics should only be removed from the model in the event that other measures provided compelling evidence that they did not contribute to the model. In the researcher's opinion, the large "minor" grouping indicated that withdrawal of any of these characteristics would have approximately equivalent impact, and that other measures would therefore have decisive importance in streamlining the data collection model.

Additionally, the researcher believed that the validity of the Top Three rankings should be placed in context. The researcher had observed that some respondents used the same Top Three characteristics for most of the sample services. While the ordinal ranking of the Top Three would change, it seemed that many respondents had "favorites" that were consistently one of the Top Three. They may have selected these because they were valid, but they may favored characteristics that were either: well-known as common discriminators in the procurement profession; self-explanatory, so their application was more readily ascertainable, or; selected first since they appeared first to the eye. In the case of visual appearance, for instance, it may be common for scorers to scan left-to-right and choose those on the left of the data collection matrix. The

researcher analyzed the unweighted Top Three totals, and constructed an ascending 1-12 ranking. The Top Three characteristics are presented in that order in Table 5-3. If, instead, the characteristics had been listed in the same left-to-right sequence as they were in the data collection matrix, the relative rank of Top Three totals would have been 5-2-7-1-4-12-6-9-10-3-8-11. The sum of these rankings for the first six (left-hand) matrix characteristics is 31, while the sum for the second six (right-hand) is 47. The difference may indicate a left-to-right bias, or merely the bias of the researcher in sequencing them on the data collection matrix. If rankings were influenced by a left-to-right bias, the researcher could not have avoided it since some characteristics would naturally appear "on first glance", regardless of the method of visual presentation. The researcher similarly concluded that there was no method to preclude the fact that some characteristics are more well-known or self-explanatory than others. These possible biases, however, served as examples to caution the researcher to be conservative with the use of Top Three rankings to remove characteristics from the model.

4. The Discrimination and Concordance of Characteristics in the Current Classification Model

The general ability of characteristics to differentiate between services has been measured during the "a priori" interviews, pre-test, and in the selection of Top Three characteristics by data collection model scorers. By selecting strategically important characteristics, these professionals have indicated their perception of which characteristics distinguish the procurement of one service from another. As noted in section V.B.3.c. above, however, these perceptions may be subject to diverse forms of bias. A statistical measure may, at this particular stage of the classification effort, provide another evaluation of the potential ability of characteristics to contribute to the classification model.

The potential discrimination provided by the characteristics may be measured by calculating the mean value of a characteristic for each cluster. In this case, the mean value of a characteristic would be calculated separately for each of the five clusters. If there were a characteristic "...whose mean is almost the same across all clusters", it may be inessential and is a prime candidate for removal from the model (Romesburg, 1984, p. 273). Conversely, if a characteristic shows a large difference in its mean value,

relative to the standard deviations, across two or more clusters it may be an important differentiator for the classification scheme.

The researcher calculated a range based solely on the mean characteristic values for each of the five clusters. The means were calculated by adding the mean values of a characteristic for every service in a cluster, and dividing this sum by the number of services in the cluster. These values are provided in Table 5-5. The range of the mean values was computed by subtracting the lowest mean value of any cluster from the highest mean value of any cluster. For instance, 1.62 is the lowest mean value of any cluster for the characteristic of **Perishability**, and this sum was subtracted from 4.43, the highest mean value to obtain a mean range of 2.81. In this case the highest mean value is from Cluster Five, but the lowest mean value of any cluster is from Cluster Two. For most clusters the extreme mean values were calculated by subtracting the mean of Cluster One from the mean of Cluster Five. The values of this mean range, as well as the ascending (lowest-to-highest) rank of these values, are enumerated for each characteristic in Table 5-5.

This mean range provides an approximate statistical indicator of the *potential* contribution of a characteristic to a clustering. The *actual* contribution of a characteristic to

TABLE 5-5

MEAN OF CHARACTERISTIC VALUES PER CLUSTER
Source: Researcher's Analysis

MEAN VALUE

Cluster	C.1	C.2	C.3	C.4	C.5	C.6	C.7	C.8	C.9	C.10	C.11	C.12
One	1.67	2.05	1.64	3.18	1.76	1.25	2.05	2.11	1.95	1.56	1.66	2.06
Two	2.45	2.08	1.45	3.95	1.93	1.64	3.05	2.7	2.51	1.92	1.89	1.62
Three	2.06	3.63	3.18	3.49	2.8	1.84	2.91	2.99	2.95	2.62	2.32	2.64
Four	2.98	3.96	3.48	4.29	3.52	3.09	3.51	3.44	3.46	2.69	3.14	3.39
Five	4.01	4.47	4.46	3.47	4.39	4.39	4.18	4.32	4.39	3.73	3.95	4.43
MEAN RANGE:	2.34	2.42	3.01	1.11	2.64	3.14	2.14	2.21	2.44	2.17	2.28	2.81
*RANK:	6	7	11	1	9	12	2	4	8	3	5	10

*Sequence of Rank is ascending, from lowest to highest value of **Mean Range**.

All numbers are rounded to nearest 1/100th. Clustering performed with numbers extended to 1/1,000,000.

KEY:
C.1 = Customization
C.2 = Expertise
C.3 = Complexity
C.4 = Labor % of Cost
C.5 = Measurability
C.6 = Confidentiality
C.7 = Risk to Govt
C.8 = Govt Attention
C.9 = Negotiation
C.10 = Competition
C.11 = Stability
C.12 = Perishability

a scheme may be observed by removing a characteristic from a scheme and recording any change in clusters. If the clustering changes, the characteristic was an important contributor in a statistical sense. Its presence was mathematically consequential, insofar as the measures of similarity changed enough to produce a different outcome at a particular cluster level. Conversely, if a characteristic is removed and a clustering does not change, it may be said to be *concordant* in the way in which it discriminates populations (Jardine & Sibson, 1971, p. 171).

While a characteristic may be found to be concordant with others at all clustering levels of a particular population, it may not be inferred that such a characteristic is redundant. The issue of logical redundancy was addressed in section III.C.1., but a detailed discussion of correlation was deferred due to a lack of data. A characteristic "...can be considered redundant only if it is *both* statistically correlated....and concordant with the same attribute or set of attributes." (Jardine & Sibson, 1971, p. 171) It is apparent from inspection of the mean value matrix (Table 4-3) and standard deviation matrix (Table 5-1) that no pair of characteristics is perfectly statistically correlated. It may be possible that a combination of characteristics is perfectly correlated with another, but such an analysis is beyond the

scope of this effort. Even if two characteristics were found to be statistically correlated and concordant within the population of services that has been sampled, they may cease to be so when further populations are considered. Removal of characteristics without a change in clustering should therefore only be construed as indicating as an approximate measure of concordance with the five category model that has been developed thus far in the effort. Future taxonomic efforts will be necessary for valid examination of redundancy among the characteristics of Government procured services.

C. STREAMLINING THE MODEL

As discussed in section V.B.1., simplification of the model is necessary in order to yield a scheme of value to potential users. Reduction of the number of characteristics would serve to streamline the model, since the number of definitions and scales that must be read and the number of scores that must be provided would be reduced. If useful information is discarded, however, elimination of characteristics may reduce the validity of the model. The researcher therefore developed measures of characteristic desirability as discussed in section V.B., and endeavored to apply them in a gradual, controlled methodology.

1. Description of the Characteristic Removal Process

The researcher used the mean range to provide a logical progression of characteristics to test for removal. The mean range indicated characteristics which were less likely to contribute to the model, and its derivation directly from the data provided an unbiased starting point for the characteristic review process. The researcher used the rank of mean value ranges, provided in Table 5-5, as the order in which characteristics would be reviewed. For example, the characteristic **Labor % of Cost** was the first characteristic that was reviewed since its mean range of 1.11 was the lowest of all characteristics, giving it a rank of "1".

Using the mean range sequence, the researcher removed the values of a characteristic from the data input and examined the cluster analysis results. Ward's minimum variance method was used since it had been recognized as a superior procedure during the initial cluster analysis and had been used to produce the current five cluster level model (provided in the Ward's column of Table 4-4). The researcher determined beforehand that a characteristic would not be removed from the model if its absence resulted in new clustering outcomes at the five-category level. For example, if removal of the characteristic **Confidentiality** caused the service "legal consultation" to change to a different

grouping, that characteristic would be retained. While it may not appear to be statistically important according to the mean range rank, or relatively strategically important according to the Top Three rankings, its statistical importance, as signified by the change of groupings, would indicate that it is an important discriminator for "legal consultation". The change in groupings would signify that **Confidentiality**, while not "popular" as a discriminator, is an important discriminator for a service in this model at the five cluster level. In the researchers opinion, removal of a characteristic that changed the groupings would result in the premature disposal of information.

If the constituent groupings did not change at the five cluster level, the researcher then considered the measures of characteristic desirability. Specifically, the researcher first consulted the relative strategic importance of a characteristic as indicated by the clustering of unweighted and weighted Top Three measures (Table 5-3). As discussed in section V.B.3.c., the researcher had decided that members of the cluster of "major" strategic importance were clearly distinguished by this measure. The researcher resolved beforehand that only compelling evidence from all of the rest of the indicators in the process would lead to removal of a characteristic of "major" importance.

Conversely, removal of characteristics from the large number of "moderate" or "minor" importance would be considered as having relatively minor impact. The predisposition would be to remove such characteristics if other measures indicated they could be removed for the sake of streamlining the model.

The final measure of the contribution of a characteristic to the model would be its relative ascertainability, as indicated by the groupings in Table 5-2. If a characteristic were "superior", that grouping would add to any evidence that retention of the characteristic is desirable. If the characteristic were "average", that would not be considered a negative attribute. If the characteristic were retained, however, review of its scales and definitions would be appropriate in order to improve its desirability. The same recommendation holds for characteristics that are retained which have "inferior" ascertainability, except that removal of such characteristics will be considered more favorably. Since the definitions and scales of these characteristics resulted from an interview and review process, it will be presumed that "inferior" characteristics may, by nature, be extremely difficult to define or scale.

If, at the end of this screening process, the measures indicated that a characteristic should be removed from the model, the researcher would remove its input from the model

when the next characteristic was examined. For example, if the removal process indicated that the characteristic **Government Attention** should be withdrawn from the model, the survey data pertaining to **Government Attention** would be removed *prior* to the cluster analysis of the characteristic that was ranked next in ascending order by the mean range measure. At that stage in the review process, if three characteristics had been removed from the model, the clustering would not include the data from those three characteristics and the data of the characteristic to be examined would also be removed. The cluster analysis in such a situation would be based upon a total of eight characteristics, (the original twelve minus three minus one) and the outcome would be examined to see if the constituent services remained in the same groupings at the five cluster level as they had been in the original groupings (see the Ward's minimum variance column of Table 4-[illegible]

2. Selection of Characteristics f[illegible]val

This narrative will descr[illegible] researcher's decisions concerning those character[illegible]s whose removal did not change the five cluster level gro[illegible]s. A summary of the entire iterative process is provided by Table 5-6.

The first removal of input that did not change the groupings was for the characteristic with the third mean range

TABLE 5-6

SUMMARY OF THE PROCESS OF CHARACTERISTIC RETENTION/REMOVAL
Source: Researcher's Analysis

START: CHARACTERISTIC	#1 Temporarily remove: are GROUPINGS UNCHANGED?	#2 Is Charac.: OF "MINOR" IMPORTANCE?	#3 Does Charac. have: "INFERIOR" ASCERTAIN -ABILITY?	#4 CHARAC. *REMOVED?* (BASIS)
Labor % of Cost	No ->			No (#1)
Risk to Government	No ->			No (#1)
Competition	*Yes*---->	-----No ->	(No)	**No** (#2 & #3)
Govt Attention	No ->			No (#1)
Stability	*Yes*---->	-----*Yes*--->	----(No)-->	--*Yes* (#1 & #2)
-*after removal of Stability*-				
Customization	No ->			No (#1)
Expertise	*Yes*---->	-----No ->	(No)	**No** (#2 & #3)
Negotiation	No ->			No (#1)
Measurability	*Yes*---->	-----*Yes*--->	----(No)-->	--*Yes* (#1 & #2)
-*after removal of Measurability*-				
Perishability	No ->			No (#3)
Complexity	*Yes*---->	-----*Yes*--->	----(No)-->	--*Yes* (#1 & #2)
-*after removal of Complexity*-				
Confidentiality	*Yes*---->	-----*Yes*--->	----(No)-->	--*Yes* (#1 & #2)

RETAINED CHARACTERISTICS		**REMOVED CHARACTERISTICS**
Labor % of Cost	Expertise	Stability
Risk to Government	Govt Attention	Measurability
Competition	Negotiation	Complexity
Customization	Perishability	Confidentiality

rank, **Competition**. Since **Competition** was of "moderate" strategic importance as measured by the Top Three groupings, the researcher had to consider its removal with care. The researcher had decided beforehand that the preponderance of indicators would have to indicate that removal was appropriate in order to remove a characteristic of "moderate" importance. The researcher reviewed the ascertainability of **Competition** and observed that it was in the "superior" grouping. The researcher therefore decided to retain the characteristic **Competition**, despite the fact that there was no change at the five cluster level when it was removed.

Stability was the next characteristic whose removal did not change the groupings. **Stability** is of "minor" importance according to the measure of Top Three ratings. Its ascertainability was "superior", but the researcher decided to remove the characteristic due to its relative lack of statistical and strategic importance.

The subsequent removal of input from the characteristic **Expertise** also did not change the groupings. **Expertise**, however, was of "major" importance according to the Top Three groupings. The researcher had decided beforehand that only compelling unity of all other indicators would be sufficient to remove a characteristic of "major" importance. The ascertainability of **Expertise**, however, was "superior" and

the researcher therefore decided to retain this characteristic in the scheme.

The groupings did not change when the input from **Measurability** was subsequently removed. Moreover, **Measurability** was of "minor" importance, and the researcher therefore decided to remove it from the scheme despite its "superior" ascertainability.

The groupings did not change when the input from **Complexity** was then removed. **Complexity** was also of "minor" importance, and the researcher therefore decided to remove it from the scheme despite its "superior" ascertainability.

Finally, the groupings did not change when the input from **Confidentiality** was removed. **Confidentiality** was also of "minor" importance and the researcher therefore removed it despite its "average" ascertainability.

3. Analysis of Characteristic Removal Results

Listings of retained and removed characteristics are provided at the bottom of Table 5-6. A total of four characteristics were removed, one-third of the original twelve. While removal of these characteristics would not reduce the amount of time required to read any basic instructions and become familiar with the model, the reduction of one-third of the characteristics could otherwise be expected to reduce the time required to complete the model by

one-third. Moreover, in the opinion of the researcher, the perceived complexity and intimidating appearance of the model would be reduced via the reduction of the number of characteristics.

The characteristic removal process had produced conservative results, insofar as many lesser characteristics were retained. Five of the retained characteristics were of "minor" importance, and three had "inferior" ascertainability. This outcome, however, was in harmony with the researcher's goal of removing characteristics i· a gradual, controlled manner. Removing a characteristic meant the reduction of potentially valuable information from the classification scheme, and the researcher had intended to remove characteristics conservatively in order to maintain the validity of the model.

The researcher calculated summary values to describe the categories after characteristic removal, and they are provided in Table 5-7. The mean values of each characteristic for the constituent services in each category are listed, and demonstrate how cluster analysis is able to categorize attributes when they do not always covary. For example, the mean value of the characteristic **Labor % of Cost** decreases from Category Two to Category Three.

TABLE 5-7

MEAN VALUES FOR THE CATEGORIES AFTER REMOVAL OF CHARACTERISTICS
Source: Researcher's Analysis

	CATEGORY				
CHARACTERISTIC	ONE	TWO	THREE	FOUR	FIVE
Customization	1.67	2.45	2.06	2.98	4.01
Expertise	2.05	2.08	3.63	3.96	4.47
Labor % of Cost	3.18	3.95	3.49	4.29	3.47
Risk to Government	2.05	3.05	2.91	3.51	4.18
Government Attention	2.11	2.70	2.99	3.44	4.32
Negotiation	1.95	2.51	2.95	3.46	4.39
Competition	1.56	1.92	2.62	2.69	3.73
Perishability	2.06	1.62	2.64	3.39	4.43
RANGE OF CHARACTERISTIC MEANS (lowest-highest)	1.62	2.33	1.57	1.60	0.96
NUMBER OF SERVICES	11	2	2	3	2
CATEGORICAL MEAN	2.08	2.54	2.91	3.47	4.13
HIGHEST MEAN OF SERVICE IN CATEGORY (matrix #)	2.25 (#5)	2.66 (#17)	3.00 (#18)	3.53 (#19)	4.34 (#3)
LOWEST MEAN OF SERVICE IN CATEGORY (matrix #)	1.91 (#4)	2.41 (#2)	2.83 (#15)	3.39 (#9)	3.91 (#1)

The categorical mean, however, ascends from Category One to Category Five.

The range of characteristic means is calculated by subtracting the lowest of the characteristic means for a category from the highest mean from that category. For example, the characteristic mean for Category One is calculated by subtracting 1.56 (for **Competition**) from 3.18 (for **Labor % of Cost**) to yield 1.62. The highest values of this range are from Categories One and Two. The researcher

also observed that in every case where the groupings changed during the removal process (see Table 5-6), the number of services in these categories changed. The constituencies of these categories were the most sensitive to changes in the classification characteristics.

These facts confirmed the efficacy of selecting the five cluster level to produce the categories. Migration between the first two categories occurred for six of the twelve characteristics after they were iteratively removed from the model. If a clustering level with a narrower width of range had been selected, such migration would likely have increased. The large number of services in Category One were not differentiated by the model, but it may be that this is a "natural" grouping. Another explanation for unequal dispersion between categories may simply be that the selection of services was not adequately diversified. As discussed in section III.E.1., the researcher sought to select a variety of services, but the primary concern was to provide services with generally recognizable and self-explanatory titles. Such services may, by virtue of theirconventional nature, tend to be on the "simple" end of the simple-to-complex procurement spectrum.

The outcome of this streamlining process was determined by many subjective decisions. The researcher had

to make decisions at every step in the process regarding which measures were needed, and how they were used in the evaluation of categories and characteristics. Future study with broader populations and different methods will be necessary to gauge the overall validity of the streamlining.

D. RESULTANT TAXONOMICAL SCHEME

In order to facilitate the ease of using the scheme, the researcher produced categorical boundaries and changed the labelling of the categories.

1. Category Boundaries

The ranges of scores for each category is represented by the services with the highest and lowest mean values. These services are listed at the bottom of Table 5-7. The researcher decided to determine the boundaries by finding the mid-point between the highest service mean value of one category and the lowest of the next. For example, the highest service mean of Category One was 2.25, from the service of "roofing repair" (which was the #5 service from the data collection matrix in Appendix C). The lowest service mean of Category Two was 2.41, from the service "Dining Facility Ops". The mid-point between these two values is 2.33. In order to facilitate ease of use, the researcher then adjusted these values slightly to the nearest value that could be evenly

divided by five. The resultant boundary between Categories One and Two was therefore 2.35. The four boundaries between Categories One through Five were, respectively, 2.35, 2.75, 3.20, and 3.70, with ranges starting at any value greater than or equal to the lower boundary and ending at any value less than the next boundary. The range of these categories were thus 1.35 for Category One, .40 for Category Two, .45 for Category Three, .50 for Category Four, and 1.30 for Category Five. The large ranges of Categories One and Five resulted from their being extreme categories in the scheme.

The researcher believed that these ranges or boundary values could be relaxed slightly due to the subjective nature of the scoring of services. The researcher, however, decided not to use ranges of even width. The "natural" groupings of services had occurred at these boundaries. The ranges of Categories Two, Three, and Four could have been standardized at .50 without changing the constituent services of each category, but the researcher deemed that standardization was premature. Future research may reveal that "natural" boundaries are consistently around certain values, in which case adjustments may be made in the interest of standardization to increase the functional appeal of the model.

2. Category Titles

In order to yield a scheme that was more self-explanatory, the researcher opted to use nomenclature to replace the numbers that had thus far described the five categories. Since the classification effort was based on a range from those services that are simple to procure to those that are quite complex, the appropriate titles would describe and distinguish the services in each category across this spectrum.

The researcher examined the titles developed by Wenger for the classification of Government procured goods (Wenger, 1990, p. 87). Wenger had developed five categories that were labeled **Simple**, **Basic**, **Moderate**, **Advanced**, and **Complex**. The researcher first applied these titles to the word service to consider whether, in his opinion, the usage was awkward. For example, the meaning of the words "simple service" seemed to be both readily recognizable and a natural usage in the English language. The words "basic service", "advanced service", and "complex service" also were deemed to be recognizable and natural. The words "moderate service", however, seemed awkward. Upon initial examination, the combination could suggest that a service is not strenuous or intemperate, instead of merely being moderate in procurement complexity. Unlike the combination of "moderate" with "good",

services, as deeds or acts, suggest different connotations than an inanimate object. The words "ordinary", "common", "normal", and "average" were considered as replacements, but the use of these terms would suggest that services in these categories are "ordinary", "common", "normal", or "average". Services in this category, however, may not be ordinary, common, normal, or average except in respect to the complexity of their procurement. These titles could therefore also be misleading, so-called "loaded" words with usage that is burdened with other possible meanings.

The researcher therefore selected the title "intermediate" to describe the middle category of service procurement complexity. While its use, in the opinion of the researcher, is not as habitual in everyday speech as the alternatives described above, the combination "intermediate service" did not suggest any meaning that would mislead a user.

Next, the researcher considered these words as they differed at each stage along the simple-to-complex spectrum. In the researcher's opinion, the words "simple" and "basic" are synonymous in everyday speech. The researcher considered the words "simple service" and "basic service" to see if a user would naturally distinguish between the two combinations, but they also seemed synonymous. The researcher considered

several alternatives to both words, and decided that the word "noncomplex" would be a suitable replacement for the word "simple". While "noncomplex" is not used in everyday speech, it is self-evident that its meaning relates to complexity.

The researcher also deemed that the lack of habitual usage of was actually advantageous since the combinations "noncomplex service" and "basic service" differed mainly due to the novelty of the word "noncomplex". The combination "noncomplex service" did not suggest any other meaning than that of a service which is extremely simple.

Finally, the researcher reviewed the progression **Noncomplex**, **Basic**, **Intermediate**, **Advanced**, and **Complex** as a whole. **Intermediate** provided a neutral, central categorization of services in the middle of the complexity spectrum. **Basic** and **Advanced** bracketed the central category in balanced increments, and **Noncomplex** and **Complex**, as a pair, were balanced extremes. The researcher concluded that these titles described a progression of complexity, as intended, in self-explanatory increments from the relatively simple to the complex.

3. Using the Classification Scheme

The resultant individual service classification scheme is provided in Fig 5-1. This provides a final tool for representing data collected in a taxonomy of Government

procured services. Along with characteristic definitions, scales, and data collection and analysis methods, this scheme constitutes the researcher's proposed method for classifying Government procured services.

Application of this model could begin with a survey that utilized the eight characteristics and their respective

FIGURE 5-1

Individual Service Classification Scheme
Source: Wenger, 1990. p. 85 and Researcher's Analysis

SERVICE: **N=**

CATEGORY

	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization						
Expertise						
Labor % of Cost						
Risk to Government						
Government Attention						
Negotiation						
Competition						
Perishability						

KEY:
+ = mean value for a characteristic is in the upper 1/3 of a category range
0 = mean value for a characteristic is in the middle 1/3 of a category range
- = mean value for a characteristic is in the lower 1/3 of a category range

definitions and scales. Alternate definitions and scales may also be tested by using them in a survey utilizing a different population of scorers. Survey populations may be selected on a nation-wide basis, using the U.S. mail as the primary medium for communication to parallel the effort used for model development. An alternative method would be to select specific Government procurement offices and gather data by using concurrent interviews in addition to data collection matrices.

Respondents could score the same population of sample services, an entirely new population, or a hybrid to evaluate them in relation to the eight characteristics. The data could then be averaged to produce a mean value for a service for each characteristic. As suggested by Wenger, a grid of the sort provided in Figure 5-1 could then be used to display the mean scoring values and classify each service into a particular category (Wenger, 1990, p. 88). The service title would be recorded in the upper left-hand corner and the number of respondents ("N") in the upper right. Mean numeric values would be listed in the "Avg Value" column and a "+", "0", or "-" could be recorded in the category range that applied for the service for each characteristic. A "+" would represent a score that was in the upper one-third of a categorical range, a "0" the middle one-third, and a "-" the lower one-third.

Utilization of these symbols would allow an analyst to quickly scan the results to observe similarities and differences between services. Individual service classification grids for each of the twenty sample services are provided in Appendix D. Specific application of the insights produced by analysis and categorization of Government procured services is discussed in section I.C..

E. SUMMARY

This chapter has described the effort to simplify the taxonomical model. The need for simplification was validated by a review of survey feedback. The streamlining process was then started with an analysis of characteristics in terms of the variability of survey scores and respective priority rankings. Removal of characteristics was then tested by using cluster analysis to gauge the contribution of characteristics in terms of their impact on the grouping of services at the five cluster level. Characteristics were removed in a conservative manner in order to retain potentially valuable analytical capabilities in the model.

Once the appropriate characteristics were selected for retention, categorical ranges and titles were formulated in order to increase the utility of the scheme. A proposed

mechanism and procedures for further classification of Government procured services were also detailed.

The next chapter highlights the resultant research conclusions of this effort and lists recommendations for further research efforts.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. INTRODUCTION

This chapter will present conclusions and recommendations based on the research effort. The primary and subsidiary research questions will be addressed and the chapter will conclude with recommendations for areas of further research.

B. CONCLUSIONS

Several conclusions may be deduced from this research effort.

1. It is possible to classify services by characteristics other than functional area.

The researcher's goal at the outset of this study was to develop a classification scheme for Government procured services that offers strategic insight. Current Government classifications, such as those used for OMB Circular A-76, rely on functional areas of application and fail to concentrate on characteristics which have strategic implications for the procurement of services.

Many of the characteristics selected in this study are novel and may be more difficult to precisely assess than functional areas. They do, however, differentiate between various types of Government procured services in relative

degrees that may gain more precision through continued usage and refinement of the scheme.

2. **The eight characteristics which were developed and selected in this study were Expertise, Labor % of Cost, Risk to the Government, Government Attention, Negotiation, Competition, and Perishability.**

Classifying the sample Government procured services demonstrated that five categories of services exist with a demonstrable degree of internal similarity and inter-category differentiation. Boundaries between categories were approximations, yet most services could be clearly identified applying to one category with limited potential for migration between the categories, as constructed.

3. **Cluster analysis is a useful approach for constructing a classification system for Government procured services.**

Cluster analysis provided a framework for development and analysis of the classification scheme. Different cluster analysis techniques were used to categorize 20 sample services into five groups of services that exhibited similar characteristics. Additionally, cluster analysis provided a mechanism for evaluation of removal of characteristics from the model to produce a simplified scheme.

Researchers should not deduce that cluster analysis is an exact or standardized tool for producing conclusive results. Judgment is required at every stage of application and analysis may therefore produce varied results. This lack

of standardization, however, also demonstrates the flexibility of cluster analysis techniques in application to a variety of phenomena and investigation of different aspects of the same phenomena. The magnitude of potential application of cluster analysis to the classification of Government procured services has not been fully evaluated.

4. Any classification scheme developed for the purpose of categorizing services will be subjective in development and application.

Subjective evaluations were a factor in every phase of model development. "A priori" selection of meaningful characteristics involves evaluations that are difficult to quantify. Development of characteristics' definitions and scales are, especially, areas where "a priori" evaluation is incorporated into subsequent quantitative evaluations of the degree to which a characteristic is "present" in a service. The subjective nature of a characteristic, however, is largely endemic to human selection and evaluation of characteristics of a deed, act or performance whose results are mainly intangible and very perishable.

The consistency of personal evaluations of the presence of characteristics may also now be tested in a systematic manner which utilizes methods for quantifying, aggregating and comparing subjective input. The reviewer and scorers are nonetheless responsible for providing evaluations

that are as objective as circumstances allow. A high degree of knowledge and experience with service acquisition will produce commensurately reliable classification results.

5. **Various methods may be used to develop a classification scheme.**

The model as proposed is based on the results from categorizing 20 sample services. More varied and rigorous testing is required before the scheme may be expected to produce convincing results.

Additionally, the basis for model formulation and methodologies used were not the only ways in which a taxonomy may be generated. Categorization on a common-to-unique basis, for example, may produce different insights from those developed in the simple-to-complex focus of this study.

C. RECOMMENDATIONS

The researcher developed several recommendations as a result of this study.

1. **The model developed as a result of this study should be established as a proposed taxonomy for classifying Government procured services.**

This research has demonstrated a capability to classify services into five different categories using characteristics other than area of functional application. The potential insights from use of the proposed scheme warrant

its consideration as a strategic model for analyzing Government procured services.

Classifying services based on their customization, expertise required, risk to the Government, and other characteristics determined in this study may provide a user with additional insight about the relationship between a service and the way it is or should be procured. Such insights may lead to the refinement of procurement policy, organizational structure, staffing and management, and improved training and education of the acquisition workforce.

2. **The model should be thoroughly tested using various populations of sample services.**

The model development effort used a heterogeneous sampling of services procured by the Government. Tests of the scheme could use the same population to confirm the validity of these results, an entirely new population of services, or a mixture of previously surveyed and new service populations. The validity of model application to a variety of services could be further examined, and refinements made, based on findings of such studies. For example, if a new population of services is sampled, a combination of the new data and that collected during this effort may be large enough to evaluate how many categories are appropriate by using the *k*th-nearest-neighbor method.

Furthermore, sampling may be focused to ascertain the extent to which the model differentiates between services that are relatively homogeneous. This tactic may highlight weaknesses and produce refinements that would be otherwise overlooked with the use of a heterogeneous population.

3. Future research efforts should continue the examination of those characteristics which impact the procurement of services.

The characteristics by which services are classified are critical in determining what categorizations, and what consequent insights, are produced by a classification scheme. Additional research is required to reveal if any crucial characteristics have been completely overlooked or were removed from the final classification scheme. Only rigorous and repetitive testing will produce an accurate assessment of what characteristics are critical for classifying Government procured services on a strategic basis.

Moreover, the definitions and scales of characteristics must be closely examined. Those characteristics which have a high degree of variability in scoring, for example, may be ascertained with greater precision if definitions are more tightly focused or scales differentiated on a more quantitative basis. Present definitions and scales may be compared with alternatives in concurrent studies which gauge the relative variability of

scoring. Greater ascertainability may, in turn, allow for consistent differentiation of the population of Government procured services into supplementary categories.

4. The model should be evaluated when applied to a subpopulation of services rather than the entire population.

Different results may be produced if classification efforts are focused on a subpopulation from one functional area, such as information processing services. Study of services in a common grouping of Standard Industrial Classification Codes, for example, may test the ability of the model to distinguish between specialized services with meaningful results. Another focused approach would be to examine the subpopulation of services that is procured by a single organization, and compare these results with previously surveyed populations. Presumably, an advantage of using an organizational basis for subpopulation and survey selection would be the high degree of scorer knowledge and experience with the procurement of the services. Such studies may be expected to produce commensurately reliable classification results.

D. RESEARCH QUESTIONS

This section provides responses to the research questions posed in Chapter I.

The primary research question this thesis attempted to answer was:

What would be the essential characteristics or features of a taxonomical structure that would classify the services procured by the Federal Government?

The essential features of the proposed taxonomical structure begin with establishing the basis on which the scheme was developed. The next feature would be the delineation of characteristics and their definitions. The final feature would be that the classification scheme produces a categorization of the services.

Subsidiary research questions included:

1. **What steps or procedures are appropriate in developing a classification scheme for Government procured services?**

An end-purpose or reason for classifying services must first be established as a basis for other steps. The most appropriate characteristics of the services, in relation to the purpose, must then be determined. These characteristics must themselves have attributes that allow for the accomplishment of the classificatory purpose. Next, a systematic methodology for the comparison of services with the characteristics must be developed and employed. Then, based on the results of the comparison, categories of services that exhibit the most similarity should be determined. Finally,

any resultant scheme should be tested to ensure that it may serve its original purpose.

2. **What are some of the distinguishable characteristics of the services procured by the Federal Government?**

From a preliminary listing of 34 candidates, 12 characteristics were selected and applied to sample services. These characteristics were Customization, Expertise, Complexity, Labor % of Cost, Measurability, Confidentiality, Risk to the Government, Government Attention, Negotiation, Competition, Stability and Perishability. While not all of these are internal to the production of a service, they are so closely related to the procurement of a service that they may distinguish services into two or more categories. Procurement experts and professionals were able to ascertain the degree of presence of a characteristic for each of the sample services.

3. **Which characteristics of Government procured services are the most important for classification purposes?**

This research effort suggests that eight characteristics may be defined and applied to 20 sample services to produce five categories that differ on a strategic basis. These characteristics were Customization, Expertise, Labor % of Cost, Risk to the Government, Government Attention, Negotiation, Competition, and Perishability. While these characteristics were sufficient for differentiating the selected sample services, further testing will be necessary to

gauge their validity for differentiating different populations of services.

4. **What should be the decision criteria for classifying Government procured services?**

The decision criteria that applied for this study were based on the end-purpose of differentiating between services in a manner that may potentially produce the most strategic insight. This was achieved by developing a scheme that allows for classification of services across a range from simple-to-complex. Other methods exist and may be considered in future studies.

5. **What are the various homogenous categories of services procured by the Government?**

The research identified five categories of services that exhibited the presence of certain characteristics to relatively different degrees. These categories consisted of "Noncomplex", "Basic", "Intermediate", "Advanced" and "Complex" services. These groups are not entirely homogenous, but the differences exhibited by the services allowed for a determination that a measurable degree of differentiation did exist between the categories.

6. **In what areas of Government procurement will this classification scheme be most useful?**

Interviews with experts in the procurement profession indicated that the greatest potential application lies in the area of management of procurement organizations.

Specifically, the scheme could be used to evaluate staffing and distribution of buying functions related to the procurement of various types of services. If a sufficient number of services are categorized, a secondary benefit would be to provide a test "market", or category, for the implementation of new procedures or policies. The characteristics of certain categories may indicate that they are either most promising or least likely to provide a successful environment for the implementation of new service procurement policies. Depending on the strategy of the implementing organization, a "market" may be selected due to likely success or, conversely, its presumed ability to clearly and rapidly identify any shortcomings in the new policy. Identification of the entire range of service categories may also allow for the testing of policies or hypotheses on a limited selection of services that is nonetheless fundamentally representative of the universe of Government procured services.

7. **What would a taxonomical structure for classifying Government procured services consist of?**

The structure that resulted from this research effort involved three essential elements. First would be the eight characteristics as defined and scaled. Next would be the services versus characteristics scoring matrix. Third, categorical ranges and boundaries provide for the

categorical ranges and boundaries provide for the classification of a service into one of five categories.

E. RECOMMENDATIONS FOR FURTHER RESEARCH

The following recommendations fall outside of the present research area but may have a significant influence on procurement classification efforts.

1. An entirely different methodology may be used to construct a classification scheme.

The scheme proposed in this study was based primarily on the results from cluster analysis. Cluster analysis, however, is only one of several methods that may be used to construct classification systems. A recommendation would be to determine if it is possible to use a decision tree approach by concentrating on characteristics of one "dimension" during each iterative classification step. Another approach may be to obtain some degree of expert agreement on what an "a priori" model would be. Possibly through the use of delphi techniques, characteristics and ultimately the structure of the scheme could be developed. Results from any of these approaches may then be compared with the scheme from this study.

2. **The taxonomy of Government procured services may be tested in tandem with a scheme for classifying goods on "hybrid" populations.**

To scope this thesis, goods were purposely omitted. There are, however, very few "pure" services, and some products such as construction services may be considered as a "hybrid" of the characteristics of goods and services. Joint application of taxonomies of services and goods to a "hybrid" such as construction services, for instance, may provide strategically meaningful insights concerning the commonality of the procurement of construction with the procurement of goods and services. Joint use may also result in the eventual production of a "master" classification scheme, which may apply to all procurements performed by the Government.

F. SUMMARY

The conclusions of this study, as well as recommendations that were directly associated with the research effort, were presented in the chapter. Answers to the primary and subsidiary questions were also provided. The chapter closed with enumeration of recommendations that fall outside of the present research area but may have a significant influence on future procurement classification efforts.

LIST OF REFERENCES

Carper, William B., and William E. Snizek, "The Nature and Types of Organizational Taxonomies: An Overview," Academy of Management Review, 5:1 (1980): 65-75.

"Congratulations to New Certified Professionals," Contract Management, (June 1991): 55-59.

Dunn, G. and Everitt, B.S., An Introduction to Mathematical Taxonomy, Cambridge University Press, 1982.

Eysenck, Michael W., Human Memory: Theory, Research and Individual Differences, Oxford, England: Pergamon Press, 1977.

Fleishman, Edwin A. and Marilyn K. Quaintance, Taxonomies of Human Performance: The Description of Human Tasks, New York: Academic Press, 1984.

Fowler, Clark D., Lieutenant, USCG, Development of a Procurement Task Classification Scheme, Master's Thesis, Naval Postgraduate School, Monterey, California, December 1987.

Hunt, Shelby D., Marketing Theory, Homewood, Illinois: Richard D. Irwin, 1983.

Interview between Mr. John J. Mulhern, President, Spring Garden College, Philadelphia, Pennsylvania, and the researcher, 20 March 1991.

Jardine, Nicholas and Robin Sibson, Mathematical Taxonomy, London, England: John Wiley & Sons, 1971.

Lovelock, Christopher H., "Classifying Services to Gain Strategic Marketing Insights," Journal of Marketing, (Summer 1983): 9-20.

Margulis, Lynn and Karlene V. Schwartz, Five Kingdoms, San Francisco, California: W. H. Freeman, 1982.

McKelvey, Bill., Organizational Systematics, Berkeley, California: University of California Press, 1982.

Miller, George A., "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 63 (1956) 90.

Mills, Peter K. and Newton Margulies, "Toward a Typology of Service Organizations," Academy of Management Review, 5:2 (1980): 255-265.

Page, Asa H., Lieutenant, SC, USN, A Taxonomic Approach to Contracting Officer Tasking. Master's Thesis, Naval Postgraduate School, Monterey, California, December 1989.

Park, Steven A., Lieutenant Commander, SC, USN, The Possibility of A Contracting Science, Master's Thesis, Naval Postgraduate School, Monterey, California, December 1986.

Public Law 89-306, 79 Stat. 1127 (Codified as amended 40 U.S.C. 759), Brooks Act, 1982.

Romesburg, H. Charles, Cluster Analysis for Researchers, Belmont, CA: Lifetime Learning Publications, 1984.

SAS User's Guide: Statistics, Version 5 Edition, Cary, North Carolina: SAS Institute, 1985.

Sobczack, Thomas V., "Industrial Classification: Its Role In Increased Productivity," (Unpublished Report for Waldes Kohinoor, Inc. 1978), DLSIE, LD 042331A, microfiche.

Sokal, Robert R., "Classification: Purposes, Principles, Progress, Prospects," Science, 185:4157 (27 September 1974): 1115-1123.

Sokal, Robert R. and P. H. A. Sneath, Principles of Numerical Taxonomy, San Francisco, California: W. H. Freeman, 1963.

Sweeney, Richard F., Lieutenant, SC, USN, A Classification and Analysis of Contracting Literature, Master's Thesis, Naval Postgraduate School, Monterey, California, December 1989.

Thomas, Dan R. E., "Strategy is Different in Service Businesses," Harvard Business Review, (July-August) 1978.

U.S. Department of Defense, Deputy Assistant Secretary of Defense (Procurement), Competition Advocate Listing, Fifth Edition, NAVSO P-3668, 18 October, 1990.

U.S. Department of Defense, Department of Defense 4100.33-INV, "Commercial Activities Inventory Report and Five Year Review Schedule FY 1987," Washington, D.C.: Government Printing Office 1988.

U.S. Department of Defense, Defense Logistics Agency, Federal Supply Classification Cataloging Handbook H2-1, Washington, D.C.: U.S. Government Printing Office, January 1989.

U.S. Department of Defense, General Services Administration, National Aeronautics and Space Administration, Federal Acquisition Regulation, Washington, D.C.: Government Printing Office, 25 May 1990.

U.S. Department of Defense, Office of the Assistant Secretary of the Navy (Research, Development and Acquisition), Deputy for Acquisition Policy, Integrity, and Accountability, Department of Defense Procurement Conference Directory, "Defense Procurement in a Dynamic World", Williamsburg, Virgina: March 19-22 1991.

U.S. Department of Labor, Employment Standards Administration, Wage and Hour Division, Service Contract Act Directory of Occupations, Second Edition, July 1986.

U.S. Executive Office of the President, Office of Management and Budget, Office of Management and Budget Circular No. A-76 (Revised), "Performance of Commercial Activities," Washington, D.C.: Government Printing Office, August 1983.

U.S. Executive Office of the President, Office of Management and Budget, Standard Industrial Classification Manual, Washington, D.C.: U.S. Government Printing Office, 1987.

U.S. General Services Administration, Office of Information Resources Management, Federal Information Resources Management Regulation, 1984, with Amendments through December 1989.

Welsh, Bill, "CM Interview: Dr. Allan V. Burman," Contract Management, March 1991.

Wenger, Brian L., Lieutenant, SC, USN, A Taxonomical Structure for Classifying the Goods Purchased by the Federal Government, Master's Thesis, Naval Postgraduate School, Monterey, California, December 1990.

1990-1991 NCMA CPCM Directory, Vienna, Virginia: National Contract Management Association, 1990.

APPENDIX A

List of Expert Panel Members

Note: Asterisk (*) indicates members were interviewed in-person instead of by telephone. Double Asterisk (**) indicates group, in-person interviews conducted with at least two people in the member's procurement organization.

Billings, Jay, Ph.D., Instructor, Defense Systems Management College, Huntsville, Alabama

Coates, Elinor Sue, Instructor & Consultant, University of California extension, Berkeley, California *

Duvall, Phillip, Lieutenant Colonel, Directorate of Contracting, Sacramento Air Logistics Center, McClellan Air Force Base, Sacramento, California **

Goodwin, Janice C., Chief, Services Contracting Branch, Directorate of Contracting, Sacramento Air Logistics Center, McClellan Air Force Base, Sacramento, California **

Hampton, Richard J., Colonel, United States Air Force, Directorate of Contracting, Andrews Air Force Base, Morningside, Maryland

Haugh, Leroy J., Vice-President, Aerospace Industry Association (AIA), McLean, Virginia

Hearn, Emmett E., Former Supervisor of Contracts, Lockheed Missile and Space Corporation, Sunnyvale, California, Instructor & Consultant, University of California extension, Berkeley, California *

Macfarlan, W. Gregor, Vice-President, Harbridge House, Inc., Alexandria, Virginia

McCarthy, Patrick J., Lieutenant Commander, United States Navy, Section Head, Acquisitions Systems Section, Naval Sea Systems Command, Washington, District of Columbia

Osborne, Otis T., Director of Contracting, Sacramento Army Depot, Sacramento, California **

Pinkerton, Richard L., Ph.D., Professor, California State University, Fresno, California

Socik, Robert C., Major, United States Air Force, Chief, Operational Contracting Division, Beale Air Force Base, Marysville, California **

Sowle, Donald E., Former Director of the Office of Federal Procurement Policy (OFPP), McLean, Virginia

Trimble, Robert, Former Vice-President of Contracts, Martin-Marietta Corporation, Gaithersburg, Maryland

Zemansky, Stanley D., Retired Director of Purchasing for the city of Baltimore, Baltimore, Maryland

---

APPENDIX B

Contents of Background Package Sent to Expert Panel
Source: Researcher's analysis and Wenger, 1990, p. 106-111. Please note: original package used left, right, and bottom margins of one inch, and spacing has therefore been altered and condensed. The list of preliminary characteristics is provided in Table 3-6, and is excluded in the interest of economy. Written contents are otherwise identical.

SMC #2611
Naval Postgraduate School
Monterey, CA 93943

_____ [name],

In accordance with our phone conversation of _____ [date], this information packet should acquaint you with my effort to develop a classification scheme for Government purchased services. My name is Captain Scott T. Allen and I am working towards obtaining a Masters Degree in Acquisition and Contract Management at the Naval Postgraduate School in Monterey, California. I am using an expert-panel approach to researching my thesis, and would like to solicit your views on the subject of classification of Government procured services.

My thesis is entitled "A Taxonomical Structure for Classifying the Services Purchased by the Federal Government". My objective is to develop a scheme for classifying services on a basis that offers the best strategic insights. In other words, which characteristics of the services, acquisition process, and the procurement environment provide the most information for the purposes of defining contracting policies and methods? Enclosure (1) provides additional information about the potential uses of this scheme along with some necessary principles for classification schemes.

One of the important steps in developing a classification structure is the generation of the characteristics by which the services are judged. This is where I need your help. I would like your feedback on my approach to classifying Government purchased services, and, if you agree with it, which characteristics are the most significant.

Therefore, I would like to telephone you during the week of _____ [date] and conduct an interview. During the interview, I expect to ask the questions listed in enclosure (2). The questions are aimed at defining those characteristics which

most clearly differentiate services into homogenous categories.

I have listed the attributes the characteristics must possess in enclosure (3), along with several preliminary (candidate) characteristics. These characteristics are based on a consolidation of my literature research, interviews, and qualitative judgment and are by no means exhaustive. The list is intended to stimulate your thinking and serve as a common point of reference during interviews. The order of characteristics does not coincide with any presumed order of importance, nor is the grouping necessarily logical. I would like to discuss these characteristics with you in order to narrow the list or add to it, as appropriate. Ultimately, I would like to end up with a workable number of characteristics, in the range of 10 to 15.

Your extensive contracting background and knowledge will be invaluable to me as I develop ideas into a functional classification structure. I look forward to talking with you and incorporating your expertise in an effort to advance the science of Government contract management.

Sincerely,

SCOTT T. ALLEN
Captain
United States Marine Corps

Enclosed:

(1) Government Services Classification Scheme: Objective, Conceptual Basis, Uses, and Principles

(2) Interview Questions to Determine Appropriate Characteristics for a Services Classification Scheme

(3) Attributes Each Characteristic Must Possess, and Preliminary (Candidate) Characteristics

GOVERNMENT SERVICES CLASSIFICATION SCHEME: OBJECTIVE, CONCEPTUAL BASIS, USES, AND PRINCIPLES

MAIN OBJECTIVE:
Develop a Government purchased services classification scheme on a basis other than that provided by functional descriptions, such as Commercial Activities titles accumulated under OMB Circular A-76, or the Service Contract Act Directory of Occupations, or current legal classification.

CONCEPTUAL BASIS FOR PROPOSED SCHEME:
Use of an appropriate scheme should enable persons to scrutinize and define contracting policies and methods in light of their strategic implications. To determine which characteristics are most appropriate for this scheme, the classifier could ask him/herself the following general question:

Which characteristics of services, their acquisition environment, and acquisition process offer the greatest strategic insights for the purpose of defining contracting policies and methods?

SPECIFIC USES:
Specifically, such a classification scheme could be used:

a. For the purposes of determining the appropriate contract instrument to use. The structure should allow for a better relationship between service and contract instrument. The new fixed-price award-fee contract type, for instance, may have an optimal application to one category (taxon) of services, and there may be "borderline" categories where different contract types should be considered.

b. For the purposes of developing and utilizing new methods of contract administration and organization. Classification may provide a target "market" on which to test new methods. The Government way wish to restrict competitive contracting offices, for example, to purchasing certain categories of services.

c. For the purposes of dividing omnibus, "umbrella" contracts for services into categories which may be properly grouped together for contract acquisition and administration.

d. For the purposes of highlighting those categories of services which require less statutory and regulatory oversight during contract acquisition or administration.

e. For the purposes of determining appropriate competitive environment elements, such as design competition versus price competition. Enclosure (1)

FOUR PRINCIPLES NECESSARY FOR A SUCCESSFUL CLASSIFICATION SCHEME:

1. The classification scheme must adequately specify the phenomenon to be classified, and must serve a purpose (end-use goal).

2. The classification scheme must adequately specify the properties or characteristics that will be used in classifying.

3. The classification scheme must have categories that are mutually exclusive, i.e., any item can be classified only in one place.

4. The classification scheme must have categories that are collectively exhaustive: it must be capable of defining all existing items needed to meet end-use goals, and; it must be able to accept new items as defined without violating any principle given herein or causing the generation of a new classification system. Enclosure (1)

INTERVIEW QUESTIONS TO DETERMINE APPROPRIATE CHARACTERISTICS FOR A SERVICES CLASSIFICATION SCHEME

1. What are some of the characteristics that distinguish services purchased by the Federal Government?

2. Which properties or characteristics of the services are the most important for classificatory purposes?

3. What should be the decision criteria for classifying Government purchased services?

4. What are the various homogenous categories of services purchased by the Government?

5. Which classes or categories of services are the most meaningful for classification and research?

6. In what specific areas of Government procurement will this classification scheme be useful? (Enclosure 2)

ATTRIBUTES EACH CHARACTERISTIC MUST POSSESS

1. Differentiation (of at least 2 classes of services).
2. Concomitance (must be exclusive).
3. Relevance (to end-use goal).
4. Ascertainability (understandable to the user).
5. Permanence (definable and unchangeable so long as the end-use goal is unchanged).
6. Consistency (fixed meaning, readily adhered to).

(Enclosure 3)

APPENDIX C

Data Collection Survey Model Package

Source: Researcher's analysis. Please note: 1 inch left, right, and bottom margins were used for survey. Spacing therefore differs from original package but written contents are otherwise identical.

Naval Postgraduate School
Monterey, CA 93943-5000
408/372-2755 (hm)

My name is Scott Thomas Allen, Captain, USMC, and I am a student in contract management at the U.S. Naval Postgraduate School. I earnestly need your assistance in an effort to develop a classification scheme for Government procured services. Briefly:

• A Preliminary Classification Model (enclosed separately) has been developed into a Matrix based on the advice of an expert panel of Government acquisition personnel, academics, and consultants. Its end-use goal is to classify services strategically, in order to study current Government policies and potential modifications.

• This Matrix needs to be tested and refined. It would be of TREMENDOUS help if, *based on your expertise in Government procurement, you would spend 20 to 45 minutes to fill the Matrix out and mail it (by October 11th) in the enclosed envelope.* (If you do not have the time to fill it out, and know a procurement professional who does, *please* forward this package to that person.)

• The Matrix contains a list of twenty services, selected from Standard Industrial Classification (SIC) codes, as well as twelve characteristics with which to grade them. By grading each service with those characteristics, and listing your Top Three Characteristics, you will enable me to run a computer comparison to select an optimal list of characteristics. If you choose to assist in this effort, the following procedure is suggested:

(1) Read the definition (attached) of the first characteristic;

(2) Grade each service (1-5) using the scale that follows the characteristic's definition. Please note - *scales should be read closely since some may appear to be counter -intuitive;*

(3) Repeat steps (1) & (2) for each of the twelve characteristics;

(4) Write your Top Three Characteristics (in order of strategic importance) on the right side of the Matrix for each service.

If you wish to provide comments on characteristic definitions or scales, please write them on the right side or back of the Matrix.

Your input will be used to develop a taxonomy (classification) that will increase the body of knowledge of Government contract management. Your assistance would also be invaluable to me, personally, and in any event I would like to sincerely thank you for taking the time to consider this effort.

Very Respectfully,

S. T. Allen

CHARACTERISTIC DEFINITIONS & SCALES

The following characteristic definitions, and their associated scales, are designed to classify services on a strategic range, from the relatively simple to the complex.

1. **Customization** is the degree to which the production of a service is modified from standard commercial practice to conform with a buyer's unique specifications. All services are modified to some degree in consideration of circumstances unique to each customer, but they will differ on the magnitude to which important procedures, or the entire service process, are exceptionally customized for a buyer. In general, a greater degree of customization will increase the amount of buyer attention, and contract cost, necessary to ensure successful service performance.

   Scale
   1 - No customization
   2 - Customization does not substantively alter service production
   3 - Customization substantively alters a few important elements of service production
   4 - Customization substantively alters the bulk of important elements of service production
   5 - The service is produced exclusively for the Government

2. **Expertise** is the degree of professional certification, skill, and experience required of the principal service production personnel to produce a service at an acceptable quality level. Higher levels of required expertise will usually increase the difficulty of evaluating service performance, as well as the extent to which a buyer should validate the qualifications of service provider personnel.

   Scale
   1 - No expertise needed by principal service production personnel
   2 - Expertise needed requires brief or inexpensive training/qualification
   3 - Expertise needed requires moderately lengthy or moderately expensive training/qualification
   4 - Expertise needed requires very lengthy or very expensive training/qualification
   5 - Expertise needed requires extremely lengthy or extremely costly training/qualification

3. **Complexity** is the degree of technical complexity of techniques or equipment used in the scope of service production. Typically, a high degree of technical complexity will require that a buyer devote substantial attention to evaluating the skill level or equipment required to produce a service, as well as evaluating potential providers for those capabilities.

   Scale
   1 - Technical complexity is rudimentary
   2 - Technical complexity is modest
   3 - Technical complexity is sophisticated
   4 - Technical complexity is advanced
   5 - Technical complexity is on the frontier of human knowledge and capabilities

4. **Labor Percentage of Cost** is the degree to which total service cost is expended on provider labor (as opposed to material and equipment). Buyer validation of provider qualifications, especially in the realm of financing, should be affected by the proportion of labor to material and equipment required to perform a service.

   Scale
   1 - A modest amount of total service cost is expended on labor
   2 - A moderate amount of total service cost is expended on labor
   3 - The bulk of total service cost is expended on labor
   4 - The vast preponderance of total service cost is expended on labor
   5 - Almost all of total service cost is expended on labor

5. **Measurability** is the degree of effort necessary to describe and measure acceptable service performance. While performance of some services is obvious and readily measured, others may necessitate extensive description and detailed review by a buyer to determine if service performance satisfies buyer requirements.

   Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
   1 - Description and measurement of acceptable service performance is obvious and almost effortless
   2 - Description and measurement of acceptable service performance is uncomplicated
   3 - Description and measurement of acceptable service performance is moderately difficult
   4 - Description and measurement of acceptable service performance is quite complex
   5 - Description and measurement of acceptable service performance is profoundly perplexing and intricate

6. **Confidentiality** is the degree to which release of information produced by, or required to produce, a service may be detrimental to either the buyer or service provider. The magnitude of potential damage, whether it be financial, competitive, related to reputation, or to national security, from a release of service information determines the level of service confidentiality. A high grade of confidentiality should necessitate extensive buyer validation of provider qualifications for controlling confidential information.

   Scale
   1 - Release of service production information is not at all potentially detrimental to the provider or Government
   2 - Release of service production information would potentially cause inconsequential damage to the provider or Government
   3 - Release of service production information would potentially cause notable damage to the provider or Government
   4 - Release of service production information would potentially cause extensive damage to the provider or Government
   5 - Release of service production information would potentially cause enormous damage to the provider or Government

7. **Risk to the Government** is the likelihood and magnitude of potential harm to the Government that would result if a service is not completed in accordance with cost, schedule, or performance specifications. Buyer attention should increase throughout the entire procurement process as the degree of risk to the Government escalates.

   Scale
   1 - The likelihood and magnitude of potential harm to the Government due to service performance failure is insignificant
   2 - The likelihood and magnitude of potential harm to the Government due to service performance failure is slight
   3 - The likelihood and magnitude of potential harm to the Government due to service performance failure is modest
   4 - The likelihood and magnitude of potential harm to the Government due to service performance failure is substantial
   5 - The likelihood and magnitude of potential harm to the Government due to service performance failure is enormous

8. **Buyer Attention** is the degree of time and effort that buyer personnel typically dedicate to procuring a service. Personnel allocation, work assignments, and other buyer organization plans and policies should vary with the distinctive degree of buyer attention customarily required by different types of services.

   Scale
   1 - Service procurement requires inconsequential time and effort from buyer personnel
   2 - Service procurement requires minor time and effort from buyer personnel
   3 - Service procurement requires moderate time and effort from buyer personnel
   4 - Service procurement requires considerable time and effort from buyer personnel
   5 - Service procurement requires extraordinary time and effort from buyer personnel

9. **Negotiation** is the degree to which price, schedule, and performance criteria are discussed and adjusted by the buyer and potential service providers during the service procurement process. More negotiation will generally require a longer and more detailed procurement effort.

   Scale
   1 - There is no negotiation between buyer and potential providers during the service procurement process
   2 - Negotiation is insignificant between buyer and potential providers during the service procurement process
   3 - Negotiation is meaningful between buyer and potential providers during the service procurement process
   4 - Negotiation is extensive between buyer and potential providers during the service procurement process
   5 - Negotiation is critical and comprehensive between buyer and potential providers during the service procurement process

10. **Competition** is the degree to which multiple, autonomous providers are willing and able to produce a service. Typically, the intensity of competition will influence buyer selection of contract type, as well as the extent to which price is the dominant source-selection factor.

   Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
   1 - Numerous autonomous providers are willing and able to produce the service and are very aggressive in their willingness to do so
   2 - It is quite easy to find several providers who are willing and able to produce the service
   3 - It is uncomplicated to find a few autonomous providers who are willing and able to produce the service
   4 - It is difficult to find a few autonomous providers who are willing and able to produce the service
   5 - It is extremely difficult to find a provider willing and able to produce the service

11. **Stability** is the degree to which important schedule and performance criteria of a service remain the same over a period of time. A more stable service will typically require less attention on the part of the buyer.

   Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
   1 - Any alteration to schedule or performance criteria is, at most, trivial for extremely lengthy periods of time
   2 - Important schedule or performance criteria seldom undergo significant alteration
   3 - Important schedule or performance criteria infrequently undergo significant alteration
   4 - Important schedule or performance criteria frequently undergo significant alteration
   5 - Important schedule or performance criteria almost constantly undergo significant alteration

12. **Perishability** is the length of time that the product of service performance is beneficial to, or consumed by, the buyer organization. A service with a relatively high degree of perishability will be consumed almost instantaneously, while the product of other services may provide benefits for many years.

   Scale *NOTE: SCALE MAY APPEAR TO BE COUNTER-INTUITIVE*
   1 - The period of benefit/consumption is immediate
   2 - The period of benefit/consumption is brief
   3 - The period of benefit/consumption is moderate
   4 - The period of benefit/consumption is lengthy
   5 - The period of benefit/consumption is extremely lengthy

THANK YOU FOR USING THESE DEFINITIONS AND SCALES!

MATRIX-PRELIMINARY CLASSIFICATION MODEL

PLEASE PROVIDE:

Your name: _____

Title: _____

_____

Organization: _____

_____

PLEASE GRADE EACH SERVICE 1, 2, 3, 4, OR 5 USING SCALES & DEFINITIONS PROVIDED FOR EACH OF THE 12 CHARACTERISTICS

SERVICE \ CHARACTERISTICS	1. CUSTOMER SERVICE	2. LEADERSHIP	3. TECHNICAL	4. INTERNAL	5. TECHNOLOGY	6. CONTROL	7. TOTAL	8. NON-LOCAL	9. VOLUME	10. OPERATIONS	11. TOTAL	12. PERSONNEL	TOP THREE Characteristics in order of Strategic Importance	Comments
1.Biological Research														
2.Dining Facility Ops														
3.Weapons Engineering														
4. Linen Supply														
5. Roofing Repair														
6. Garbage Collection														
7. Non-local Trucking														
8. Indoor Painting														
9. Legal Consultation														
10. Automotive Repair														
11. Printing/Copying														
12. Furniture Repair														
13. Architect Design														
14. Grounds Maint.														
15. Computer Maint.														
16. Television Repair														
17. Guard Services														
18. Dentistry Clinics														
19. ADPE Programming														
20. Packing & Crating														

PLEASE MAIL MATRIX IN ENCLOSED ENVELOPE BY OCTOBER 11th.
THANK YOU FOR YOUR ASSISTANCE!

APPENDIX D

Individual Service Classification of the Twenty Sample Services

Source: Wenger, 1990, p. 85 and Researcher's analysis. Please note: values are from mean value matrix (Table 4-3). The following key applies to all of the individual service classifications. For further explanation see section V.D.3..

KEY:
+ = mean value for a characteristic is in the upper 1/3 of a category range
0 = mean value for a characteristic is in the middle 1/3 of a category range
- = mean value for a characteristic is in the lower 1/3 of a category range

SERVICE: Bio. Research N= 85

		CATEGORY				
	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	3.35				-	
Expertise	4.48					0
Labor % of Cost	3.53				+	
Risk to Government	3.81					-
Government Attention	4.01					-
Negotiation	4.05					-
Competition	3.73					-
Perishability	4.34					0

SERVICE: DiningFacOps. N= 85

	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
		CATEGORY				
Customization	2.36		-			
Expertise	2.12	+				
Labor % of Cost	3.51				0	
Risk to Government	2.61		+			
Government Attention	2.82			-		
Negotiation	2.53		0			
Competition	1.84	0				
Perishability	1.49	0				

SERVICE: Weapons Eng. N= 85

	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
		CATEGORY				
Customization	4.66					+
Expertise	4.46					0
Labor % of Cost	3.41				0	
Risk to Government	4.55					0
Government Attention	4.62					+
Negotiation	4.73					+
Competition	3.73					-
Perishability	4.52					0

SERVICE: Linen Supply N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.56	0				
Expertise	1.62	0				
Labor % of Cost	3.08			+		
Risk to Government	1.95	+				
Government Attention	2.02	+				
Negotiation	1.92	+				
Competition	1.56	0				
Perishability	1.53	0				

SERVICE: Roofing Repair N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.64	0				
Expertise	2.24	+				
Labor % of Cost	2.96			0		
Risk to Government	2.45		-			
Government Attention	2.31	+				
Negotiation	1.98	+				
Competition	1.36	-				
Perishability	3.11			+		

SERVICE: Garbage Coll. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.41	-				
Expertise	1.41	-				
Labor % of Cost	3.39				0	
Risk to Government	2.15	+				
Government Attention	2.18	+				
Negotiation	1.91	+				
Competition	1.72	0				
Perishability	1.52	0				

SERVICE: Non-loc.Truck. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.74	0				
Expertise	2.04	+				
Labor % of Cost	2.89			-		
Risk to Government	2.27	+				
Government Attention	2.15	+				
Negotiation	1.99	+				
Competition	1.54	0				
Perishability	1.62	0				

SERVICE: IndoorPainting N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.71	0				
Expertise	1.87	0				
Labor % of Cost	3.35				-	
Risk to Government	1.75	0				
Government Attention	2.06	+				
Negotiation	1.84	0				
Competition	1.26	-				
Perishability	2.61		+			

SERVICE: Legal Consult. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	2.85			-		
Expertise	4.18					0
Labor % of Cost	4.52					0
Risk to Government	3.61				+	
Government Attention	3.21				-	
Negotiation	3.25				-	
Competition	2.65		+			
Perishability	2.89			-		

SERVICE: Auto Repair N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.64	0				
Expertise	2.71		+			
Labor % of Cost	2.89			-		
Risk to Government	2.33	+				
Government Attention	2.25	+				
Negotiation	2.06	+				
Competition	1.55	0				
Perishability	2.32	+				

SERVICE: Printing/Copy N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.82	0				
Expertise	2.05	+				
Labor % of Cost	2.89			-		
Risk to Government	2.06	+				
Government Attention	1.91	+				
Negotiation	1.81	0				
Competition	1.47	0				
Perishability	1.56	0				

SERVICE: Furniture Rep. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.47	0				
Expertise	2.31	+				
Labor % of Cost	3.36				-	
Risk to Government	1.64	0				
Government Attention	1.88	0				
Negotiation	1.80	0				
Competition	1.65	0				
Perishability	2.41		-			

SERVICE: Arch. Design N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	2.87			-		
Expertise	3.99					-
Labor % of Cost	4.22					0
Risk to Government	3.21				-	
Government Attention	3.45				0	
Negotiation	3.52				+	
Competition	2.67		+			
Perishability	3.84					-

SERVICE: Grounds Maint. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	1.62	0				
Expertise	1.67	0				
Labor % of Cost	3.68				+	
Risk to Government	1.75	0				
Government Attention	2.27	+				
Negotiation	2.12	+				
Competition	1.55	0				
Perishability	1.80	0				

SERVICE: Computer Maint N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	2.12	+				
Expertise	3.22				-	
Labor % of Cost	3.44				0	
Risk to Government	3.26				-	
Government Attention	3.00			0		
Negotiation	2.88			-		
Competition	2.19	+				
Perishability	2.49		0			

SERVICE: TV Repair N= 85

	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
		CATEGORY				
Customization	1.38	-				
Expertise	2.68		+			
Labor % of Cost	3.22				-	
Risk to Government	1.86	0				
Government Attention	2.07	+				
Negotiation	1.95	+				
Competition	1.66	0				
Perishability	2.35		-			

SERVICE: Guard Services N= 85

	Avg Value	Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
		CATEGORY				
Customization	2.53		0			
Expertise	2.04	+				
Labor % of Cost	4.40					0
Risk to Government	3.48				0	
Government Attention	2.58		0			
Negotiation	2.48		-			
Competition	2.01	+				
Perishability	1.75	0				

SERVICE: Dental Clinics N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	2.00	+				
Expertise	4.04					-
Labor % of Cost	3.54				+	
Risk to Government	2.55		0			
Government Attention	2.98			0		
Negotiation	3.02			0		
Competition	3.06			+		
Perishability	2.79			-		

SERVICE: ADPE Program. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	3.21				-	
Expertise	3.72					-
Labor % of Cost	4.13					-
Risk to Government	3.69				+	
Government Attention	3.67				+	
Negotiation	3.61				+	
Competition	2.76			-		
Perishability	3.45				0	

SERVICE: Packing&Crati. N= 85

	Avg Value	CATEGORY Noncomplex 1.00-2.34	Basic 2.35-2.74	Intermediate 2.75-3.19	Advanced 3.20-3.69	Complex 3.70-5.00
Customization	2.34	+				
Expertise	1.94	+				
Labor % of Cost	3.19			+		
Risk to Government	2.29	+				
Government Attention	2.11	+				
Negotiation	2.05	+				
Competition	1.80	0				
Perishability	1.78	0				

INITIAL DISTRIBUTION LIST

1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Defense Logistics Studies Information Exchange US Army Logistics Management Center Fort Lee, VA 23801	2
3.	Library, Code 52 Naval Postgraduate School Monterey, CA 93943-5000	2
4.	Dr. David V. Lamm, Code AS/LT Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943-5000	3
5.	Dr. Kishore Sengupta, Code AS/SE Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Captain Scott T. Allen MCLB, Systems Support Division, IRMD (Code 743) Albany, GA 31704-5000	1
7.	Ms. Elinor Sue Coates 1705 Eleventh Ave. San Francisco, CA 94122	1
8.	Ms. Janice C. Goodwin Services Contracting Branch, Directorate of Contracting Sacramento Air Logistics Center McClellan Air Force Base, CA 95652-5320	1
9.	Mr. Emmett E. Hearn 1270 Covington Rd. Los Altos, CA 94024	1

10. Commandant of the Marine Corps 1
Code TE 06
Headquarters, U.S. Marine Corps
Washington, D.C. 20380-001

END

FILMED

DATE: 3-92

DTIC